

Centre international d'études supérieures en sciences agronomiques

THESE

Pour obtenir le grade de

DOCTEUR DU CENTRE INTERNATIONAL D'ETUDES SUPERIEURES EN SCIENCES AGRONOMIQUES

Ecole doctorale : Sciences des Procédés – Sciences des Aliments
Spécialité : Biochimie, chimie et technologie des aliments

Présentée et soutenue publiquement par

Francis Canon

le 30 Septembre 2010

Contribution de la spectrométrie de masse à l'étude des interactions entre les protéines salivaires riches en proline et les tanins

Directeur de thèse : Dr. Véronique Cheynier

Encadrant : Dr Pascale Sarni-Manchado

JURY

E. de Pauw	Professeur (Université de Liège, Belgique)
J.M. Schmitter	Professeur (Université de Bordeaux)
S. Longhi	Directeur de Recherche (AFMB, Marseille)
O. Dangles	Professeur (Université d'Avignon)
V. Cheynier	Directeur de Recherche (INRA, Montpellier)
P. Sarni-Manchado	Chargée de Recherche (INRA, Montpellier)

Contribution de la spectrométrie de masse à l'étude des interactions entre les protéines salivaires riches en proline et les tanins

Résumé

L'astringence résulte de l'interaction des tanins, polyphénols abondants dans les végétaux, avec les protéines salivaires et plus particulièrement les protéines salivaires riches en proline (PRP), appartenant à la famille des protéines peu structurées. Les tanins participent aux mécanismes de défense des végétaux et présentent des effets antinutritionnels dus à leur capacité à inhiber les enzymes digestives. La synthèse de PRP salivaires constitue un mécanisme d'adaptation à la consommation d'aliments riches en tanins. Ce travail vise à caractériser les complexes établis en solution entre les PRP et les tanins par une approche basée sur la spectrométrie de masse (MS). Pour cela, les protéines salivaires humaines IB5, PRP basique, et II-1, PRP glycosylée, ont été produites par voie hétérologue. Après purification, les deux protéines ont été caractérisées par MS avec une source d'électrobulbion (ESI-MS) et avec une source MALDI (Matrix-Assisted Laser Desorption/Ionisation). L'étude des interactions par ESI-MS a confirmé la présence en solution de complexes non-covalents IB5•tanin et permis de préciser leurs stœchiométries. Des expériences de compétition entre différents tanins et de dissociation des complexes IB5•tanin ont mis en évidence l'influence des principales caractéristiques structurales des tanins sur cette interaction. L'étude structurale des édifices IB5•tanin, par différentes techniques de MS/MS (Collision Induced Dissociation et Electron Capture Dissociation) et par spectrométrie de mobilité ionique couplée à la MS, a mis en évidence la présence de plusieurs sites d'interaction sur IB5 ainsi que des changements conformationnels liés à l'interaction.

Mots clé : astringence, protéines riches en proline, polyphénols, tanins, spectrométrie de masse, spectrométrie de mobilité ionique, interactions non-covalentes, ESI-MS, MALDI-MS, MS/MS, CID, ECD.

Mass spectrometry contribution to the study of interactions occurring between proline rich proteins and tannins.

Abstract

Astringency is an important organoleptic property of plant-based food. It is attributed to interactions of tannins, which are polyphenolic compounds, with salivary proteins and especially proline rich proteins (PRPs), which belong to the group of intrinsically unstructured proteins (IUPs). Tannins play an important part in plant defence mechanisms. Indeed, they have an antinutritional effect as they inhibit digestive enzymes. Production of salivary PRPs is thus an adaptation process to tannin-rich diets. The purpose of this work is to provide a closer look at PRP•tannin supramolecular edifices in solution, using a mass spectrometry (MS) approach. The human salivary proteins IB5, a basic PRP, and II-1, a glycosylated PRP, have been produced by heterologous expression. After purification, both proteins have been characterized by MS using electrospray (ESI) and Matrix-Assisted Laser Desorption Ionisation (MALDI) sources. The study of the interaction between IB5 and model tannins by ESI-MS confirmed the presence of IB5•tannin non-covalent complexes in solution and provided new information on their stoichiometries. Competitive interaction experiments between IB5 and two tannins, along with IB5•tannin complexes dissociation studies revealed the impact of the main tannin chemical features on this interaction. Structural studies performed on IB5•tannin edifices by Collision Induced Dissociation (CID) and Electron Capture Dissociation (ECD) MS/MS experiments and by ion mobility spectrometry coupled with MS showed the presence of several interaction sites on IB5 and conformational changes arising from the interaction.

Keywords : astringency, proline rich proteins, polyphenols, tannins, mass spectrometry, ion mobility spectrometry, non-covalent interactions, ESI-MS, MALDI-MS, MS/MS, CID, ECD.

DISCIPLINE: Biochimie, chimie et technologie des aliments

LABORATOIRE D'ACCUEIL: INRA – UMR 1083 Sciences Pour l'Oenologie – Equipe Polyphénols et Interactions – 2, place Pierre Viala – F34060 Montpellier Cedex

Valorisation des travaux de recherche

Publications

Boze H., Marlin T., Durand D., Pérez J., Vernhet A., Canon F., Sarni-Manchado P., Cheynier V. et Cabane B. (2010) Proline-rich salivary proteins have extended conformations. *Biophysical Journal*, 99, 656-665.

Canon F., Giuliani A., Paté F. et Sarni-Manchado P. (2010) Ability of a salivary intrinsically unstructured protein to bind different tannin targets revealed by mass spectrometry. *Analytical and Bioanalytical Chemistry*, 398, 815-822.

Canon F., Paté F., Meudec E., Marlin T., Cheynier V., Giuliani A. et Sarni-Manchado P. (2009) Characterization, stoichiometry and stability of salivary protein-tannin complexes by ESI-MS and ESI-MS/MS. *Analytical and Bioanalytical Chemistry*, 395, 2535-2545.

Canon F., Ballivian R., Chirot F., Antoine R., Sarni-Manchado P., Lemoine J. et Dugourd P. Folding of a salivary intrinsically unstructured protein upon binding to tannins. Soumis à *Journal of American Chemical Society*.

Tarascou I., Souquet J.M., Mazauric J.P., Carrillo S., Coq S., Canon F., Fulcrand H. et Cheynier V. (2010) The hidden face of food phenolic composition. *Archives of Biochemistry and Biophysics*, 501, 16-22.

Participation à des congrès nationaux et internationaux

Canon F., Giuliani A., Ballivian R., Chirot F., Albrieux F., Paté F., Cheynier V., Lemoine J., Antoine R., Dugourd P. et Sarni-Manchado P. (2010) Tannin interactions with a human basic salivary proline rich protein studied by mass spectrometry, XXVth International Conference on Polyphenols, Montpellier. Communication orale.

Canon F., Paté F., Giuliani A., Cheynier V. et Sarni-Manchado P. (2010) Understanding functions of a human intrinsically unstructured protein involved in astringency, Journée de l'école doctorale Sciences des Aliments - Sciences des Procédés, Montpellier, France. **Prix de la meilleure communication orale.**

Canon F., Paté F., Giuliani A., Cheynier V. et Sarni-Manchado P. (2009) Mass spectrometry contribution to understand function of an human intrinsically unstructured protein involved in astringency, 18th International Mass Spectrometry Conference, Bremen, Germany. Communication orale.

Canon F., Paté F., Giuliani A., Boze H., Marlin T., Cheynier V. et Sarni-Manchado P. (2009) Etude des interactions tanins – protéines riches en proline salivaires humaines par spectrométrie de masse Journée de l'école doctorale Sciences des Aliments - Sciences des Procédés. **Prix du meilleur Poster.**

Canon F., Giuliani A., Marlin T., Boze H., Mazauric J.-P., Bouchut C., Paté F., Meudec E., Cheynier V. et Sarni-Manchado P. (2008) Contribution de la spectrométrie de masse à l'étude d'une protéine non structurée riche en proline impliquée dans le phénomène d'astringence, 25èmes Journées Françaises de Spectrométrie de Masse, Grenoble, France. **Prix du meilleur Poster.**

Canon F., Giuliani A., Marlin T., Boze H., Paté F., Bouchut C., Meudec E., Cheynier V. et Sarni-Manchado P. (2008) Tannins interactions with a human basic salivary proline rich protein studied by mass spectrometry, XXIVth International Conference on Polyphenols, vol. 1, Salamanca, Spain. pp. 67-68. Communication orale.

Canon F., Meudec E., Marlin T., Sarni- Manchado P. et Cheynier V. (2008) Basic salivary proline rich proteins involved in astringency studied by mass spectrometry, 8th European Symposium on saliva, Egmond aan Zee, The Netherlands. pp. 33. Communication orale.

A mes parents,

Pour votre soutien tout au long de ces années d'études

A Sophie,

Pour ton soutien et tes encouragements au quotidien et pour m'avoir aidé à me dépasser

On est obligé à présent de regarder l'imposant spectacle de l'évolution de la vie comme un ensemble d'événements extraordinairement improbables, impossibles à prédire et tout à fait non reproductibles.

Stephan Jay Gould

Merci à ...

Madame le Docteur Véronique Cheynier,

Pour m'avoir accordé ta confiance et accepté comme doctorant dans ton équipe polyphénols-interactions et comme participant au projet ANR PROTANIN. Je te remercie également pour m'avoir permis de représenter le laboratoire lors de plusieurs congrès internationaux au cours desquels j'ai eu l'opportunité d'exposer mon travail.

Madame le Docteur Pascale Sarni-Manchado,

Pour m'avoir encadré durant ces trois années, pour ta disponibilité au quotidien, pour nos discussions scientifiques chaleureuses, pour m'avoir transmis ta rigueur et ta passion de la biblio et de la science, pour m'avoir formé aux fonctions du métier de chercheur, pour avoir soutenu mes initiatives, et pour avoir su fédérer la SARNITEAM.

Monsieur le Professeur Edwin de Pauw,

Pour l'honneur que vous m'avez fait en acceptant d'être rapporteur de ces travaux.

Monsieur le Professeur Jean-Marie Schmitter,

Pour avoir accepté d'être rapporteur de cette thèse et apporter votre expertise sur le sujet.

Monsieur le Professeur Olivier Dangles,

Pour avoir accepté d'examiner ces travaux.

Madame le Docteur Sonia Longhi,

Pour avoir accepté de participer à ce jury de thèse en apportant votre expertise sur les protéines non structurées.

Nos collaborateurs,

Sans vous une grande partie de ce travail n'aurait pas été possible.

Monsieur le Docteur Alexandre Giuliani,

Pour ton expertise en spectrométrie de masse et tes nombreux conseils très productifs. Merci également pour la confiance que tu m'as accordée en m'ayant recruté en tant que post-doctorant au synchrotron SOLEIL à la suite de ma thèse.

Messieurs les Docteurs Matthieu Réfrégiers, Frédéric James et Franck Wien et Valérie Rouam,

Pour votre accueil au synchrotron SOLEIL et votre bonne humeur.

Messieurs les Docteurs Philippe Dugourd, Rodolphe Antoine, Jérôme Lemoine et Fabien Chirot et Renaud Ballivian,

Pour notre travail en commun sur l'étude des complexes IB5•tanin par spectrométrie de mobilité ionique couplée à la spectrométrie de masse.

Monsieur le Docteur Guillaume Van der Rest,

Pour avoir pris le temps de réaliser avec moi les expériences d'ECD sur les complexes IB5•tanin.

Toutes les personnes qui ont participé de près ou de loin à ce projet.

Monsieur le Docteur Franck Paté,

Pour nos discussions scientifiques, tes connaissances en informatique et pour nos délires partagés qui auront pu exaspérer la pauvre Isabelle souvent cible et procurer à Pascale quelques merveilleux photos montages et fonds d'écran.

Thérèse Marlin, surnommée la « fée colonne »,

Pour m'avoir formé à la purification des protéines et pour avoir été si patiente au cours de mes mésaventures avec les gels d'électrophorèse qui m'ont valu le surnom de « casseur de gel ».

Madame le Docteur Hélène Boze,

Pour avoir apporté ses connaissances sur *Pichia pastoris*, sur les techniques de purification des protéines et sur les caractéristiques des glycosylations assurées par *Pichia pastoris*.

Monsieur le Docteur Nicolas Sommerer,

Pour m'avoir formé à l'utilisation des sources d'ionisation de type MALDI et pour son aide au cours des expériences de caractérisation de II-1.

Messieurs les Docteurs Bernard Cabane, Marc-André Delsuc, Jérôme Golebiowski, Serge Antonczak et Juan Fernandez, Madame le Docteur Dominique Durand et Matthieu Tanty,

Pour les discussions très enrichissantes que nous avons eues lors des réunions entre partenaires de l'ANR PROTANIN.

Messieurs les Docteurs Alain Buléon et Eric Forest,

Pour avoir accepté d'être membres de mon comité de pilotage de thèse et pour vos précieux conseils et critiques.

Emmanuelle Meudec,

Pour m'avoir enseigné les secrets de la trappe ionique du laboratoire et assisté lors de sessions expérimentales à SOLEIL.

Claire Bouchut,

Pour m'avoir formé à l'utilisation du TOF et emmené tester les autres sources d'ionisation de chez Jeol.

Madame le Docteur Laëtizia Mouls, devenue Madame le Docteur RICARD,

Pour m'avoir également aidé à dompter les spectromètres de masse du laboratoire et pour m'avoir fait découvrir quelques uns des endroits mythiques de la vie nocturne montpelliéraine.

Patrick Chamardin,

Pour avoir veillé à la bonne tenue des fermenteurs.

Mesdames les Docteurs Virginie Hougouvieux et Céline Poncet,

Pour nos discussions scientifiques autour des polyphénols et de leurs interactions avec les protéines, mais également pour nos séances de natation dans les différentes piscines de Montpellier.

Mes collègues de travail,

Madame le Docteur Isabelle Tarascou,

Pour avoir partagé mon bureau pendant plus de deux ans, supporté mes blagues et pour les nombreuses soirées que vous avez souvent organisées avec Pascale et pour participants : Jérôme, Carine, Vincent, Franck et Claire.

A ceux qui ont partagé mon bureau : Marie, Aurélie, Benoît, Céline, Stéphanie, Pandi-Panda,

Pour votre bonne humeur, les discussions et les rires partagés.

Aux thésards et post-doc rencontrés à SPO : Marie-Agnès, Anysia, Jérem, Hélène, Chloé, Laurent, Maëva, Yung-Fen, Camilla, Alexandre, Pandi-panda, Tiphaine, Maud, Claire, Fida, Magalie, Pouki et Lucie,

Pour les discussions, les séances piscines, les délires et les soirées partagées.

A l'ensemble des personnes travaillant à SPO et plus particulièrement dans l'équipe Polyphénols-Interactions : Jean-Marc, Jean-Paul, Pascale (Williams), Agnès, Nancy, Clothilde, Thierry, Jean-Luc, Aude, Didier, Martine, Arnaud, Fred, Virginie (Galléote), les frères Picou...

Pour votre accueil tout au long de ces trois années passées à SPO.

Plus personnellement j'aimerais remercier :

Mes parents,

Pour votre soutien, tant moral que financier, tout au long de ces années. Merci de m'avoir laissé libre dans mes choix.

Martine et Paul,

Pour votre accueil, les moments passés en votre compagnie et vos encouragements.

Mes frères : Baptiste et Charles et leurs amies respectives Marie et Marie-Jeanne,

Pour supporter mes humeurs, pour les moments passés ensemble lorsque l'occasion se présentait.

Marsouin, Ronron, Lauren, Alex, Prout-Prout, Greg, Elodie, Mat, Mélane, Gueil, Anaïs, Ben, Steph, Simon, Céline,

Pour les moments de détente passés ensemble...

Plus chèrement je remercie :

Sophie, sans qui je n'en serais pas là,

Merci pour m'avoir aidé à croire en moi, pour m'avoir stimulé, motivé et encouragé tout au long de mon parcours. Ton soutien et ton esprit volontaire m'ont aidé à avancer tout au long de ma thèse. Merci pour ton affection, pour ta tendresse et tes attentions quotidiennes malgré notre éloignement. Les week-end et vacances passés ensemble ont été des moments de détente et de réconfort très précieux pour moi.

Sommaire

Valorisation des travaux de recherche	5
Sommaire	15
Liste des figures	19
Liste des tableaux	25
Liste des abréviations	27
Constantes physiques fondamentales	29
Nomenclature des acides aminés	30
Introduction générale	31
Etat de l'art	33
1. L'astringence	33
1.1. Mécanisme physiologique de l'astringence	33
1.2. Astringence et évaluation sensorielle	33
1.3. Astringence et évolution	34
2. Les Tanins	35
2.1. Tanins et structure	35
2.1.1. Tanins hydrolysables	35
2.1.2. Tanins condensés ou proanthocyanidines	36
2.2. Tanins et astringence	38
2.3. Tanins et écologie	39
2.4. Tanins et protéines	40
2.4.1. Effet de la structure du tanin	42
2.4.2. Nature des interactions	43
2.4.3. Influence de l'environnement	43
3. La salive	44
3.1. La salive : composition et sécrétion	44
3.2. La salive : fonctions	45
4. Les protéines riches en proline ou PRP	45
4.1. PRP : aspect génétique	46
4.2. PRP : caractéristiques et structures	46
4.3. PRP : fonctions	47
4.4. PRP : synthèse	48
4.5. PRP : relation avec l'astringence	49
5. Les protéines non structurées : IUP, IDP ou PID	50
5.1. PID : le réaménagement du paradigme séquence-structure-fonction	51
5.2. PID : caractéristiques	52
5.3. PID : fonctions	53
5.4. PID : bénéfice fonctionnel	54
6. Les méthodes d'étude utilisées en biologie structurale	57
7. La spectrométrie de masse ou MS	60
7.1. Les sources d'ionisation	62
7.1.1. L'électrospray (ESI ou ES)	63
7.1.2. La Désorption et Ionisation par Laser Assistées par Matrice (MALDI)	67
7.2. Les analyseurs	70
7.2.1. Les analyseurs de type temps de vol « TOF »	72
7.2.2. Les analyseurs de type quadripôle	77
7.2.2.1. Structure des quadripôles	77
7.2.2.2. Stabilité des ions dans un quadripôle	78
7.2.3. Les analyseurs de type trappe ionique de Paul ou 3D	81
7.2.3.1. Structure de la trappe	81
7.2.3.2. Piégeage des ions	82
7.2.3.3. Expulsion des ions de la trappe	84

7.2.4. Les analyseurs de type trappe linéaire ou 2D	84
7.2.5. Les analyseurs de type FT-ICR.....	86
7.3. La spectrométrie de masse en tandem ou MS/MS	93
7.3.1. La dissociation induite par collision (CID)	95
7.3.2. La dissociation par capture d'électrons (ECD).....	98
7.4. La spectrométrie de masse couplée à la spectrométrie de mobilité ionique (IMS-MS).....	99
7.4.1. Instrumentation.....	100
7.4.2. Principe	101
Matériel et méthodes	105
1. Le matériel biologique	105
1.1. Les PRP modèles choisies : IB5 et II-1	105
1.2. Les tanins modèles	106
1.3. Le « ligand » témoin	106
1.4. La protéine témoin BPTI.....	107
2. Production des PRP modèles	108
3. Purification des PRP modèles	110
3.1. Protocole de purification n°1	110
3.2. Protocole de purification n°2	112
3.2.1. Adsorption/désorption des PRP sur une colonne échangeuse d'ions.....	112
3.2.2. Séparation des PRP par gel filtration	113
3.3. Caractérisation des PRP	114
3.3.1. Caractérisation des PRP par SDS-PAGE	114
3.3.2. Caractérisation de la séquence N-terminale des PRP	114
3.3.3. Caractérisation des PRP par ESI-MS	114
3.4. Déglycosylation de II-1	115
4. Solutions d'interaction PRP - tanins	116
4.1. Les milieux utilisés	116
4.2. Les solutions mères	116
4.3. Les solutions d'interaction	116
5. Etude par spectrométrie de masse.....	118
5.1. Spectromètres de masse utilisés avec source ESI	118
5.2. Spectromètres de masse utilisés avec source MALDI	118
5.3. Préparation des échantillons pour les analyses MALDI-MS	118
Chapitre I. Caractérisation des protéines II-1, II-1NG et IB5	121
I.1. Production et purification des PRP	122
I.2. Caractérisation des protéines II-1 et II-1NG par ESI-MS et MALDI-MS.....	125
I.2.1. Caractérisation des chaînes peptidiques des protéines II-1NG et II-1	129
a. Caractérisation de la chaîne peptidique de II-1NG	129
b. Caractérisation de la chaîne peptidique de II-1.....	132
I.2.2. Caractérisation de la protéine II-1	134
I.2.3. Détermination du nombre de glycanes portées par II-1	135
I.2.4. Détermination de la localisation des glycanes	136
I.3. Caractérisation de la protéine IB5	140
Chapitre II. Etude des interactions IB5-EgCG	143
II.1. Introduction	143
II.2. Etudes des interactions IB5-EgCG par ESI-MS et ESI-MS/MS CID.....	143
Publication n°1 : Characterization, Stoichiometry and Stability of Salivary Protein-tannin Complexes by ESI-MS and ESI-MS/MS	144
II.3. Discussion	170
II.4. Conclusions et perspectives.....	172
Chapitre III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin	175
III.1. Introduction	175

III.2 Etude des interactions IB5-tanin par ESI-MS et ESI-MS/MS CID	175
Publication n°2 : Ability of a Salivary Intrinsically Unstructured Protein to Bind Different Tannin Targets Revealed By Mass Spectrometry	176
III.3. Etude des interactions IB5-tanin par compétition par ESI-MS	194
III.4. Etude des complexes de stœchiométrie 1•2 par ESI-MS/MS CID.....	202
III.5. Conclusion et Perspectives	207
Chapitre IV. Etude structurale de l'interaction IB5-tanin	211
IV.1. Introduction	211
IV.2. Structure des protéines et formation de complexes : IB5-EgCG Vs. BPTI-EgCG	212
IV.3. Nombre de sites d'interaction et position	216
IV.3.1. Introduction	216
IV.3.2. Etude du nombre de sites d'interaction par ESI-MS/MS CID	217
IV.3.3. Etude de la position des sites d'interaction par ESI-MS/MS ECD	219
IV.4. Etude des modifications de conformation de la protéine IB5 en interaction avec les tanins.....	225
IV.4.1. Introduction	225
IV.4.2. Etude des conformations de IB5 et de ses complexes IB5•EgCG par ESI-IMS- MS	225
Publication n°3 : Folding of a Salivary Intrinsically Unstructured Protein Upon Binding to Tannins.....	227
IV.5. Conclusions et perspectives	230
Conclusion générale et perspectives	233
Références bibliographiques	243
Annexe 1	265

Liste des figures

Figure 1. Exemple d'enregistrement de l'intensité ressentie en fonction du temps et paramètres calculés.	34
Figure 2. Structure d'un tanin hydrolysable, le pentagalloylglucose.	36
Figure 3. Squelette des monomères des proanthocyanidines.	36
Figure 4. a. structure des monomères de flavan-3-ols ; b. exemple de tétramère comportant les deux types de liaisons interflavaniques de type B C4-C6 et C4-C8 et une liaison de type A C2-O-C7.	37
Figure 5. Le continuum des structures des protéines, tiré de (Dyson et Wright, 2005).	52
Figure 6. Transition désordre-ordre couplée à une interaction protéine-protéine, tiré de (Dyson et Wright, 2005).	56
Figure 7. Représentation schématique d'un spectromètre de masse.	61
Figure 8. Principe de l'électrospray.	64
Figure 9. Libération des ions à partir des gouttelettes émises lors de la nébulisation ; a. libération directe des ions à partir de la surface des gouttelettes ; b. éjection d'un ion encore solvaté (nanogouttelette) et dont le solvant résiduel sera évaporé.	66
Figure 10. Exemple de dépôt : « <i>Dried-droplet method</i> ».	68
Figure 11. Principe de l'ionisation de type MALDI.	69
Figure 12. Formation de la plume MALDI, tiré de (Poretzky <i>et al.</i> , 1999).	69
Figure 13. La résolution en masse.	71
Figure 14. Principe d'un analyseur de type TOF linéaire.	72
Figure 15. Principe d'un analyseur de type TOF à réflectron.	75
Figure 16. Principe de a. l'extraction continue et de b. l'extraction retardée.	77
Figure 17. Principe des analyseurs de type quadripôle : a. schéma d'un analyseur quadripolaire ; b. évolution des ions chargés positivement dans un quadripôle au cours d'un cycle de tension alternative.	78
Figure 18. Zones de stabilité pour un ion de masse m donnée suivant x ou y , adapté de (Hoffmann, 2005).	80
Figure 19. Zones de stabilité A en fonction de U et V pour des ions de différents rapports m/z ($m/z_1 < m/z_2 < m/z_3$), adapté de (Hoffmann, 2005).	81
Figure 20. Photographie et schéma d'une trappe ionique 3D, adapté de (March, 1997).	82
Figure 21. Trajectoire d'un ion sous forme d'un 8 tridimensionnel dans une trappe ionique.	82
Figure 22. Diagramme de stabilité suivant r et z d'une trappe ionique. Les régions de stabilité simultanée sont marquées A et B, adapté de (March, 1997).	83
Figure 23. Domaines de stabilité A de trois ions de rapport m/z différents en fonction de U et V dans une trappe ionique, adapté de (Hoffmann, 2005).	84
Figure 24. Trappe linéaire ou 2 D avec fentes dans les barres opposées.	85
Figure 25. Schéma d'un QIT, spectromètre de masse hybride composé de 2 quadripôles en Q1 et Q3 et d'une trappe linéaire en q2, d'après (Hager, 2002). Les électrodes d'entrée et de sortie de la trappe linéaire sont marquées IQ2 et IQ3. Q0 représente un quadripôle fonctionnant en mode « <i>RF only</i> » (guide d'ion).	86

Figure 26. Schéma d'une cellule ICR. Cette cellule est placée à l'intérieur d'un aimant qui génère le champ magnétique B .	87
Figure 27. Schéma d'un piège de Penning contenant une particule de charge positive en son centre (en rouge). Le champ électrique E (en bleu) est produit par deux coupelles chargés positivement (a) et par une électrode circulaire chargée négativement (b). Le champ magnétique homogène B (en rouge) est produit par un électro-aimant toroïdal (c).	89
Figure 28. Trajectoire cyclotron classique d'un ion dans le plan radiale d'un piège de Penning. Elle est composée d'un mouvement cyclotron réduit v_+ et d'un mouvement magnétron v_- .	89
Figure 29. Etapes de la mesure de masse en FT-ICR : a. Ions avant excitation ; b. excitation des ions jusqu'à atteindre une certaine orbite ; c. mouvement cohérent des ions de même m/z créant le courant induit.	91
Figure 30. Principe de la spectrométrie de masse en tandem (MS/MS).	93
Figure 31. Principe de la MS/MS de type CID.	95
Figure 32. Nomenclature des ions issus de la fragmentation peptidique en MS/MS.	96
Figure 33. Utilisation de la CID pour l'étude de la dissociation des complexes non-covalents.	96
Figure 34. Principe de la fragmentation de type ECD appliquée à l'étude d'un complexe non-covalent.	98
Figure 35. Principe du couplage IMS-MS.	99
Figure 36. Spectre bidimensionnel IMS-MS. Le spectre de masse est en ordonnée et le spectre de mobilité ionique est en abscisse. L'intensité des signaux est indiquée par les variations de couleur.	100
Figure 37. Structure d'un appareil de IMS-MS à une seule cellule de mobilité avec un analyseur de type Q-TOF.	101
Figure 38. Distribution des temps d'arrivée en mobilité ionique.	101
Figure 39. Corrélation entre le temps d'arrivée des ions et le rapport pression de gaz dans le tube de mobilité sur la tension appliquée.	103
Figure 40. Séquence primaire de la proprotéine PRB4S.	105
Figure 41. Clivage de la proprotéine PRB4S par les protéases similaires à la furine.	106
Figure 42. Séquence primaire des protéines a. II-1 et b. IB5 issues du clivage de PRB4S.	106
Figure 43. a. Structure des tanins modèles (M : monomères, D : dimères, G : galloyl) ; b. structure de la réserpine.	107
Figure 44. Structure et séquence primaire de BPTI.	108
Figure 45. Schéma simplifié des étapes de purification des protéines II-1 et IB5.	111
Figure 46. Schéma des étapes de la chromatographie en lit expansé.	113
Figure 47. Représentation schématique de l'action des N-glycannases utilisées.	115
Figure 48. Détails des deux expériences de compétition.	117
Figure 49. Profils SDS-PAGE (12,5% acrylamide, 0,1% SDS) des différentes étapes de la production et de la purification des PRP excrétées par la souche recombinante <i>Pichia pastoris-PRB4S</i> (coloration au bleu de Coomassie. MqM : marqueurs de masse moléculaire en kDa).	122

Figure 50. Espèces protéiques sécrétées par <i>P.pastoris-PRB4S</i> . Les protéines ont été séparées sur un gel SDS-PAGE à 12,5% d'acrylamide et colorées au bleu de Coomassie. Piste 1 : surnageant brut ; piste 2 : fraction de PRP après adsorption/désorption sur une colonne échangeuse d'ion ; les pistes 3 à 5 correspondent aux fractions purifiées par chromatographie d'exclusion sur une colonne HiLoad Superdex 75 HR ; piste 3 : II-1 lyophilisée ; piste 4 : II-1 non-glycosylée lyophilisée ; piste 5 : IB5 lyophilisée ; piste MqM : marqueurs moléculaires de faible masse moléculaire (LMW, Pharmacia).	124
Figure 51. Profil d'éluion de la colonne HiLoad Superdex 75.	125
Figure 52. Les deux types de glycosylation des protéines : a. liaison de type O-glycosidique ; b. liaison de type N-glycosidique.....	126
Figure 53. Approche développée pour l'étude de II1.	128
Figure 54. a. spectre ESI-MS de la protéine II-1NG ; b. déconvolution du spectre ESI-MS de la protéine II-1NG ; c. spectre MALDI-MS de la protéine II-1NG. Les 9 formes présentes sur ces spectres sont nommées de A à F (N.B. ce code de lettre est différent de celui des figures 55 à 63). Sur le spectre ESI-MS, les lettres d'identification sont accompagnées d'un chiffre correspondant à l'état de charge de l'ion correspondant.	130
Figure 55. Position des débuts et des fins de séquence des différents peptides de II-1 et de II-1NG identifiés.	131
Figure 56. SDS-PAGE de la protéine II-1 avant et après traitement enzymatique. Piste 1 : MqM LWM ; piste 2 : II-1 purifiée ; piste 3 : II-1 déglycosylée par l'enzyme PNGaseF ; piste 4 : II-1 déglycosylée par l'enzyme endoH ; piste 5 : II-1 après action des enzymes PNGaseF et GluC ; piste 6 : II-1 après action des enzymes endoH et GluC.....	132
Figure 57. Spectre MALDI-MS de la protéine II-1 déglycosylée par la PNGaseF (se reporter à la figure 55 pour le code lettre utilisé).	133
Figure 58. Spectre MALDI-TOF de la protéine II-1 acquis en mode positif.	134
Figure 59. Spectre MALDI-MS de la protéine II-1 déglycosylée par l'endoH. Les espèces marquées * présentent une seule glycosylation et celle marqués ** en possèdent deux (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).....	136
Figure 60. Spectre MALDI-MS en mode linéaire de la protéine II-1 déglycosylée par l'enzyme PNGaseF et clivée par l'enzyme GluC (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).	137
Figure 61. Spectre MALDI-MS en mode réflectron de la protéine II-1 déglycosylée par l'enzyme PNGaseF et clivée par l'enzyme GluC (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).	138
Figure 62. Spectre MALDI-MS en mode linéaire de la protéine II-1 déglycosylée par l'enzyme endoH et clivée par l'enzyme GluC. Les espèces marquées * présentent une seule glycosylation et celle marqués ** en possèdent deux (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).....	139
Figure 63. Spectre MALDI-MS en mode réflectron de la protéine II-1 déglycosylée par l'enzyme endoH et clivée par l'enzyme GluC. Les espèces marquées * présentent une seule glycosylation et celle marqués ** en possèdent deux (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).....	140
Figure 64. a. spectre de masse d'une fraction d'IB5 purifiée selon le protocole n°2 ; b. séquence des chaînes polypeptidiques identifiées (les séquences répétées sont soulignées).	141
Figure 65. Evolution des proportions de protéine libre et de protéine complexée en fonction de la concentration en EgCG présent dans le mélange (ions dont état de charge est +7)..	169

Figure 66. Spectre MS du mélange IB5:EgCG au rapport molaire 1/30 (5 μM /150 μM). Les pics correspondant à $[\text{IB5}]^{7+}$ et aux complexes $[\text{IB5}\cdot n\text{EgCG}]^{7+}$ sont légendés.	169
Figure 67. Principe du test de validation de l'observation des complexes non-covalents par spectrométrie de masse. La protéine est dénaturée à pH acide et son site de liaison n'est plus adapté à la forme du ligand.	171
Figure 68. Région d'un spectre de masse d'une solution IB5-EgCG (5 μM -50 μM) eau-éthanol 88/12 (v/v) non acidifiée pH 5,7.....	172
Figure 69. Voies évolutives suivies par expansion des séquences répétées chez les PID, tiré de (Tompa, 2003b).....	173
Figure 70. Spectre MS de la solution d'interaction IB5 5 μM - EgCG 25 μM - EgC 25 μM	195
Figure 71. Spectre MS de la solution d'interaction IB5 5 μM - EgCG 25 μM - ECG 25 μM	195
Figure 72. Spectre MS de la solution d'interaction IB5 5 μM - EgCG 25 μM – B2 25 μM . Le point d'interrogation indique que la présence du pic de m/z correspondant à $[\text{IB5}\cdot 2\text{B2}]^{7+}$ ne peut être déterminée car ce pic a le même rapport m/z que $[\text{IB5}]^{6+}$	196
Figure 73. Spectre MS de la solution d'interaction IB5 5 μM - EgCG 25 μM – B2 3'OG 25 μM	196
Figure 74. Spectre MS de la solution d'interaction IB5 5 μM – EgG 25 μM – ECG 25 μM . .	197
Figure 75. Spectre MS de la solution d'interaction IB5 5 μM – EgC 25 μM – B2 25 μM . Le point d'interrogation indique que la présence du pic de m/z correspondant à $[\text{IB5}\cdot 2\text{B2}]^{7+}$ ne peut être déterminée car ce pic a le même rapport m/z que $[\text{IB5}]^{6+}$	197
Figure 76. Spectre MS de la solution d'interaction IB5 5 μM – EgC 25 μM – B2 3'OG 25 μM	198
Figure 77. Spectre MS de la solution d'interaction IB5 5 μM – ECG 25 μM – B2 25 μM . Le point d'interrogation indique que la présence du pic de m/z correspondant à $[\text{IB5}\cdot 2\text{B2}]^{7+}$ ne peut être déterminée car ce pic a le même rapport m/z que $[\text{IB5}]^{6+}$	198
Figure 78. Spectre MS de la solution d'interaction IB5 5 μM – ECG 25 μM – B2 3'OG 25 μM	199
Figure 79. Spectre MS de la solution d'interaction IB5 5 μM – B2 3'OG 25 μM – B2 25 μM . Le point d'interrogation indique que la présence du pic de m/z correspondant à $[\text{IB5}\cdot 2\text{B2}]^{7+}$ ne peut être déterminée car ce pic a le même rapport m/z que $[\text{IB5}]^{6+}$	199
Figure 80. Courbes de dissociation des tanins du complexe $[\text{IB5}\cdot \text{EgCG}\cdot \text{B2 3'OG}]^{6+}$ et de libération d'IB5.....	203
Figure 81. a. courbes de dissociation du premier tanin des complexes $[\text{IB5}\cdot \text{EgC}\cdot \text{T}]^{6+}$ avec T= EgC, ECG, EgCG, B2 ou B2 3'OG ; b. courbes de dissociation du second tanin des complexes $[\text{IB5}\cdot \text{EgC}\cdot \text{T}]^{6+}$	204
Figure 82. a. courbes de dissociation du premier tanin des complexes $[\text{IB5}\cdot \text{B2 3'OG}\cdot \text{T}]^{6+}$ avec T= EgC, ECG, EgCG, B2 ou B2 3'OG ; b. courbes de dissociation du second tanin des complexes $[\text{IB5}\cdot \text{B2 3'OG}\cdot \text{T}]^{6+}$	205
Figure 83. Spectre de masse ESI-TOF d'une solution de BPTI à 5 μM dans un solvant eau-éthanol 88/12 acidifié à pH 3,2 avec de l'acide acétique.	213
Figure 84. Spectre du mélange BPTI 5 μM - EgCG 50 μM (seuls les pics les plus importants sont légendés).....	214
Figure 85. Techniques de fragmentation applicables aux protéines et types de fragments produits.	216

Figure 86. Courbes de dissociation des complexes IB5•EgCG de stœchiométries 1•1 à 1•9 présents dans la solution IB5:EgCG 5 µM:200 µM dans le solvant H ₂ O/EtOH 88/12 à pH 3,2 acidifiée avec de l'acide acétique.....	218
Figure 87. a. Spectre MS de la solution IB5 5 µM ; b. Spectre MS/MS ECD de l'ion parent [IB5] ⁷⁺ (<i>m/z</i> 990).....	220
Figure 88. a. Spectre MS de la solution IB5 : EgCG 5 µM : 50 µM ; b. spectre MS/MS ECD de l'ion parent [IB5•EgCG] ⁷⁺ (<i>m/z</i> 1055).	221
Figure 89. a. Spectre MS de la solution IB5 : B2 3'OG 5 µM : 50 µM ; b. Spectre MS/MS ECD de l'ion parent [IB5•B2 3'OG] ⁷⁺ (<i>m/z</i> 1094).	223
Figure 90. Fragments obtenus en MS/MS ECD pour a. [IB5] ⁷⁺ (<i>m/z</i> 990) et b. [IB5•1B2 3'OG] ⁷⁺ . Les fragments ayant retenu le B2 3'OG sont indiqués avec le symbole ♣.	224
Figure 91. Mécanisme de fragmentation ECD pour un résidu proline.	224
Figure 92. Sections efficaces de collision des états de charge 6+ à 9+ de la protéine IB5.	226
Figure 93. Deux exemples de conformation de la protéine IB5 compatibles avec les données de SAXS ; d'après (Boze <i>et al.</i> , 2010).	238
Figure 94. Courbes de dissociation des tanins du complexe [IB5•EgC•EgC] ⁶⁺ et de libération d'IB5.	265
Figure 95. Courbes de dissociation des tanins du complexe [IB5•EgC•ECG] ⁶⁺ et de libération d'IB5.	265
Figure 96. Courbes de dissociation des tanins du complexe [IB5•EgC•EgCG] ⁶⁺ et de libération d'IB5.....	266
Figure 97. Courbes de dissociation des tanins du complexe [IB5•EgC•B2] ⁶⁺ et de libération d'IB5.	266
Figure 98. Courbes de dissociation des tanins du complexe [IB5•EgC•B2 3'OG] ⁶⁺ et de libération d'IB5.....	267
Figure 99. Courbes de dissociation des tanins du complexe [IB5•ECG•ECG] ⁶⁺ et de libération d'IB5.	267
Figure 100. Courbes de dissociation des tanins du complexe [IB5•ECG•EgCG] ⁶⁺ et de libération d'IB5.....	268
Figure 101. Courbes de dissociation des tanins du complexe [IB5•ECG•B2] ⁶⁺ et de libération d'IB5.	268
Figure 102. Courbes de dissociation des tanins du complexe [IB5•ECG•B2 3'OG] ⁶⁺ et de libération d'IB5.....	269
Figure 103. Courbes de dissociation des tanins du complexe [IB5•EgCG•EgCG] ⁶⁺ et de libération d'IB5.....	269
Figure 104. Courbes de dissociation des tanins du complexe [IB5•EgCG•B2] ⁶⁺ et de libération d'IB5.....	270
Figure 105. Courbes de dissociation des tanins du complexe [IB5•EgCG•B2 3'OG] ⁶⁺ et de libération d'IB5.....	270
Figure 106. Courbes de dissociation des tanins du complexe [IB5•B2•B2] ⁶⁺ et de libération d'IB5.	271
Figure 107. Courbes de dissociation des tanins du complexe [IB5•B2•B2 3'OG] ⁶⁺ et de libération d'IB5.....	271

Figure 108. Courbes de dissociation des tanins du complexe [IB5•EgCG•EgCG]⁶⁺ et de libération d'IB5..... 272

Liste des tableaux

Tableau 1. Prédiction de la présence de désordre au sein du protéome des 3 règnes du vivant. Les pourcentages indiqués correspondent aux valeurs les plus hautes et les plus basses observées (rem : seules les protéines ayant plus de 60 acides aminés ont été prises en compte).	50
Tableau 2. Les cinq grandes classes de PID.	54
Tableau 3. Principales techniques utilisées en biologie structurale (Receveur-Brechot <i>et al.</i> , 2006).	58
Tableau 4. Analyseurs utilisés au cours de cette étude.	92
Tableau 5. Tableau de comparaison MS/MS spatiale versus MS/MS temporelle.	94
Tableau 6. Résultats du séquençage N-terminal des protéines IB5 et II-1. Les acides aminés en gras sont majoritaires. La lettre X indique que le signal était trop faible ou trop diffus pour permettre l'identification.	123
Tableau 7. Séquences théoriques obtenues après clivage de la protéine II-1 par l'endoprotéinase GluC et leurs masses moléculaires correspondantes.	128
Tableau 8. Informations attendues à partir des différentes digestions enzymatiques.	129
Tableau 9. Tableau récapitulatif des différentes formes de chaînes polypeptidiques de la protéine II-1NG et des masses correspondantes (attention le code de lettre de la figure 54 (code à une seule lettre) est différent de celui des figures 55 à 63 (code à deux lettres))...	131
Tableau 10. Concentrations en IB5 et en EgCG testées durant les expériences de titration.	168
Tableau 11. Récapitulatif des stœchiométries maximales observées pour chacune des expériences de compétition.	200
Tableau 12. E_{CoM} de demi dissociation des tanins des complexes [IB5•EgC•T] ⁶⁺	204
Tableau 13. E_{CoM} de demi dissociation des tanins des complexes [IB5•B2 3'OG•T] ⁶⁺	205
Tableau 14. Rappel des E_{CoM} de demi dissociation des complexes [IB5•1T] ⁶⁺	206
Tableau 15. Expériences réalisées pour étudier l'impact de la structure des tanins sur leur interaction avec IB5.	207
Tableau 16. Pourcentages de BPTI et d'IB5 libres et complexés calculés à partir des spectres MS des solutions IB5:EgCG et BPTI:EgCG.	215

Liste des abréviations

II-1	Protéine II-1
II-1NG	Protéine II-1 sous sa forme non glycosylée
Ω	Section efficace de collision (<i>collision cross section</i>)
a.a.	Acides aminés
ADN	Acide DésoxyriboNucléique
ADNc	ADN complémentaire
APPI	Photoionisation à pression atmosphérique (<i>Atmospheric-Pressure Photoionization</i>)
aPRP	Protéines riches en proline acides
ANR	Agence Nationale de la Recherche
ARN	Acide RiboNucléique
ARNt	ARN de transfert
ARNm	ARN messenger
ATD	Distribution des temps d'arrivées (<i>Arrival Time Distribution</i>)
B2	(-)-épicatechine (4- β -8)-(-)-épicatechine
B2 3'OG	B2 3'O gallate, (-)-épicatechine (4- β -8)-(-)-épicatechine-3-O-gallate
BPTI	Inhibiteur pancréatique bovin de la trypsine (<i>Bovine Pancreatic Trypsin Inhibitor</i>)
BSA	Albumine de sérum bovin (<i>bovine serum albumin</i>)
bPRP	Protéines riches en proline basiques
CD	Dichroïsme circulaire (<i>Circular Dichroism</i>)
CI	Ionisation chimique (<i>Chemical Ionization</i>)
CID	Dissociation induite par collision (<i>Collision-Induced Dissociation</i>)
CRM	Modèle du résidu chargé (<i>Charge Residue Model</i>)
Cryo-ME	CryoMicroscopie Electronique
C-term	C-terminal
Da	Dalton
DLS	Diffusion dynamique de la lumière (<i>Dynamic Light Scattering</i>)
DP	Degré de polymérisation
DPm	Degré de polymérisation moyen
DTT	Dithiothréitol, (2S,3S)-1,4-Bis-sulfanylbutane-2,3-diol
ECD	Dissociation par capture d'électrons (<i>Electron Capture Dissociation</i>)
E_{CoM}	Energie au centre des masses (<i>Center of Mass Energy</i>)
EI	Ionisation par impact électronique (<i>Electron Ionization</i>)
EgC	(-)-épigallocatechine
EgCG	(-)-épigallocatechine gallate
EGC	(-)-épicatechine gallate
E_{lab}	Energie cinétique dans le référentiel du laboratoire (<i>kinetic Energy in the laboratory frame</i>)
EndoH	Endo- β -N-acetylglucosaminidase H
ESI	Electronébulisation (<i>ElectroSpray Ionization</i>)
ETD	Dissociation par transfert d'électrons (<i>Electron Transfert Dissociation</i>)
EtOH	Ethanol
e	Charge de l'électron
eV	Electronvolt
FT-ICR	Résonance cyclotron à transformée de Fourier (<i>Fourier Transform Ion Cyclotron Resonance</i>)
FTIR	Spectroscopie infrarouge à transformée de Fourier (<i>Fourier Transform InfraRed spectroscopy</i>)
FWHM	Largeur à mi-hauteur (<i>Full Width at Half-Maximum Height</i>)

GluC	Endoprotéinase Glu-C
GlcNAc	N-Acetylglucosamine
gPRP	Protéines riches en proline glycosylées
IB5	Protéine IB5
IEM	Modèle de formation des ions par évaporation (<i>Ion Evaporation Model</i>)
ICR	Résonance d'ions en cyclotron (<i>Ion Cyclotron Resonance</i>)
IMS	Spectrométrie de mobilité ionique (<i>Ion Mobility Spectrometry</i>)
IMS-MS	Spectrométrie de mobilité ionique couplée à la spectrométrie de masse (<i>Ion Mobility Spectrometry – Mass Spectrometry</i>)
IR	Infrarouge
IT	Trappe ionique (<i>Ion Trap</i>)
ITC	Calorimétrie de titration isotherme (<i>Isothermal Titration Calorimetry</i>)
IT-orbitrap	Spectromètre de masse hybride constitué d'une trappe d'ion et d'une orbitrap (<i>Ion Trap – Orbitrap</i>)
IT-TOF	Spectromètre de masse hybride constitué d'une trappe d'ion et d'un temps de vol (<i>Ion Trap – Time-Of-Flight</i>)
IUP	Protéine intrinsèquement non structurée (<i>Intrinsically Unstructured Protein</i>)
LIT	Trappe ionique linéaire (<i>Linear Ion Trap</i>)
LWM	Kit de marqueurs de faibles masses moléculaires (<i>Low Weight Molecular</i>)
Man	Mannose
m_i	Masse de l'ion
MALDI	Désorption et ionisation par laser assistées par matrice (<i>Matrix-Assisted Laser Desorption Ionization</i>)
ME	Microscopie électronique
MM	Masse moléculaire (exprimée en uma ou en Dalton) (<i>Molecular Mass</i>)
MqM	Marqueurs moléculaires
MS	Spectrométrie de masse (<i>Mass Spectrometry</i>)
MS/MS	Spectrométrie de masse en tandem
m/z	Rapport masse-sur-charge (<i>mass-to-charge ratio</i>)
N-term	N-terminal
PAGE	Electrophorèse avec gel polyacrylamide (<i>PolyAcrylamide Gel Electrophoresis</i>)
pI	Point isoélectrique
PID	Protéine Intrinsèquement Désordonnée (<i>Intrinsically Disordered Protein</i>)
PNGase F	Peptidyl N-glycosidase F
<i>P. pastoris</i>	<i>Pichia Pastoris</i>
PRB4S	Proprotéine dont la maturation conduit à la formation des protéines II-1 et IB5
<i>PRB4S</i>	Gène codant pour la proprotéine PRB4S
PRP	Protéine riche en proline (<i>Proline Rich-Protein</i>)
pRP	Peptide riche en proline (<i>Proline Rich-peptide</i>)
PSD	Décomposition post-source (<i>Post-Source Decay</i>)
PTM	Modification post-traductionnelle (<i>Post-Translational Modification</i>)
Q	Quadripôle (<i>Quadrupole</i>)
q	Charge totale des ions ($q=ze$)
Q-q-Q	Spectromètre de masse hybride constitué du couplage de trois quadripôles
Q-TOF	Spectromètre de masse hybride constitué d'un quadripôle couplé à un temps de vol (<i>Quadrupole Time-Of-Flight</i>)
QIT	Trappe d'ion quadripolaire (<i>Quadrupole Ion Trap</i>)
RE	Réticulum Endoplasmique
RF	Radiofréquence
RF only	Quadripôle en mode radiofréquence sans tension continue
RMN	Résonance magnétique nucléaire
SAXS	Diffusion des rayons X (<i>Small Angle X-ray Scattering</i>)
SDS	Laurylsulfate de sodium ou dodécylsulfate de sodium (<i>Sodium Dodécyl Sulfate</i>)

SDS-PAGE	Sodium Dodécyl Sulfate - Electrophorèse avec gel polyacrylamide (<i>Sodium Dodécyl Sulfate - PolyAcrylamide Gel Electrophoresis</i>)
SEC	Chromatographie par exclusion stérique (<i>Size-Exclusion Chromatography</i>)
S/N	Rapport signal sur bruit (<i>Signal-to-Noise ratio</i>)
TBSP	Protéine salivaire fixatrice de tanin (<i>Tannin-Binding Salivary Protein</i>)
TEMED	Tétra-méthyl-éthylènediamine
TI	Méthode Temps - Intensité
TIC	Courant ionique total (<i>Total Ion Current</i>)
TMAC	Chlorure de tétraméthylammonium (<i>TetraMethylAmmonium Chloride</i>)
TOF	Temps de vol (<i>Time-Of-Flight</i>)
Tris HCl	Tris HydroChloride, 2-Amino-2-(hydroxyméthyl)-1,3-propanediol,hydrochloride, Tris (hydroxyméthyl) aminométhane Hydrochloride
uma	Unité de masse atomique
UV	Ultraviolet
V	Tension appliquée
Vcell	Tension appliquée sur la cellule de collision
z	Nombre de charge de l'ion

Constantes physiques fondamentales

Quantité	Symboles	Valeur
Constante de gravitation	G	$6,67259 \cdot 10^{-11} \text{ m}^3 \cdot \text{kg}^{-1} \cdot \text{s}^{-2}$
Charge élémentaire	e	$1,60217733 \cdot 10^{-19} \text{ C}$
Unité atomique de masse, Dalton	u, Da	$1,6605402 \cdot 10^{-27} \text{ kg}$
Constante de Planck	h	$6,6260755 \cdot 10^{-34} \text{ J} \cdot \text{s}$
Constante de Boltzmann	k	$1,380658 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1}$
Electron-volt	eV	$23,06054 \text{ kcal} \cdot \text{mole}^{-1}$
Constante d'Avogadro	N_A	$6,0221367 \cdot 10^{23} \cdot \text{mole}^{-1}$
Constante molaire des gaz	R	$8,314510 \text{ J} \cdot \text{mole}^{-1} \cdot \text{K}^{-1}$

Nomenclature des acides aminés

A	Alanine (Ala)
D	Acide aspartique ou Aspartate (Asp)
C	Cystéine (Cys)
E	Acide glutamique ou Glutamate (Glu)
F	Phénylalanine (Phe)
G	Glycine (Gly)
H	Histidine (His)
I	Isoleucine (Ile)
K	Lysine (Lys)
L	Leucine (Leu)
M	Méthionine (Met)
N	Asparagine (Asn)
P	Proline (Pro)
Q	Glutamine (Glu)
R	Arginine (Arg)
S	Sérine (Ser)
T	Thréonine (Thr)
V	Valine (Val)
W	Tryptophane (Trp)
Y	Tyrosine (Tyr)

Introduction générale

Le paradigme classique séquence-structure-fonction (Anfinsen, 1973) stipule que la structure d'une protéine induite par sa séquence définit sa fonction. La protéine doit donc posséder une structure tridimensionnelle précise pour avoir une activité. Cependant, durant la dernière décennie, l'accumulation de résultats sur un certain nombre de protéines a montré la nécessité de repenser ce paradigme et de définir une nouvelle catégorie de protéines (Wright et Dyson, 1999). Ces protéines présentent dans leur état natif une absence partielle ou complète de structure tridimensionnelle définie et possèdent une activité biologique. Elles ont été nommées « Intrinsically Unstructured Protein » (IUP) (Tompa, 2002) ou « Intrinsically Disordered Protein » (IDP) (Dunker *et al.*, 2001). En français, nous utiliserons le terme de protéines intrinsèquement désordonnées (PID). Les analyses bioinformatiques réalisées sur la base de leur signature peptidique ont révélé qu'environ 30% du génome des eucaryotes codaient pour des protéines appartenant à cette classe. Nombre de ces protéines assurent des fonctions de régulation et de signalisation cellulaire, en relation avec leur capacité à interagir avec différents partenaires. Les fonctions de ces protéines ne s'arrêtent pas là : elles interviennent également au niveau de fonctions mécaniques, d'assemblage de complexes multi-protéiques ou encore de piégeage de diverses molécules.

Les protéines riches en proline (PRP) présentes dans la salive appartiennent à cette classe des PID. Divisées en trois catégories, PRP acides, PRP basiques et PRP glycosylées, elles peuvent constituer jusqu'à 70% des protéines sécrétées par les glandes parotides. Les PRP basiques et glycosylées seraient susceptibles d'intervenir dans la sensation d'astringence qui survient en particulier lors de la consommation de produits d'origine végétale tels que le vin ou le thé. Cette sensation, qui se traduit par une impression de friction et de dessèchement dans la cavité buccale, est particulièrement associée à la consommation des tanins, molécules participant aux mécanismes de défense des plantes et appartenant au groupe des polyphénols. La répartition ubiquitaire et l'importance des tanins au sein du règne végétal font que l'astringence constitue une caractéristique importante de la qualité organoleptique des produits d'origine végétale. Si les mécanismes à l'origine de l'astringence restent à l'heure actuelle encore mal compris, l'implication des interactions entre tanins et PRP serait avérée (Baxter *et al.*, 1997). Ainsi, la problématique portant sur les relations existant entre les interactions PRP•tanin et l'astringence revêt à la fois un intérêt cognitif et un intérêt économique. Dans cette perspective, il est important d'obtenir des données sur les interactions s'établissant entre les tanins et les PRP salivaires.

Les développements apparus au cours des deux dernières décennies ont fait de la spectrométrie de masse (MS) une technique de choix dans l'étude des complexes non-covalents impliquant des protéines (Borch *et al.*, 2005). En effet, l'étude des interactions d'une protéine nécessite d'identifier les partenaires susceptibles d'interagir avec elle, de définir la structure des édifices supramoléculaires formés ainsi que de déterminer la stabilité et la cinétique d'association de ces complexes. La spectrométrie de masse est un outil puissant dans la résolution de tous ces points. L'électrospray (ESI) est une technique d'ionisation suffisamment douce pour permettre l'étude de larges complexes non-covalents (Robinson, 2002). La précision de cette technique permet d'identifier les partenaires d'interaction et de déterminer leur(s) stœchiométrie(s) d'interaction. La plupart des spectromètres de masse modernes permettent également de fragmenter sélectivement les objets détectés via une technique appelée « Collision Induced Dissociation » (CID). Cette technique permet de confirmer l'identification des objets sélectionnés grâce aux produits de dissociation détectés, d'étudier leur stabilité et d'établir leur cinétique de dissociation (Jørgensen *et al.*, 1999). D'autres techniques de fragmentation, comme l'« electron capture dissociation » (ECD), permettent de déterminer l'emplacement des sites d'interaction au niveau de la chaîne peptidique (Xie *et al.*, 2006). Enfin, le couplage de la spectrométrie de masse avec la mobilité ionique ajoute une dimension aux spectres de masse en séparant les ions en fonction de leur conformation tridimensionnelle (Dugourd *et al.*, 1997).

L'ensemble des travaux qui constituent ce mémoire porte sur l'étude des interactions PRP - tanin par spectrométrie de masse. Au cours de l'étude expérimentale, nous nous sommes particulièrement intéressés aux aspects suivants:

- la caractérisation des PRP modèles par MS, travaux préliminaires à la suite de l'étude ;
- la caractérisation de l'interaction de la PRP modèle, IB5, avec un tanin modèle, l'épigallocatechine 3-O-gallate (EgCG), un intérêt particulier étant porté aux stœchiométries, à la stabilité des complexes IB5•EgCG et à la spécificité de cette interaction ;
- le rôle joué par les différents motifs structuraux des tanins sur les stœchiométries et la stabilité des complexes IB5•tanins ;
- les caractéristiques structurales des sites d'interactions impliqués et les modifications de conformation induites par la fixation des tanins.

Avant d'exposer les résultats obtenus, il convient de faire l'état de l'art sur l'astringence, les partenaires impliqués (tanins et PRP qui appartiennent aux PID) et les techniques de spectrométrie de masse utilisées.

Etat de l'art

1. L'astringence

L'astringence est définie comme étant l'ensemble des sensations comprenant le dessèchement de la bouche, la rugosité des tissus buccaux ainsi qu'une impression d'étirement des joues et des muscles du visage (ASTM, 1989). Elle est ressentie lors de la consommation d'un certain nombre de produits d'origine végétale. Parmi ces produits, on citera le vin rouge, le thé ainsi qu'un certain nombre de fruits, en particulier lorsque ces derniers ne sont pas encore mûrs. C'est leur richesse en tanins qui est à l'origine de cette sensation.

1.1. Mécanisme physiologique de l'astringence

Du point de vue physiologique, l'astringence s'apparente plus à une sensation tactile qu'à un goût (Breslin *et al.*, 1993). La transmission nerveuse de la muqueuse buccale vers le cerveau se ferait via les nerfs trijumeaux (Green, 1993). Elle résulterait de l'activation de mécanorécepteurs due à des changements des caractéristiques de friction de la surface de l'épithélium buccal (Guinard et Mazzucchelli, 1996). Toutefois, l'activation de ces mécanorécepteurs n'est pas encore complètement comprise.

1.2. Astringence et évaluation sensorielle

En analyse sensorielle classique, les approches utilisées consistent soit à déterminer les seuils de perception de composés ou de fractions isolés, soit à réaliser une analyse descriptive de l'échantillon. Dans les études les plus récentes, l'utilisation d'approches descriptives a été préférée car elles permettent d'obtenir des valeurs d'intensité d'un caractère pour lequel des descripteurs ont été générés. Toutefois ces approches descriptives classiques ne donnent aucun renseignement sur la dynamique des sensations. Basée sur l'enregistrement de l'intensité d'un caractère en fonction du temps, la méthode dite « temps-intensité » (TI) permet de prendre en compte le caractère temporel d'un caractère isolé (Pangborn *et al.*, 1983 ; Robichaud et Noble, 1990 ; Lee et Lawless, 1991 ; Naish *et al.*, 1993). Le juré procède à l'enregistrement informatique de l'intensité de l'astringence qu'il ressent au cours du temps, générant une courbe TI (figure 1) (Guinard *et al.*, 1986b ; Robichaud et Noble, 1990 ; Lee et Lawless, 1991 ; Naish *et al.*, 1993 ; Fischer *et al.*, 1994).

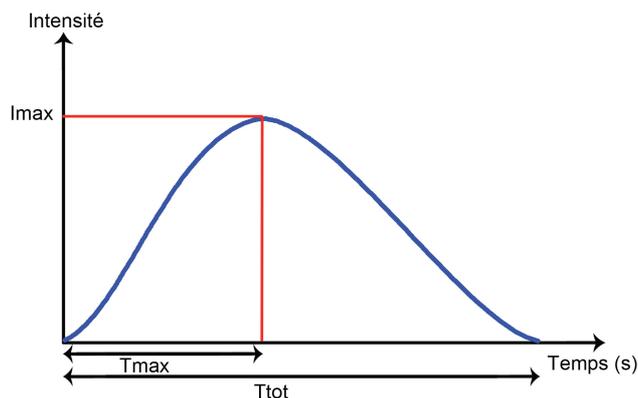


Figure 1. Exemple d'enregistrement de l'intensité ressentie en fonction du temps et paramètres calculés.

Ces courbes permettent d'obtenir plusieurs informations : l'intensité maximale (I_{max} en unités relatives), la durée totale de la perception (T_{tot} en s), le temps nécessaire avant de ressentir l'intensité maximale (T_{max} en s) et l'astringence totale (aire comprise sous la courbe).

La différenciation et la caractérisation de l'astringence peuvent varier d'un individu à l'autre. Cela pose un problème pour l'évaluation sensorielle de cette sensation. L'origine de ce problème peut être due à l'absence de définitions claires, renforcée par le fait que certains composés astringents peuvent présenter en même temps un caractère amer ou acide. Il est donc nécessaire d'avoir recours à un jury entraîné à distinguer ces différents caractères sensoriels. L'astringence reste également délicate à évaluer (Guinard *et al.*, 1986b ; Lyman et Green, 1990) en raison de son caractère cumulatif et rémanent. En effet, l'astringence est une sensation persistante et il est donc recommandé d'espacer dans le temps les différents échantillons à tester (quelques minutes) (Kallithraka *et al.*, 1997). Ceci permet de retrouver un état de lubrification à l'initial et d'éviter « l'effet mémoire » (Guinard *et al.*, 1986b).

1.3. Astringence et évolution

D'un point de vue écologique, l'astringence constitue un mécanisme d'évitement vis à vis d'aliments trop riches en tanins, ces derniers présentant des propriétés toxiques et des effets anti-nutritionnels.

2. Les Tanins

Le terme « tanins » désignait à l'origine les composés végétaux utilisés pour convertir les peaux animales, matière putrescible, en cuir, matière imputrescible. En effet, les tanins forment des complexes stables avec le collagène des peaux. En chimie, la première définition des tanins, proposée par Bate-Smith, précise que le terme « tanins » désigne les composés phénoliques hydrosolubles de masses molaires comprises entre 500 et 3000, capables de précipiter les alcaloïdes et les protéines (Bate-Smith et Swain, 1962). Les tanins présentent des structures variables comportant toujours une partie phénolique.

2.1. Tanins et structure

Les tanins appartiennent à la famille des polyphénols qui regroupe les composés possédant plusieurs groupements phénoliques. Les tanins, composés largement répandus dans le règne végétal, sont des métabolites secondaires qui présentent des fonctions de protection des plantes contre les parasites et les prédateurs. Ils sont très abondants dans certains organes végétaux consommés ou utilisés par l'homme, de nombreux fruits (pomme, prune, fraise...) ou des produits dérivés (thé, vin, cidre...).

Sur la base de leur structure chimique, les tanins sont classiquement divisés en deux groupes : les tanins hydrolysables et les tanins condensés. Les tanins condensés sont des oligomères et des polymères de flavan-3-ols qui sont liés par des liaisons C-C difficiles à hydrolyser. Les tanins hydrolysables sont facilement hydrolysables, que ce soit en condition acide ou basique ou en présence d'estérases.

2.1.1. Tanins hydrolysables

Les tanins hydrolysables sont divisés en deux sous-classes : les gallotanins et les ellagitanins. En effet, leur hydrolyse produit respectivement de l'acide gallique et de l'acide ellagique. Ils possèdent un noyau polyol (dans la plupart des cas le D-glucopyranose) dont les fonctions hydroxyles sont estérifiées par des unités d'acide gallique. Une forme simple de ce type de tanins est le pentagalloylglucose (figure 2), molécule très réactive à l'origine de la plupart des formes plus complexes (par exemple la castalagine chez le châtaignier).

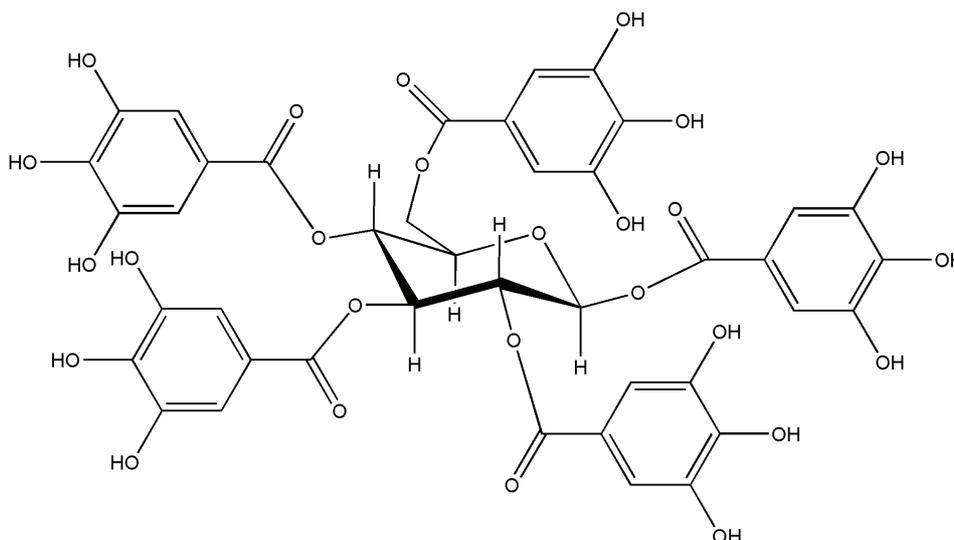


Figure 2. Structure d'un tanin hydrolysable, le pentagalloylglucose.

Les tanins hydrolysables sont relativement abondants dans le bois et l'écorce des arbres et des arbustes mais sont rares dans les autres organes végétaux. Ils sont également présents dans quelques fruits, comme la grenade et certaines baies du genre *Rubus*, telles que la framboise. On en trouve donc peu dans les aliments et les boissons à l'exception de celles qui ont été en contact avec du bois.

2.1.2. Tanins condensés ou proanthocyanidines

Le système de nomenclature le plus souvent utilisé par les auteurs a été initialement suggéré par Hemingway (Hemingway *et al.*, 1982) puis étendu par Porter (Porter, 1988).

Les tanins condensés (ou tanins catéchiques) font partie de la famille des flavonoïdes qui possède un squelette en C6-C3-C6 comprenant deux cycles aromatiques, les cycles A et B (C6) et un hétérocycle oxygéné, le cycle C (C3) (figure 3). Par traitement acide à chaud, ils se transforment en anthocyanidines, pigments rouges. Ils sont donc nommés *proanthocyanidines* (Porter, 1989).

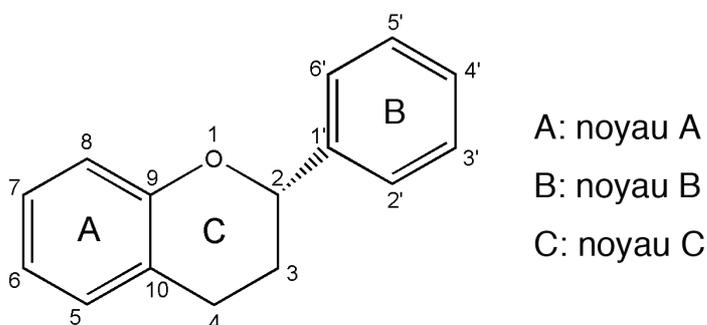


Figure 3. Squelette des monomères des proanthocyanidines.

Les proanthocyanidines sont des oligomères et des polymères (figure 4.b). Elles présentent une diversité structurale importante, relative à la nature des flavan-3-ols constitutifs, à leur nombre (longueur des chaînes ou degré de polymérisation noté DP) et à la position des liaisons qui relient ces unités (figure 4.b). La stéréochimie du carbone asymétrique C3 (catéchine et épicatechine) et les différents degrés d'hydroxylation sur les cycles A et B sont à l'origine de différents monomères. Les tanins condensés se divisent ainsi en plusieurs classes selon le schéma d'hydroxylation de leurs unités constitutives. Parmi elles, on peut citer les prodelphinidines, constituées d'unités (épi)gallocatéchine, et les procyanidines, dérivées de l'(épi)catéchine. De plus, certaines unités sont substituées par de l'acide gallique (figure 4.a).

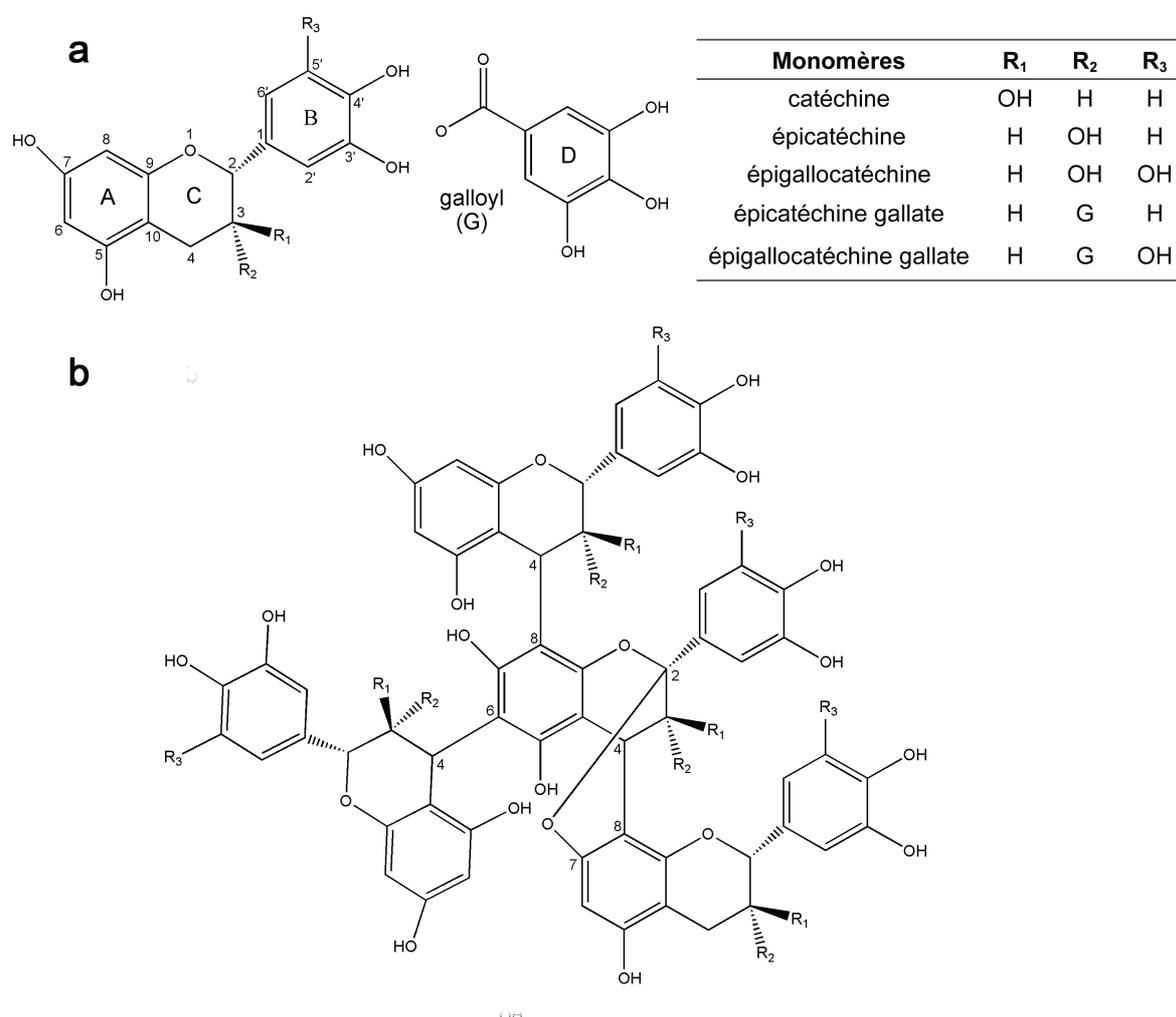


Figure 4. a. structure des monomères de flavan-3-ols ; b. exemple de tétramère comportant les deux types de liaisons interflavaniques de type B C4-C6 et C4-C8 et une liaison de type A C2-O-C7.

Lorsque les liaisons interflavaniques (inter-monomériques) sont de type C-C, positionnées en 4 sur l'unité supérieure et en 6 ou en 8 sur l'unité inférieure, les tanins sont

dits de type B (figure 4.b). Les tanins de la série A possèdent une liaison interflavanique supplémentaire de type éther entre les atomes de carbone 5 ou 7 de l'unité inférieure et l'atome de carbone 2 de l'unité supérieure. La polymérisation conduit à des structures de plus en plus complexes qui restent cependant hydrosolubles (Maury *et al.*, 2001b).

2.2. Tanins et astringence

L'astringence est une des caractéristiques des tanins. De nombreuses études ont été réalisées afin d'évaluer leur impact sur l'astringence (structure et quantité).

Une des principales caractéristiques structurales rapportée comme ayant une influence sur les propriétés astringentes des tanins est leur degré de polymérisation (DP). Il est généralement admis que les polyphénols de haute masse (>3000) ne présentent pas de caractère astringent du fait de leur insolubilité (Lea et Arnold, 1978). Cependant, des tanins d'une masse moyenne proche de 5000 se sont révélés solubles en solution hydro-alcoolique (Cheynier *et al.*, 1997) et ont été mis en évidence dans les vins (Sarni-Manchado *et al.*, 1999b ; Maury *et al.*, 2001b) où ils pourraient contribuer à l'astringence. L'astringence des tanins augmente généralement avec leur DP (Rossi et Singleton, 1966 ; Arnold *et al.*, 1980 ; Robichaud et Noble, 1990 ; Peleg *et al.*, 1999 ; Vidal *et al.*, 2003 ; Lesschaeve et Noble, 2005). Peleg *et al.* ont également montré une relation entre la position de la liaison interflavanique sur l'astringence ressentie. Le dimère de catéchine avec une liaison C4-C6 a en effet été ressenti comme plus astringent que son homologue présentant une liaison C4-C8 (Peleg *et al.*, 1999). La structure des unités monomères a également un impact sur l'astringence. Ainsi, l'estérification de l'hydroxyle du C3 par un acide gallique (galloylation) augmente le potentiel astringent des tanins (Vidal *et al.*, 2003). Dans cette étude, un score similaire d'astringence globale a été attribué à des tanins de DPm (DP moyen) 20 présentant 5% d'unités galloylées et à des tanins de DPm 10 présentant 22% d'unités galloylées (Vidal *et al.*, 2003). Ainsi, la présence d'unités galloylées peut compenser l'effet d'une diminution de DP sur l'astringence. Le niveau d'hydroxylation du noyau B et l'isomérisation des liaisons 2,3 sont deux autres caractéristiques structurales influençant l'astringence des tanins. Ceux présentant une isomérisation 2,3 *cis* sont ressentis comme plus astringents que leurs homologues ayant une isomérisation 2,3 *trans* (Thorngate et Noble, 1995 ; Kallithraka *et al.*, 1997 ; Peleg *et al.*, 1999).

On doit également noter que la matrice contenant les tanins peut avoir un impact sur l'astringence : la teneur en alcool, l'acidité et la viscosité de la matrice alimentaire sont rapportées comme ayant une influence sur l'astringence ressentie (Guinard *et al.*, 1986a ; Lyman et Green, 1990 ; Cheynier et Sarni-Manchado, 2006).

2.3. Tanins et écologie

La répartition ubiquitaire des tanins, leur abondance et leur capacité à interagir avec les protéines font qu'ils ont une importance économique et écologique considérable. Ils jouent un rôle majeur dans la perception organoleptique des produits d'origine végétale et finalement dans leur acceptation par le consommateur. D'un point de vue écologique, les végétaux peu riches en tanins sont généralement préférés par les herbivores. Les tanins, métabolites secondaires, exercent un rôle de défense également contre les micro-organismes (bactéries et mycètes) (Gonzalez-de-Colmenares *et al.*, 1998 ; Dixon *et al.*, 2005). Leur activité antimicrobienne serait en partie due à leur capacité à lier les molécules de fer grâce à la présence de fonction *ortho*-diphénol (*o*-diphénol : deux groupements hydroxyles sur des carbones voisins d'un cycle aromatique) et résulterait principalement de la capacité des tanins à se lier aux protéines et à gêner le fonctionnement de l'équipement enzymatique des organismes.

Chez les herbivores mammaliens, l'ingestion de tanins est connue pour être la cause de divers effets indésirables tels que la réduction de la digestibilité du bol protéique (Robbins *et al.*, 1987 ; Chung-MacCoubrey *et al.*, 1997 ; Shimada et Saitoh, 2003), l'altération de la muqueuse gastro-intestinale (Blytt *et al.*, 1988) et la perte d'azote endogène. Cette dernière est causée par l'excrétion des protéines complexées aux tanins dans les fèces (Blytt *et al.*, 1988 ; Shimada et Saitoh, 2003). Les tanins agiraient comme des inhibiteurs de la digestion protéique (Zucker, 1983) en se liant à la fois aux enzymes digestives et aux protéines du bol alimentaire. De plus, les acides phénoliques libérés lors de l'hydrolyse des tanins peuvent être absorbés au niveau de l'intestin et entraîner une toxicité sévère, telle que des nécroses et des défaillances du foie et des reins (Fowler et Richards, 1965 ; Niho *et al.*, 2001). L'ensemble de ces phénomènes conduit à considérer les tanins à la fois comme des inhibiteurs de la digestion et comme des toxines (Shimada, 2006) qui peuvent avoir des effets nocifs sévères sur les herbivores allant jusqu'à la mort (Mehansho *et al.*, 1987a). En accord avec la théorie de la reine rouge décrite par Van Valen¹ (Van Valen, 1973), les herbivores mammaliens auraient développé différents mécanismes de défense pour se protéger d'un régime riche en tanins. Ces mécanismes de défense peuvent être classés en deux catégories : les mécanismes pré-ingestion et les mécanismes post-ingestion. Les mécanismes pré-ingestion consistent à adopter des comportements d'évitement vis-à-vis d'aliments trop riches en tanins. Ce type de comportement a été décrit chez de nombreuses

¹ L'hypothèse de la reine rouge énoncée par Van Valen est un modèle de coévolution dans lequel proie et prédateur réalisent une course aux "armements" l'un envers l'autre.

familles d'herbivores mammaliens (Cooper et Owen-Smith, 1985 ; Clausen *et al.*, 1990 ; Shimada et Saitoh, 2003). Des comportements visant à contrôler le taux de tanins ingérés ont aussi été observés afin de ne pas surcharger les voies de détoxification (Dearing et Cork, 1999 ; Burritt et Provenza, 2000 ; Foley et Moore, 2005). La sensation d'astringence ressentie lors de l'ingestion d'aliments ou de boissons riches en tanins peut être considérée comme un mécanisme pré-ingestion. Les aliments présentant une astringence trop forte sont perçus comme désagréables et donc impropres à la consommation. Les mécanismes post-ingestion correspondent à la sécrétion de « tannin-binding salivary proteins » (TBSPs), à l'augmentation de la sécrétion de mucus gastro-intestinal, à la dégradation des tanins par la microflore intestinale, à l'activation d'enzymes détoxifiantes et à l'augmentation de la perméabilité gastro-intestinale aux glycoprotéines. Ces mécanismes sont supposés être les éléments clés de la protection des herbivores mammaliens contre les tanins (Shimada, 2006). Dans le but de mieux comprendre le mode d'action des tanins, différents travaux ont porté sur l'étude des interactions protéine-tanin.

2.4. Tanins et protéines

De nombreuses études ont été réalisées, afin de répondre à différentes questions relatives aux interactions protéine-tanin :

- quel est l'impact de la structure des tanins sur la formation des complexes ?
- quel est le rôle de la nature de la protéine, son comportement au cours de l'interaction ?
- quelle est la nature des liaisons qui s'établissent ?

Beaucoup de systèmes impliquant des protéines, des tanins et des milieux différents ont été étudiés. Un mécanisme général d'interaction a été dégagé. Dans celui-ci, il est admis qu'il existe une interaction réversible entre les polyphénols et les protéines en solution. Ainsi, un équilibre s'établit entre les complexes protéine•tanin solubles et les partenaires libres en solution. Si cet équilibre est modifié dans une seconde étape par l'augmentation de la concentration en l'un des deux partenaires, les complexes solubles peuvent s'agréger et précipiter (Luck *et al.*, 1994). La formation de ces objets insolubles est généralement réversible. Le plus souvent, ils peuvent être redissous suite à l'ajout d'un excès de protéine qui permet la redistribution des tanins entre les molécules protéiques (Haslam *et al.*, 1992).

L'ensemble des protéines est capable d'interagir avec les tanins. Cependant, il a été montré que toutes les protéines n'interagissent pas de la même façon avec les tanins (Hagerman et Butler, 1981). Ainsi, si l'on peut considérer que l'interaction tanin-protéine n'est

pas spécifique à une protéine ou à une famille de protéines particulière, une différence d'affinité envers les tanins existe entre les protéines. Des expériences de compétition vis à vis de tanins entre la « Bovine Serum Albumin » (BSA), protéine de référence, et des protéines de diverses origines ont montré que l'affinité des protéines pour les tanins pouvait varier d'un facteur 1000 (inhibition de 50% de la précipitation de la BSA) (Hagerman et Butler, 1981). Les auteurs concluent que la masse moléculaire, la conformation et la nature de la protéine influencent l'interaction protéine-tanin (Hagerman et Butler, 1981). Les protéines de haute masse moléculaire (>20 kDa) ont une affinité pour les tanins supérieure à celle des protéines de faible masse moléculaire. Cependant, des résultats contradictoires ont également été rapportés. Des protéines de faible masse moléculaire, comme les histatines salivaires, s'avèrent plus affines pour les tanins que les plus grosses protéines présentes dans la salive (Yan et Bennick, 1995). La même observation a été faite lors d'expériences de modélisation du processus de collage protéique. Le collage consiste à incorporer dans un vin plus ou moins trouble ou plus ou moins instable des « substances capables de flocculer et de sédimenter en entraînant les particules en suspension » (Caillet, 1994). Plusieurs objectifs sont poursuivis à travers la pratique du collage dont l'amélioration des caractéristiques organoleptiques du vin en éliminant par exemple certains tanins « trop astringents » (Salgues et Razungles, 1983). Les gélatines de masse molaire moyenne faible se sont révélées plus efficaces dans la formation de précipités protéine-tanin que des fractions de même nature mais de masse molaire plus élevée (Maury *et al.*, 2001a). En ce qui concerne la nature des protéines, les affinités les plus importantes ont été observées pour des protéines et des peptides riches en proline (PRP et pRP). De plus, il a été montré que les agrégats PRP•tanin insolubles ne se redissolvent pas (une fois formés) (Haslam *et al.*, 1992). La comparaison entre une PRP et un peptide composé de la séquence répétée de cette dernière a montré que la PRP entière a une plus grande affinité pour les tanins (Charlton *et al.*, 1996). Les auteurs ont suggéré que la PRP entière est capable de s'enrouler autour des tanins. Les modifications post-traductionnelles des protéines sont également susceptibles de modifier le déroulement de l'interaction. Ainsi, la glycosylation retarde la précipitation des protéines en augmentant la stabilité des complexes en solution (Asquith *et al.*, 1987 ; Sarni-Manchado *et al.*, 2008). Des résultats contradictoires concernant l'influence de la glycosylation des PRP sur leur affinité envers les tanins ont été publiés. Ainsi, la déglycosylation d'une PRP de souris a diminué son affinité envers les tanins (Asquith *et al.*, 1987), alors que le résultat contraire a été obtenu sur une PRP humaine (Lu et Bennick, 1998). Cette différence de résultats est certainement due aux méthodes utilisées qui ne sont pas comparables (il s'agit dans un cas d'expériences de compétition avec la BSA, basées sur la détermination de la concentration de protéine nécessaire pour inhiber 50% de la

précipitation de la BSA, alors que dans l'autre cas c'est la quantité de tanins précipités qui est évaluée).

2.4.1. Effet de la structure du tanin

Le rôle de la structure des tanins, qui influence leurs propriétés astringentes, est aussi évoqué lors de leurs interactions avec les protéines.

L'affinité des proanthocyanidines pour les protéines augmente avec leur nombre d'unités constitutives (Prieur *et al.*, 1994 ; Baxter *et al.*, 1997 ; Sarni-Manchado *et al.*, 1999a) et leur taux de galloylation (Guyot *et al.*, 1996 ; Baxter *et al.*, 1997 ; Zhu *et al.*, 1997 ; Sarni-Manchado et Cheynier, 2002). Ces observations ont conduit à considérer l'affinité des tanins pour les protéines comme dépendant du nombre de fonctions *o*-diphénols portées par les tanins. Ces fonctions chimiques semblent être les principaux sites d'interactions des tanins avec les protéines (Haslam, 1974 ; McManus *et al.*, 1985 ; Haslam et Lilley, 1988). La présence d'un troisième groupement hydroxyle sur le noyau B augmente également l'affinité des tanins pour les protéines : les cycles aromatiques portant trois fonctions hydroxyles vicinales semblent particulièrement actifs (Spencer *et al.*, 1988 ; Hagerman, 1989 ; Siebert, 1999). Il apparaît que le nombre et la position des groupements hydroxyles portés par les tanins sont des paramètres importants dans ces interactions (Bacon et Rhodes, 1998). L'établissement des interactions semble aussi sous la dépendance de la configuration spatiale adoptée par les polyphénols (McManus *et al.*, 1985). En effet, les molécules présentant une liaison interflavanique C4-C6 (figure 4.b) ont une affinité pour les protéines plus forte que celle de leurs régio-isomères à liaison C4-C8 (Ezaki-Furuichi *et al.*, 1987 ; Ricardo da Silva *et al.*, 1991a). De plus, les monomères ayant une configuration 2,3 *cis* auraient une affinité pour les protéines légèrement supérieure à leurs isomères 2,3 *trans* (Bacon et Rhodes, 1998).

Il faut signaler que même si les monomères sont capables d'interagir avec les protéines, le phénomène de précipitation est théoriquement limité aux tanins ayant une masse moléculaire supérieure à 500 (à partir des dimères) (Gustavson, 1956 ; Haslam, 1974). En effet, l'augmentation du nombre de noyaux phénoliques autoriserait la formation de « ponts » entre les molécules protéiques (McManus *et al.*, 1985 ; Siebert, 2006) ou entre les polyphénols (Baxter *et al.*, 1997) conduisant à la formation d'agrégats, puis à leur précipitation. Par conséquent, les tanins doivent être perçus comme des ligands ayant la capacité d'interagir avec plusieurs sites de liaison simultanément (Haslam, 1996).

2.4.2. Nature des interactions

Malgré le grand nombre de travaux réalisés, les mécanismes et la nature des interactions protéines-tanins sont encore mal compris et sujets à discussion.

Afin d'expliquer les fortes interactions des tanins avec les gélatines qui sont des protéines chargées positivement et utilisées dans le collage des vins, l'implication d'un mécanisme ionique a été invoquée. Cependant l'augmentation du pH entraîne une diminution de la précipitation malgré la plus forte concentration en polyphénols présents sous forme ionisée : ce processus est donc improbable (Hagerman et Butler, 1978 ; McManus *et al.*, 1985). Il a été montré qu'au pH du vin (3,5), les tanins ne portent pas de charge : ainsi les mécanismes ioniques et électrostatiques ne jouent qu'un rôle négligeable dans ces interactions (Vernhet *et al.*, 1996).

Les tanins oxydés en *ortho*-quinone, particulièrement réactifs, sont capables d'établir des liaisons covalentes avec les protéines. Les semi-quinones peuvent aussi être à l'origine de liaisons covalentes par couplage radicalaire avec un radical porté par la protéine. Cependant, la possible redissolution des précipités et des troubles exclut l'intervention de liaisons covalentes dans ces interactions. Les effets hydrophobes et les liaisons hydrogène sont les deux principaux types d'interactions non-covalentes évoqués dans ces systèmes. Certains auteurs parlent d'une implication simultanée de ces deux types d'interactions (Spencer *et al.*, 1988 ; Murray *et al.*, 1994). La part de ces mécanismes est alors déterminée par la nature des deux partenaires (Artz *et al.*, 1987 ; Hagerman *et al.*, 1998). L'existence unique de liaisons hydrogène a également été rapportée dans certains systèmes (Hagerman et Butler, 1980), notamment en présence d'éthanol (Simon *et al.*, 2003).

2.4.3. Influence de l'environnement

La présence d'autres molécules peut modifier les phénomènes d'association et de précipitation. Ainsi, l'interaction de certains polysaccharides comme les xanthanes, la gomme arabique ou les pectines solubles avec le pentagalloylglucose prévient sa précipitation par les gélatines (Luck *et al.*, 1994). Ce phénomène s'explique, soit par la formation de complexes ternaires solubles, soit par l'existence d'une compétition entre polysaccharides et protéines vis-à-vis des tanins.

De plus, certains paramètres environnementaux comme le pH, la force ionique ou la teneur en éthanol influent sur la précipitation et l'agrégation (Poncet-Legrand *et al.*, 2003).

Ceci nous amène à considérer la salive, premier milieu avec lequel les tanins entrent en contact lors de leur ingestion. C'est également dans ce milieu que se déroule l'interaction PRP•tanin et où est générée la sensation d'astringence.

3. La salive

La salive est la première sécrétion digestive. C'est un milieu aqueux dilué qui remplit de nombreuses fonctions.

3.1. La salive : composition et sécrétion

La composition de la salive est difficile à déterminer car elle varie avec le débit salivaire, selon la présence et la nature d'une éventuelle stimulation et l'heure du prélèvement. La salive totale ou salive mixte est composée de l'ensemble des salives sécrétées au sein de la cavité buccale principalement par les glandes parotides, sub-linguales et sous-maxillaires. On distingue deux types de salive : la salive en conditions basales et la salive stimulée. Dans les conditions de repos, le débit salivaire est assuré à 70%, 20% et 5% respectivement par les glandes sous-maxillaires, parotides et sub-linguales. En conditions stimulées, ces proportions changent. La contribution des glandes parotides augmente fortement et peut atteindre jusqu'à 60% de la sécrétion. Les diminutions consécutives des contributions des glandes sub-linguales et sous-maxillaires sont respectivement de 2% et de plus de 40% dans le cas d'une mastication. Il n'est pas surprenant d'observer des différences notables de cette sécrétion en fonction du moment de la journée chez un même individu. Ainsi, la mastication stimule la sécrétion salivaire. De plus, l'observation de variations de sécrétion, d'un jour à l'autre mais également d'une saison à l'autre, suggère une modulation de la sécrétion salivaire par le régime alimentaire, comme cela a été démontré chez les rongeurs (Mehansho *et al.*, 1992 ; Skopec *et al.*, 2004).

La salive est composée de 99% d'eau, les autres constituants étant principalement des sels et des protéines. Les principaux électrolytes sont le sodium, le chlore, le potassium, les phosphates, le calcium et les bicarbonates. Les ions bicarbonates tamponnent la salive à un pH compris entre 6,8 et 7,4, tandis que les ions calcium et phosphates assurent la minéralisation des dents (Dodds *et al.*, 2005). La composition ionique de la salive primaire (avant excrétion par passage dans les canaux excréteurs) est proche de celle du plasma. Les canaux excréteurs sont peu perméables à l'eau et les échanges ioniques (réabsorption active de Na⁺) aboutissent à la formation d'un liquide hypotonique par rapport au plasma.

La concentration en protéines de la salive est comprise entre 1 et 3,5 mg.mL⁻¹ (Bennick, 1982), même si des valeurs supérieures ont été rapportées (Clifford, 1997 ; Ferguson, 1999). Un tiers de ces protéines est issu de la sécrétion des glandes sous maxillaires mais l'essentiel provient des parotides. Parmi ces protéines, les plus abondantes sont les Protéines Riches en Proline (PRP) qui peuvent représenter jusqu'à 70% des

protéines sécrétées par les glandes parotides. On trouve aussi de l' α -amylase, des mucines, des statherines, des histatines et différentes protéines enzymatiques.

3.2. La salive : fonctions

La salive a de nombreuses fonctions dont beaucoup sont assurées par des protéines. Elle participe à la digestion, la dilution et la clairance de nombreux composés. De plus, la salive intervient également au niveau de la minéralisation de l'émail et de la protection contre les caries. La salive participe aussi au goût en mettant les substances alimentaires en solution. Son rôle dans les mécanismes de défense de l'organisme est majoritairement un rôle anti-microbien. Sa principale fonction est de protéger la muqueuse buccale (Matsuo, 2000 ; Humphrey et Williamson, 2001). Cette propriété est imputable essentiellement à la présence de mucines et de protéines riches en proline glycosylées (gPRP) (Hatton *et al.*, 1985 ; Breslin *et al.*, 1993).

Parmi les protéines jouant un rôle au niveau de la digestion, on trouve quelques protéases ainsi que l' α -amylase, qui assure l'hydrolyse de l'amidon (Edgar, 1992).

On peut citer, parmi les protéines salivaires « anti-bactériennes », des immunoglobulines comme les IgA qui sont l'un des principaux éléments de la réponse humorale (Fagarasan et Honjo, 2003), et le lysosyme qui dégrade les parois bactériennes. La lactoferrine, glycoprotéine multifonctionnelle de la famille des transferrines qui se lie au fer, présente des effets bactériostatiques et bactéricides (Ward *et al.*, 2005). Les mucines MG2 interagissent avec les microorganismes (Tabak, 1995) et les histatines se sont révélées des antifongiques très actifs en particulier contre le genre *Candida* (Tsai et Bobek, 1998).

Parmi les protéines jouant un rôle dans la protection des dents, les statherines maintiennent une sursaturation en calcium et en phosphates.

Enfin, les protéines riches en proline (PRP) constituent une famille importante de protéines salivaires. Elles apparaissent comme des protéines multifonctionnelles remplissant de nombreuses fonctions au niveau de la salive.

4. Les protéines riches en proline ou PRP

Les PRP sont présentes chez de nombreuses espèces de mammifères.

4.1. PRP : aspect génétique

Les gènes de ces protéines ont été répertoriés chez plusieurs espèces de mammifères (souris (Ann et Carlson, 1985), hamster (Ann *et al.*, 1987), macaque (Oppenheim *et al.*, 1987), homme (Kim et Maeda, 1986)). Des homologies de séquences (géniques et protéiques) ont été observées entre les PRP salivaires de différents mammifères dont l'homme (Bennick, 1982 ; Bennick, 2002) suggérant l'existence d'un ancêtre commun à ces gènes.

Chez l'homme, l'ensemble des PRP synthétisées est codé par seulement 6 gènes (PRB1, PRB2, PRB3, PRB4, PRH1 et PRH2), tous situés sur le chromosome 12 (Maeda, 1985 ; Maeda *et al.*, 1985). Parmi ces gènes qui semblent avoir un ancêtre commun (Azen *et al.*, 1996 ; Stubbs *et al.*, 1998), deux codent pour des PRP acides (PRH1 et PRH2) et quatre pour des PRP basiques et glycosylées (PRB1, PRB2, PRB3 et PRB4) (Kim et Maeda, 1986). Le fort polymorphisme notamment au niveau de leur longueur est dû à de nombreux crossing over intra et inter géniques (Kim et Maeda, 1986 ; Lyons *et al.*, 1988b ; Kim *et al.*, 1990 ; Kim *et al.*, 1993 ; Azen *et al.*, 1996). Ces gènes sont présents sous forme de nombreux allèles dans la population. Ils codent pour des pro-protéines qui subissent divers processus de maturation au cours de leur excrétion. Ces modifications post-traductionnelles (glycosylation, phosphorylation et clivage) conduisent à la sécrétion d'une vingtaine de PRP dans la salive (Lyons *et al.*, 1988a ; Bennick, 2002).

4.2. PRP : caractéristiques et structures

Comme leur nom l'indique, les PRP montrent une richesse "anormale" en proline. Alors que cet acide aminé ne représente que 5% des acides aminés totaux chez les protéines structurées, un minimum de 20% est rencontré chez les PRP. Ce taux peut atteindre 40% pour les PRPs de certaines espèces comme l'homme, le rat et la souris (Kauffman et Keller, 1979 ; Mehansho et Carlson, 1983 ; Mehansho *et al.*, 1985 ; Mole *et al.*, 1990). La présence accrue de glycine, acide glutamique et glutamine est aussi une caractéristique de ces protéines. Le taux de ces quatre acides aminés avoisine les 70 à 90% (Mole *et al.*, 1990). Ces protéines ont ainsi un index de complexité très faible. Ces quatre acides aminés qui empêchent la formation de structures secondaires chez ces protéines (Dunker *et al.*, 2001) sont qualifiés de casseurs d'hélice. De plus, chez les PRP, ces acides aminés sont organisés sous la forme de nombreuses séquences répétées au sein de la chaîne peptidique.

Sur la base de la valeur de leur point isoélectrique (pI) et de leur éventuelle glycosylation, les PRP salivaires sont divisées en trois classes, les PRP acides (*aPRP*), les PRP basiques (*bPRP*) et les PRP glycosylées (*gPRP*).

4.3. PRP : fonctions

La fixation des tanins est rapportée comme étant la fonction majeure des PRP salivaires. Une des observations appuyant le mieux cette hypothèse vient d'expériences au cours desquelles des rats (Mehansho et Carlson, 1983), des souris et des hamsters (Mehansho *et al.*, 1985) ont été soumis à des régimes riches en tanins. Chez les deux premières espèces, l'exposition à ce régime a d'abord entraîné une perte de poids qui a ensuite été récupérée. Le début de la reprise de poids est corrélé avec l'expression des ARNm codant pour les PRP (Mehansho et Carlson, 1983 ; Mehansho *et al.*, 1985 ; Mehansho *et al.*, 1992). Il est à noter que les PRP sécrétées par les rats dans ces conditions présentent une forte affinité pour les tanins. A l'inverse, chez des hamsters incapables de stimuler la synthèse de ces protéines en réponse à ce régime, la perte de poids n'est pas compensée (Mehansho *et al.*, 1987a).

L'affinité des PRP pour les tanins est supérieure à celle des autres protéines étudiées : 5 à 80 fois supérieure à celle de la BSA et 1000 fois supérieure à celle du lysosyme (Hagerman et Butler, 1981 ; Asquith *et al.*, 1987 ; Austin *et al.*, 1989 ; Mole *et al.*, 1990 ; Bacon et Rhodes, 1998). Cette interaction préférentielle entre tanins et PRP permettrait à ces dernières d'agir comme une première ligne de défense de l'organisme contre les tanins. En effet, en présence d'un excès d'autres protéines ayant une plus faible affinité, les PRP sont capables d'établir des complexes stables avec les tanins dans la cavité orale (Asquith et Butler, 1985). Ainsi, au sein de la salive, les PRP protègent les autres protéines, leur permettant d'assurer leurs fonctions (Sarni-Manchado *et al.*, 2008) et de limiter la fixation des tanins sur l'épithélium buccal (Robbins *et al.*, 1991 ; Hagerman et Robbins, 1993). Ces complexes restent stables dans le tractus gastro-intestinal (Bennick, 1982) et sont excrétés dans les fèces (Mitaru *et al.*, 1984 ; Skopec *et al.*, 2004). Les enzymes gastro-intestinales sont ainsi protégées. Une des objections à ce rôle provient de l'interrogation concernant l'intérêt de sécréter une protéine qui sera perdue dans les fèces, entraînant une perte d'azote pour l'organisme. Pour répondre à cette objection, il est intéressant d'étudier le rapport bénéfice/coût et de noter que les quatre acides aminés qui composent principalement les PRP ne sont pas des acides aminés essentiels pour les mammifères. Ainsi des acides aminés non essentiels sont excrétés à la place d'acides aminés essentiels (McArthur *et al.*, 1995). Une autre objection à cette hypothèse vient du fait que certaines

espèces produisent des protéines qui sont riches en proline mais qui ne montrent pas une forte affinité pour les tanins (Mole *et al.*, 1990).

Ceci nous amène à revenir sur le fait qu'il existe différents types de PRP qui ont des affinités différentes pour les tanins et des rôles différents (Bennick, 2002).

La fonction principale des *bPRP* serait la fixation des tanins afin de protéger l'organisme de leur effets nocifs (Mehansho et Carlson, 1983 ; Yan et Bennick, 1995 ; Beeley, 2001 ; Bennick, 2002). Elles montrent une affinité plus importante pour les tanins que les *aPRP* (Lu et Bennick, 1998).

Les *aPRP* interviendraient dans la régulation de l'homéostasie calcique buccale (Hay *et al.*, 1987). En effet, elles sont capables de lier le calcium grâce à leur extrémité N-term très acide (Bennick *et al.*, 1981). Elles interviendraient également dans la liaison de l'hydroxyapatite, au cours de laquelle elles peuvent se retrouver intégrées à la plaque dentaire à la surface de l'émail (Bennick *et al.*, 1979 ; Bennick *et al.*, 1983).

Les PRP glycosylées (*gPRP*) interviennent dans la lubrification de la cavité buccale (Hatton *et al.*, 1985). Elles protègent les muqueuses délicates de la bouche contre l'abrasion en formant une pellicule lubrifiante à leur surface (Tabak, 1995) et assurent aussi la lubrification des aliments pour faciliter leur mastication et leur déglutition. Ces protéines sont également capables de lier certains micro-organismes (Murray *et al.*, 1992). Cela pourrait faciliter l'adhérence des bactéries à la surface de la bouche ou permettre de nettoyer la bouche des bactéries, suivant que les *gPRP* soient présentes à la surface de la plaque dentaire ou dans la salive (Gillece-Castro *et al.*, 1991 ; Azen *et al.*, 1993).

Au cours de l'évolution, les *aPRP* seraient apparues en premier pour réguler l'homéostasie orale. Ensuite, afin de répondre au stress causé par une alimentation plus riche en tanins, les *bPRP* seraient apparues (McArthur *et al.*, 1995).

4.4. PRP : synthèse

Selon les espèces, la synthèse des PRP peut être induite par l'injection de β -agoniste et l'ingestion de tanins (Shimada, 2006). Chez le rat et la souris, cette induction cesse au bout de quelques jours après l'arrêt des stimuli (Haghighat *et al.*, 1996 ; Shimada, 2006). Cela suggère que la production de ces protéines représente un coût pour l'organisme, en particulier en azote (Skopec *et al.*, 2004). Cependant, ce mécanisme de régulation de la synthèse des PRP n'a été montré que chez des rongeurs et ne semble pas exister chez d'autres espèces comme l'homme. Le bénéfice d'un système de régulation de la synthèse des PRP pourrait être primordial pour des espèces qui changent de régime alimentaire

régulièrement. Chez d'autres espèces telles que l'homme la présence continue dans leur alimentation de tanins serait à l'origine d'une synthèse constitutive importante sans mécanisme de régulation (Ann *et al.*, 1997).

L'hypothèse d'une synthèse importante de protéines salivaires capables de fixer les tanins chez les animaux qui rencontrent de fortes concentrations en tanins dans leur régime alimentaire a été émise. En effet, d'autres protéines salivaires telles les histatines peuvent intervenir dans la fixation des tanins (Yan et Bennick, 1995). Elles forment avec les PRP les deux principales classes de « Tannin-Binding Salivary Proteins » (TBSP). Cependant, les histatines ne représentent que 2,6% des protéines salivaires totales. Les PRP restent donc les principales protéines impliquées dans la défense contre les tanins. La présence de TBSP varie d'une espèce à l'autre chez les mammifères. Les feuilles de monocotylédones ainsi que les plantes herbacées sont moins riches en tanins que les fruits, les graines, les feuilles des arbres et des arbustes (Waterman et Mole, 1994). L'hypothèse initiale implique donc que les frugivores, les rongeurs et les omnivores produisent des TBSP en quantité importante, mais qu'à l'inverse la synthèse de ces protéines est réduite chez les herbivores. Cette hypothèse est en général confirmée (Shimada, 2006). Les espèces les plus exposées sont celles qui synthétisent le plus de TBSP (McArthur *et al.*, 1995 ; Fickel *et al.*, 1998). Par exemple, deux espèces de rhinocéros vivant dans deux biotopes différents présentent des différences de pool de TBSP. L'espèce de rhinocéros la moins exposée aux tanins produit le moins de TBSP (Clauss *et al.*, 2005).

4.5. PRP : relation avec l'astringence

Les interactions PRP-tanin qui s'établissent au niveau de la cavité buccale semblent être impliquées dans la sensation d'astringence. Deux théories principales sont proposées pour expliquer l'astringence. Tout d'abord, l'interaction des tanins avec les PRP réduirait le pouvoir de lubrification de ces dernières, soit via leur précipitation, soit via des modifications structurales de ces protéines à l'issue de cette interaction (Bate-Smith, 1973). La diminution de l'état de lubrification serait à l'origine d'une augmentation de la friction au niveau de l'épithélium buccal (Green, 1993 ; Clifford, 1997). La deuxième théorie postule que c'est l'interaction directe des tanins au niveau de l'épithélium qui est à l'origine de cette sensation. Dans ce cas, les PRP joueraient un rôle de protection et de prévention de l'astringence : l'interaction des composés astringents avec les PRP empêcherait leur fixation sur la muqueuse buccale (Horne *et al.*, 2002). En faveur de cette thèse, les récents travaux de Nayak et Carpenter ont montré que la perception de l'astringence diminuait avec l'augmentation de la quantité de protéines salivaires présentes dans la bouche et

inversement (Nayak et Carpenter, 2008). L'astringence résulterait donc, au moins pour partie, de l'interaction des tanins avec les cellules de la muqueuse buccale. Cette hypothèse est confortée par le fait que les flavanols monomères peuvent fixer les protéines mais sont incapables de les faire précipiter (Gustavson, 1956 ; Haslam, 1974) excepté à de fortes concentrations (McManus *et al.*, 1985), alors qu'ils sont capables d'engendrer une sensation d'astringence (Scharbert *et al.*, 2004). Cependant ces deux théories ne s'excluent pas, les deux phénomènes pouvant prendre part à la sensation. Il est donc envisageable que l'astringence résulte à la fois de la perte du pouvoir de lubrification des gPRP et de la fixation directe des tanins sur l'épithélium buccal après saturation des capacités de fixation des bPRP.

Les PRP sont des protéines multifonctionnelles qui appartiennent au groupe des protéines non structurées. Les protéines de ce groupe présentent des caractéristiques leur permettant d'assurer des fonctions qui ne sont pas accessibles aux protéines structurées.

5. Les protéines non structurées : IUP, IDP ou PID

Les noms de « Intrinsically Unstructured Protein » (IUP), de « Intrinsically Disordered Protein » (IDP) et de « Protéines Intrinsèquement Désordonnées » (PID) font référence à l'état natif non structuré des protéines appartenant à ce groupe. Ces protéines sont communes dans les trois domaines du vivant (Eubactéries, Archéobactéries et Eucaryotes). Les archéobactéries et les eubactéries présentent un taux de ces protéines relativement voisin, mais plus faible que chez les eucaryotes (tableau 1) (Dunker *et al.*, 2000a). La prédiction de la présence de désordre est supérieure chez les eucaryotes multicellulaires que chez les eucaryotes unicellulaires.

Tableau 1. Prédiction de la présence de désordre au sein du protéome des 3 règnes du vivant. Les pourcentages indiqués correspondent aux valeurs les plus hautes et les plus basses observées (rem : seules les protéines ayant plus de 60 acides aminés ont été prises en compte).

	eubactéries	archéobactéries	eucaryotes
% de protéines prédites pour avoir des régions désordonnées > à 50 acides aminés chez les espèces étudiées	2 à 21%	4 à 24%	27 à 41%
% de protéines complètement déstructurées chez les espèces étudiées	1 à 8%	2 à 11%	3 à 17%

Le taux de régions non structurées au sein du protéome augmente donc avec le niveau d'organisation du vivant. Chez les mammifères, qui apparaissent au sommet de l'arbre phylogénétique du vivant, environ 75% des protéines de signalisation sont prédites avec de longues régions désordonnées (> 30 résidus), 50% des protéines totales sont prédites avec de telles régions et environ 25% d'entre elles seraient complètement déstructurées (Dunker *et al.*, 2008). L'importance fonctionnelle de cet état non structuré est confortée par le fait que ces protéines assurent des fonctions de régulation de base dans des processus cellulaires clés (Tompa, 2002).

Si cette catégorie de protéines a été si longtemps méconnue malgré les fonctions importantes assurées par ses membres, c'est parce que les techniques de purification et de caractérisation étaient classiquement adaptées aux protéines structurées. Le développement de la génomique a permis de mettre en évidence l'importance de ces protéines (Dyson et Wright, 2005). De plus, l'émergence de ce groupe a suscité de nombreux débats car cette nouvelle catégorie de protéines remettait en cause le paradigme séquence-structure-fonction.

5.1. PID : le réaménagement du paradigme séquence-structure-fonction

Le paradigme séquence-structure-fonction classique stipulait que la fonction d'une protéine était liée à sa structure tridimensionnelle émanant de sa séquence primaire. Les protéines intrinsèquement désordonnées (PID) ou "protéines non repliées à l'état natif" sont des protéines fonctionnelles dépourvues de structures secondaire et tertiaire stables en conditions physiologiques en l'absence d'un partenaire, et dont les fonctions nécessitent la présence de régions non structurées plus ou moins grandes. L'incapacité à adopter une structure 3D précise confère à ces protéines une grande flexibilité structurale qui leur permet d'interagir avec plusieurs partenaires structurellement différents, ce qui entraîne des impacts biologiques multiples. Ces interactions sont caractérisées par des vitesses d'association et de dissociation rapides. La nature transitoire de ces interactions est particulièrement bien adaptée à des phénomènes de régulation qui nécessitent des réponses rapides à des changements environnementaux. Les PID sont beaucoup plus étendues en solution que les protéines globulaires, une propriété qui leur permet d'agir en tant que "linkers", de recruter des partenaires sur des complexes macromoléculaires de taille importante et d'exercer des effets de volume exclu.

L'émergence des PID a nécessité de réaménager le paradigme séquence-structure-fonction (Wright et Dyson, 1999). Dunker a proposé de redéfinir ce paradigme en une trinité :

la trinité structure-fonction (Dunker et Obradovic, 2001). Selon celle-ci, les protéines peuvent exister nativement en trois états différents, un état complètement replié, un état partiellement replié et un état complètement déstructuré (Dunker *et al.*, 2002). L'état natif d'une protéine dépend de sa séquence primaire, de la présence ou non de modification(s) post-traductionnelle(s), de la présence ou non de ligand et des conditions du milieu (Dyson et Wright, 2005). Ces différents paramètres sont sous un contrôle dynamique de la cellule, qui pourra alors ajuster l'état de ses protéines (Dunker *et al.*, 2002). Ainsi, dans leur milieu natif, les protéines sont présentes sous un continuum dynamique de structures, qui évolue en fonction de l'environnement sur la base de la structure induite par la séquence peptidique primaire (figure 5) (Dyson et Wright, 2005).

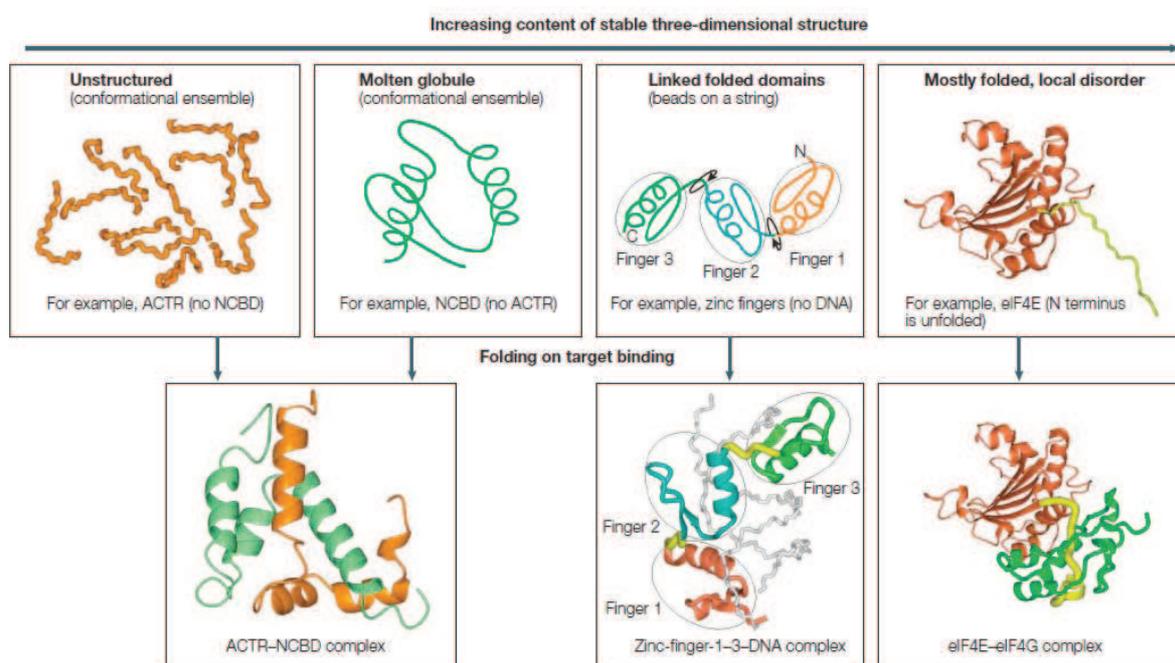


Figure 5. Le continuum des structures des protéines, tiré de (Dyson et Wright, 2005).

Les protéines non structurées présentent plusieurs caractéristiques qui les différencient des autres protéines.

5.2. PID : caractéristiques

Les PID se caractérisent par l'absence presque complète d'une structure tridimensionnelle repliée de manière précise (Wright et Dyson, 1999 ; Dunker *et al.*, 2002 ; Uversky, 2002a), l'adoption d'une conformation étendue avec une grande flexibilité intramoléculaire et la faible présence de structure secondaire (Wright et Dyson, 1999 ; Dunker *et al.*, 2002 ; Tompa, 2002 ; Uversky, 2002a). Cette inhabituelle flexibilité de la chaîne polypeptidique a été démontrée par l'absence de coordonnées de la chaîne dans les

structures cristallographiques et par une diminution de la distribution des déplacements chimiques sur les spectres RMN (Résonance Magnétique Nucléaire) de ces protéines. Leur manque de structure secondaire a été montré par des expériences de dichroïsme circulaire (circular dichroism, CD) dans l'UV lointain et de spectroscopie infra-rouge (Fourier Transform Infra-Red, FTIR). Les expériences de diffusion des rayons X aux petits angles (Small-Angle X-ray Scattering, SAXS) et de diffusion dynamique de la lumière (Dynamic Light Scattering, DLS) ont montré leur conformation ouverte et étendue, ainsi que l'absence d'un cœur hydrophobe replié. Ces résultats ont été corroborés par des observations indirectes comme une sensibilité particulière à la protéolyse, une forte stabilité thermique, un comportement aberrant en gel filtration et une mobilité inhabituelle en électrophorèse en gel de polyacrylamide en conditions dénaturantes (Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis : SDS-PAGE).

Les PID montrent une richesse particulière en certains acides aminés comme la proline ou la glutamine. Plus précisément, comparées aux protéines structurées, les PID présentent un appauvrissement significatif en Cys, Trp, Tyr, Phe, Ile, Val et Leu et un enrichissement particulier en Met, Lys, Arg, Ser, Gln, Pro et Glu (Dunker *et al.*, 2001 ; Tompa, 2005 ; Dunker *et al.*, 2008).

Les PID présentent également une distribution particulière des angles dièdres phi et psi de leurs résidus sur le diagramme de Ramachandran mettant en évidence la coexistence d'un ensemble de conformères s'interconvertissant (Dunker *et al.*, 2002 ; Receveur-Brechot *et al.*, 2006).

5.3. PID : fonctions

Les PID n'assurent pas de fonctions enzymatiques car ces dernières nécessitent la présence d'un site catalytique ayant une structure tridimensionnelle précise. Dunker a recensé vingt neuf fonctions assurées par les PID (Dunker *et al.*, 2002). Sur cette base, Tompa (Tompa, 2002 ; Tompa, 2003a) a défini cinq classes (tableau 2).

Tableau 2. Les cinq grandes classes de PID.

Classes	Traduction	Fonctions
entropic chain	chaînes entropiques	leurs fonctions prennent directement racine dans le désordre structural, leur permettant d'assurer par exemple des fonctions mécaniques telle que celles d'un ressort.
effectors	effecteurs	modifier l'activité d'une protéine ou d'un complexe protéique fonctionnel.
scavengers	collecteurs	neutraliser ou stocker des ligands.
assemblers	assembleurs	assembler, stabiliser et réguler de larges complexes protéiques tels les ribosomes.
display sites	exposition de sites de modification post-traductionnelle	réguler les modifications post-traductionnelles. Ces sites remarquables sont préférentiellement reconnus par les enzymes assurant les modifications post traductionnelles comme la phosphorylation ou la protéolyse limitée.

L'absence de structure présente un certain nombre d'avantages fonctionnels pour les PID.

5.4. PID : bénéfice fonctionnel

Les protéines des différentes classes fonctionnelles de PID décrites ci-dessus présentent une flexibilité et une malléabilité structurales leur permettant d'assurer des fonctions qui leur seraient bien souvent impossibles à réaliser si elles étaient structurées (Fink, 2005). Cet avantage est particulièrement flagrant dans la classe des chaînes entropiques. Il leur permet d'assurer des fonctions d'espaceur (MAPs et neurofilament H) ou de passeur (nucléoporines). Chez les protéines ayant plusieurs domaines globulaires liés par un domaine flexible nommé linker (Zhou, 2003), la présence de ce dernier autorise la gestion de la distance entre deux domaines et leur alloue une plus grande liberté dans leur recherche d'orientation. Ainsi, ces protéines peuvent reconnaître au niveau de leur ligand des domaines séparés.

La structure étendue des PID leur permet d'interagir avec leur partenaire sur une surface plus large que celle offerte par une protéine structurée de même masse moléculaire (Gunasekaran *et al.*, 2003). Ainsi, avec une taille plus faible, les PID ont des capacités de liaison similaires à celles de protéines structurées beaucoup plus grandes. Ceci permet de réduire la taille du génome, d'améliorer la vitesse de transcription et de traduction, de diminuer l'encombrement cellulaire et de réduire les besoins en acides aminés (Gunasekaran *et al.*, 2003 ; Fink, 2005). Le bénéfice d'une structure étendue est également

bien visible pour la classe des assembleurs, dont le rôle est d'associer plusieurs partenaires. Enfin pour les collecteurs, catégorie dont font partie les PRP salivaires, l'intérêt d'avoir une configuration étendue est de multiplier les sites de liaison. Ceci leur permet d'atteindre les hautes stœchiométries d'interaction essentielles à leur fonction qui nécessite de lier plusieurs ligands de taille réduite en même temps (Tompa, 2003b).

Pour les effecteurs, la flexibilité est un autre atout qui leur permet d'augmenter leur vitesse d'interaction avec un partenaire (Pontius, 1993). Par exemple, les effecteurs se lient à leur ligand grâce à plusieurs points de liaison séparés par des zones flexibles. L'établissement d'un premier point de liaison est assez rapide, les contraintes stériques étant faibles. La liaison établie réduit ensuite l'espace conformationnel à explorer, ce qui permet l'établissement rapide des autres points de contact. Les multiples points de liaison et l'adaptabilité des zones flexibles assurent un haut niveau de spécificité de l'interaction. De plus, grâce à leur malléabilité structurale, les PID peuvent s'adapter à différents partenaires (Kriwacki *et al.*, 1996). Ce phénomène appelé "diversité de liaison" est un comportement plastique permettant de répondre au besoin de la cellule et de l'organisme (Tompa, 2003b).

Un autre avantage fonctionnel à être non structuré pour les PID est lié à leur rôle dans les phénomènes de régulations et de transduction du signal au sein de la cellule. En effet, leur structure dépliée les rend plus sensibles à la protéolyse enzymatique que les protéines repliées. Une durée de vie réduite et un turnover important permettent un meilleur contrôle de leur activité dans le temps (Gsponer *et al.*, 2008).

Le désordre est également exploité par les PID fonctionnant par reconnaissance moléculaire. Ces protéines subissent souvent une transition désordre-ordre (figure 6), accompagnée par une forte diminution de l'entropie lors de la liaison à leur ligand. Ce terme thermodynamique non favorable déconnecte la force de liaison de la spécificité de l'interaction. Ainsi, une interaction hautement spécifique est facilement réversible, ce qui est également très utile dans les processus de régulation (Tompa, 2003b). Ce phénomène de transition désordre-ordre a été démontré pour de nombreuses PID (Daughdrill *et al.*, 1997 ; Hauer *et al.*, 1999 ; Parker *et al.*, 1999 ; Wallon *et al.*, 2000) et peut être appliqué à toutes les PID qui fonctionnent par reconnaissance moléculaire (Tompa, 2003b).

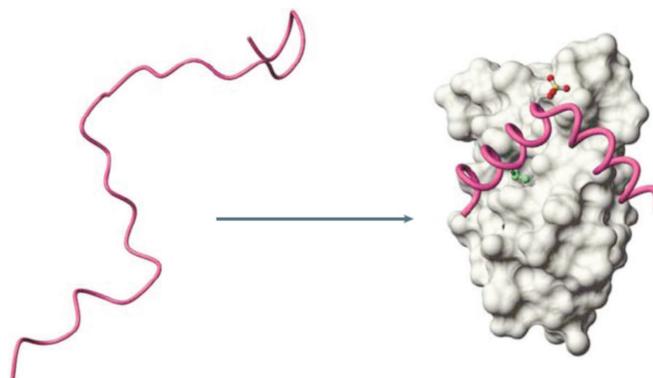


Figure 6. Transition désordre-ordre couplée à une interaction protéine-protéine, tiré de (Dyson et Wright, 2005).

Ainsi, les PID assurent souvent des fonctions nécessitant d'interagir avec un ou plusieurs ligands. Cette interaction peut entraîner une transition désordre-ordre dans laquelle la fonction des PID trouve son origine. Un modèle a été proposé pour décrire ce phénomène.

Fuxreiter *et al.* (Fuxreiter *et al.*, 2004) ont proposé un modèle de repliement pour la structuration des PID lors de leur liaison à un ou des partenaires. Ce modèle repose sur la présence de deux types de régions au sein des PID : des régions avec de faibles préférences de conformations et d'autres avec de fortes préférences. Les PID sont ainsi composées de segments stables liés par des zones flexibles. Au sein des séquences primaires des PRP, les zones stables seraient assurées par les clusters de proline et les zones flexibles par les résidus glycine qui les bordent. Les zones stables sont les régions d'interactions et d'ancrages à leurs partenaires. L'établissement d'un contact initial par l'une de ces régions réduit ensuite l'espace à explorer pour l'établissement de zones d'interactions et d'ancrages supplémentaires. Suite à la reconnaissance des sites de liaison, les régions flexibles sont amenées sur le "moule" du ligand et stabilisées. Plus le nombre de zones flexibles est important, plus rapide est l'interaction. Les faibles variations d'énergie libre entre différentes conformations font qu'une faible différence d'énergie est requise afin de produire la structure optimale pour la liaison. Ainsi la présence de régions avec de fortes préférences structurales réduit la pénalité entropique lors de la formation de complexes. Le fait que peu de points d'ancrage sont nécessaires à la stabilisation de la conformation d'interaction modifie la cinétique de l'interaction. Tous ces éléments participent à la facilité d'interaction nécessaire aux fonctions des PID.

Ce type d'interaction conduit souvent à des états plus ordonnés sans qu'il y ait forcément acquisition de structure secondaire bien définie, processus appelé repliement induit. Ce type de changement de conformation semblerait se produire lors de l'interaction des PRP avec les tanins (Pascal *et al.*, 2007), et pourraient rendre compte de la différence

d'affinité pour les tanins observée entre un peptide de PRP et une PRP entière (Charlton *et al.*, 1996).

6. Les méthodes d'étude utilisées en biologie structurale

Les tanins jouent un rôle important sur les propriétés organoleptiques des produits d'origine végétale. Il est probable que leurs interactions avec les PRP salivaires interviennent dans la sensation d'astringence. Il est donc important de mieux caractériser ces interactions et les édifices supramoléculaires PRP•tanin qui en résultent. Les informations acquises permettront de mieux définir la relation « séquence-structure-fonction » des PRP. Au delà, ces informations pourraient également contribuer à une meilleure compréhension des rôles assurés par les PID au sein des organismes vivants. Ainsi, l'objectif du travail présenté dans ce mémoire est d'apporter des informations sur :

- la conformation des PRP,
- la capacité de ces protéines à interagir avec les tanins au niveau de :
 - leur capacité à établir des complexes avec les tanins et les stœchiométries des complexes formés,
 - leur affinité pour les tanins par rapport à une protéine structurée,
 - la présence de sites d'interaction privilégiés par les tanins au niveau de leur chaîne peptidique,
 - leurs éventuelles modifications de conformation lors de leur interaction avec les tanins,
- l'impact de la structures des tanins sur :
 - l'affinité de l'interaction PRP-tanin,
 - les stœchiométries des complexes,
 - la stabilité des complexes PRP•tanin formés.

Afin d'obtenir ces informations, différentes approches sont généralement utilisées en biologie structurale (tableau 3).

Tableau 3. Principales techniques utilisées en biologie structurale (Receveur-Brechot et al., 2006).

Approche	Principe	Avantages/Inconvénients
Cristallographie	L'interaction entre les rayons X et les électrons des monocristaux de la molécule étudiée permet d'obtenir la carte de densité électronique de la molécule à l'intérieur d'un cristal.	Les cartes de densité électronique permettent de déterminer l'arrangement tridimensionnel des atomes de la structure cristalline. Il est nécessaire d'obtenir des monocristaux de la molécule, ce qui est impossible dans le cas des PID excepté dans certaines conditions d'interaction. La « structure » obtenue n'est pas forcément représentative de la structure en solution. Dans le cas des PID, c'est un inconvénient majeur puisque ce sont des molécules particulièrement dynamiques.
Résonance Magnétique Nucléaire (RMN)	L'application d'un champ magnétique puissant sur un échantillon liquide ou solide va orienter le moment magnétique de spin des noyaux qui se comporte alors comme un aimant. L'utilisation d'impulsions de radio fréquence sur ces « aimants » et l'analyse du retour à l'équilibre du moment magnétique permet de mesurer le déplacement chimique de chaque atome de l'échantillon.	Un ensemble de cartes de corrélation 2D et 3D permet de remonter à la structure primaire directement mais également aux structures secondaire et tertiaire en faisant de la modélisation sous contraintes RMN de la molécule. La restriction de la gamme de déplacement chimique et les nombreux recouvrements de signaux rendent l'analyse des PID complexe.
Diffusion de lumière	L'application d'un rayonnement incident sur la solution contenant les objets à étudier entraîne l'émission d'ondes diffusées par ces derniers. La mesure de ce rayonnement diffusé permet de déterminer différentes caractéristiques des objets suivant que l'on utilise le mode statique ou le mode dynamique.	Cette technique permet de déterminer certaines caractéristiques des objets en solution tels que leur forme ou leur taille en mode dynamique ou leur masse en mode statique. Les données obtenues permettent de modéliser la forme moyenne des objets, ne représentant pas la diversité structurale de ces derniers. Elles ne permettent pas non plus une détermination précise des structures locales d'un objet.
Diffusion des rayons X	Cette technique repose sur le même principe physique que la diffusion de lumière mais les longueurs d'ondes du rayonnement incident sont dans la gamme des rayons X et l'intensité diffusée est fonction de la densité électronique des particules.	Cette technique permet d'étudier des objets plus petits qu'en diffusion de la lumière.

Dichroïsme Circulaire (CD)	Mesure la différence d'absorption de la lumière par une molécule optiquement active (chirale) suivant que la polarisation de la lumière soit circulaire droite ou circulaire gauche.	Les structures secondaires des protéines présentent des signatures dichroïques caractéristiques : détermination des éléments de structure secondaire au sein d'une protéine. Cette technique permet seulement une mesure des proportions des éléments de structure secondaire de la molécule, sans identification des régions concernées.
Microscopie électronique (EM)	Les effets des interactions entre un faisceau d'électrons et l'échantillon donnent naissance à une image de celui-ci. Coloration négative : enrobage des protéines avec un sel d'atomes lourds. Cryo-ME : congélation rapide de l'échantillon.	Visualisation directe des protéines à partir d'une taille de 50-100 kDa. Il est nécessaire de passer par une étape de fixation des molécules. Coloration négative : simple et rapide, image contrastée de la surface de l'objet. Peut abîmer la structure native de l'objet. Résolution de l'ordre de 10Å. Cryo-ME : respecte mieux la structure native. La modélisation réalisée à partir des données permet d'accéder à la structure interne des molécules.
Spectrométrie de Masse (MS)	Après ionisation des molécules de l'échantillon, la MS mesure leur rapport m/z généralement couplée à différentes techniques de fragmentation ou/et à la mobilité ionique.	Mesure du rapport m/z de chacune des molécules et de chacun des édifices supramoléculaires ionisables présents dans l'échantillon, Détermination des stœchiométries des complexes. Mesure de la stabilité des complexes par MS/MS CID. Détermination des sites de liaison par MS/MS ECD. Mesure des changements de conformation par échanges H/D. Ne permet pas d'établir la structure précise des molécules analysées.
Spectrométrie de Mobilité Ionique (IMS)	Cette technique donne accès à la section efficace de collision d'une molécule sur la base de son temps de migration à travers un tube de mobilité.	Mesure de la section efficace de collision des molécules. Les mesures sont effectuées en phase gazeuse. Peut être couplée à la MS.

Le travail que nous avons entrepris s'inscrit dans une démarche pluridisciplinaire, mettant en jeu une partie de ces techniques (RMN, CD, diffusion de lumière et de rayons X, MS) qui apportent des informations complémentaires, pour l'étude des interactions PRP-

tanin. Ce projet, appelé PROTANIN, est supporté financièrement par l'Agence Nationale de la Recherche (ANR). Les travaux présentés dans ce mémoire portent sur l'approche par MS.

L'apparition de sources d'ionisation douces comme l'ionisation électrospray (ESI) qui conservent les édifices non-covalents lors de l'étape d'ionisation a permis l'essor d'un nouveau domaine d'application au sein la spectrométrie de masse : l'étude des complexes non-covalents. Ce nouveau domaine porte le nom de « MS supramoléculaire » (Sanglier *et al.*, 2005).

7. La spectrométrie de masse ou MS

La spectrométrie de masse (MS) est une méthode destructive donnant accès à des informations importantes sur les molécules d'un échantillon. Elle permet la mesure de la masse moléculaire d'une substance et d'obtenir des données structurales. Cette méthode mesure les rapports masse-sur-charge (m/z) de molécules ionisées ainsi que de leurs produits de fragmentation. La MS fournit des informations qualitatives et quantitatives sur la composition atomique et moléculaire. Elle présente de nombreux atouts qui ont été résumés par McLafferty par la qualification de « Méthode aux 3 S » : « Sensitivity, Specificity, Speed » (McLafferty, 1981).

Les récents développements en MS, en particulier ceux concernant les sources d'ionisation, ont permis un élargissement de son champ d'application: biopolymères naturels, protéines, ARN, ADN, polysaccharides et même complexes non-covalents (McLafferty, 2008). L'émergence de la MS comme méthode d'étude des édifices supramoléculaires non-covalents a conduit Daniel à rebaptiser la MS en « Méthode aux 4 S » avec l'ajout du terme « Stoichiometry » (Daniel *et al.*, 2002)

Pour mesurer la masse de molécules isolées, le spectromètre de masse doit assurer les opérations suivantes : 1- Volatiliser ; 2- Ioniser ; 3- Mesurer les rapports m/z

Classiquement, il est constitué de 3 parties principales (figure 7) :

- la source d'ionisation, dont le rôle est de faire passer les molécules en phase gazeuse et de les ioniser ;
- l'analyseur, qui sépare les ions en fonction de leur rapport masse-sur-charge (m/z) ;
- le détecteur.

Le détecteur est couplé à un système informatique, qui assure le traitement des données et l'élaboration des spectres de masse.

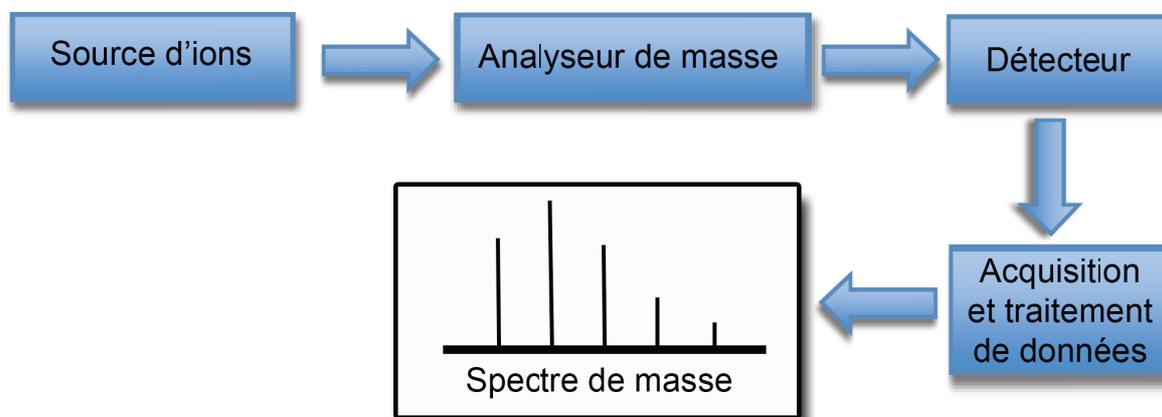


Figure 7. Représentation schématique d'un spectromètre de masse.

Le spectromètre fonctionne dans des conditions de pression allant de la pression atmosphérique au niveau de la source d'ionisation à un vide poussé au niveau de l'analyseur. Le vide est nécessaire pour optimiser la transmission des ions jusqu'au détecteur sans qu'ils subissent de collisions avec les molécules gazeuses. En effet, les collisions peuvent entraîner une déviation de la trajectoire des ions, qui iraient se charger sur la paroi de l'appareil. Les collisions peuvent également être à l'origine de réactions de fragmentation qui, si elles ne sont pas désirées, compliquent le spectre (de Hoffmann et Stroobant, 2005). Le système d'introduction d'échantillon est conçu pour une perte minimale de vide. La source d'ions crée des ions gazeux à partir de l'échantillon.

La spectrométrie de masse se caractérise par la grande diversité des appareils disponibles sur le marché ou développés au sein des laboratoires de recherche. Cette diversité se retrouve au niveau :

- des sources d'ionisation : électrobulisisation (ESI, *ElectroSpray Ionisation*), ionisation par impact électronique (EI, *Electron Ionization*), ionisation chimique (CI, *Chemical Ionization*), photoionisation à pression atmosphérique (APPI, *Atmospheric-Pressure Photolonization*) et désorption et ionisation par laser assistées par matrice (MALDI, *Matrix-Assisted Laser Desorption Ionisation*)...
- des analyseurs : quadripôles (Q), trappes ioniques (IT), temps de vols (TOF), analyseurs à résonance d'ions en cyclotron à transformée de Fourier (FT-ICR), analyseurs magnétiques, orbitrap ...

- des configurations hybrides d'analyseurs : Q-TOF, IT-TOF, IT-orbitrap, triple quadripôle ...
- du couplage avec la spectrométrie de mobilité ionique (IMS).

Ces configurations variables d'appareil permettent l'analyse d'échantillons de nature différentes et d'apporter un large panel d'informations. Elles sont également adaptées ou non à la réalisation d'expériences dites de MS/MS qui servent généralement à apporter des informations sur les objets étudiés en les fragmentant. Là encore, les techniques de fragmentations sont variées (Dissociation Induite par collision (CID, Collision Induced Dissociation), dissociation par capture d'électrons (ECD, *Electron Capture Dissociation*), dissociation par transfert d'électron (ETD, *Electron Transfert Dissociation*)...) et permettent d'obtenir différentes informations.

Pour ce travail différentes configurations d'appareils ont été envisagées : ESI-IT, ESI-TOF, ESI-Q-TOF, ESI-FT-ICR, ESI-IM-Q-TOF, MALDI-TOF et MALDI-TOF/TOF. En fonction de la configuration d'appareil utilisée, plusieurs expériences ont été réalisées dont certaines ont fait intervenir les techniques de fragmentation CID et ECD. La section bibliographique qui suit présente le fonctionnement, les avantages et inconvénients et les informations apportées par chacune des configurations d'appareils utilisées.

7.1. Les sources d'ionisation

Il existe de nombreux types de sources d'ions. Chacun repose sur un principe physique différent. L'opérateur choisit le principe physique capable de volatiliser et d'ioniser le composé étudié en fonction de ses caractéristiques.

Les étapes de volatilisation et d'ionisation se font successivement ou simultanément selon le type de source. Les critères de choix principaux sont :

- la volatilité et la stabilité thermique du composé à analyser,
- sa labilité chimique,
- les fonctions chimiques présentes et leur aptitude à induire une ionisation,
- la taille des molécules,
- les quantités de produit disponibles,
- le solvant de solubilisation du composé,
- le type d'introduction souhaitée (directe ou en couplage chromatographique).

Les types de sources les plus classiques sont : l'ionisation par impact électronique (EI) et l'ionisation chimique (IC) pour l'ionisation en phase gazeuse, l'électronébulisation (ESI), la photoionisation à pression atmosphérique (APPI) et la désorption et ionisation par laser assistées par matrice (MALDI) pour l'analyse en phase liquide, même si pour la MALDI une étape de cristallisation à l'aide d'une matrice est nécessaire (Karas et Hillenkamp, 1988).

L'étape d'ionisation est cruciale : elle assure non seulement la volatilisation, l'ionisation proprement dite et le guidage vers l'analyseur, mais elle détermine aussi l'énergie apportée aux ions parents. Un apport trop important conduit à la fragmentation de ces derniers. Longtemps, les techniques d'ionisation ont été trop énergétiques pour permettre l'étude des objets biologiques (protéines intactes, complexes ligand-protéine, ARN, etc). Au cours des dernières années, le développement de méthodes d'ionisation douce telles que l'ESI et la MALDI a permis l'essor de la spectrométrie de masse dans le monde de la biologie. La MS est devenue une technique de référence dans l'identification des protéines, dans l'analyse fine de leurs séquences et de leurs éventuelles modifications post-traductionnelles. Ces méthodes d'ionisation douce permettent aussi la caractérisation des édifices supramoléculaires et les études des changements de conformation et des repliements induits (Sanglier *et al.*, 2005).

Pour l'étude des édifices supramoléculaires tels que les complexes PRP-tanin, l'ESI constitue la technique d'ionisation la plus adaptée. En effet, contrairement aux autres sources mentionnées précédemment, l'ionisation des molécules par ESI se déroule à pression atmosphérique et température ambiante. Rapportée comme la plus douce de toutes les méthodes d'ionisation, ce type de source est capable de garder intacts les complexes non-covalents dans les conditions d'ionisation et de transfert des ions appropriées (Robinson *et al.*, 1996 ; Pramanik *et al.*, 1998 ; Loo, 2000 ; Griffiths *et al.*, 2001 ; Daniel *et al.*, 2002 ; Gabelica *et al.*, 2009). Les liaisons faibles établies en solution au niveau de ces objets supramoléculaires peuvent ainsi être conservées (Loo et Ogorzalek-Loo, 1997 ; Gabelica et De Pauw, 2001 ; Robinson, 2002).

7.1.1. L'électrospray (ESI ou ES)

Développée par le groupe de Fenn (prix Nobel de Chimie en 2002), afin de permettre l'ionisation des macromolécules biologiques, l'électrospray ou électronébulisation est la technique d'ionisation la plus utilisée. Elle autorise l'étude des échantillons liquides avec une grande sensibilité. Elle permet ainsi de « suivre » la majeure partie des réactions chimiques et biologiques qui se déroulent en phase liquide. De plus elle peut être facilement couplée à des techniques séparatives (chromatographie liquide, électrophorèse capillaire). Mais la

raison principale de son importante utilisation réside dans le fait que lors du processus d'électronébulisation, l'énergie interne ajoutée aux ions peut être contrôlée afin de ne pas entraîner de fragmentation moléculaire.

Typiquement, l'ionisation de type ESI peut être divisée en trois étapes (figure 8) (Bruins, 1998):

1°) La nébulisation : la solution à analyser est dispersée en des microgouttelettes chargées, sous l'action conjointe d'un différentiel de potentiel entre deux électrodes et de l'application d'un gaz de nébulisation.

2°) La désolvation : durant le trajet de ces microgouttelettes, le solvant qui les compose s'évapore et leur diamètre devient donc de plus en plus petit. La densité de charge au sein de chaque gouttelette augmente jusqu'à ce qu'elle devienne trop importante et qu'il se produise une explosion coulombienne, provoquant une libération des ions en phase gazeuse.

3°) Le transport des ions depuis la source, qui est à pression atmosphérique, jusqu'au vide poussé de l'analyseur.

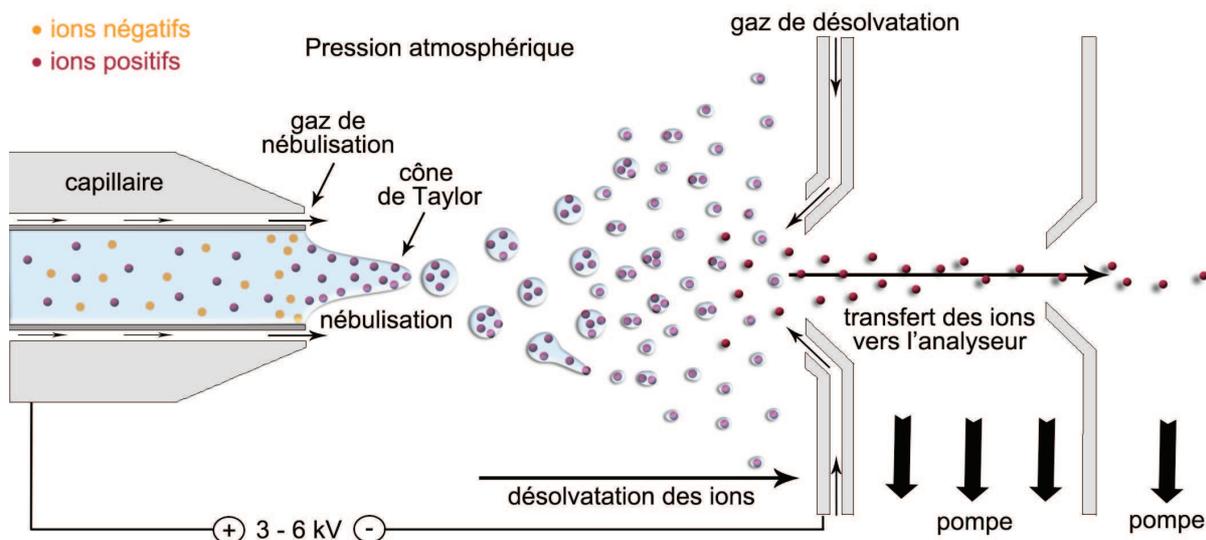


Figure 8. Principe de l'électrospray

La nébulisation :

La nébulisation correspond à la dispersion d'un liquide en gouttelettes électriquement chargées. Elle combine deux processus, la formation des gouttelettes et leur acquisition de charge.

Ces processus qui s'effectuent à pression atmosphérique résultent de l'application d'un potentiel élevé sur un capillaire au travers duquel passe la solution de l'échantillon à

analyser. Ce capillaire est enchâssé dans un tube au travers duquel circule un débit constant de gaz inerte, appelé gaz de nébulisation (N_2 la plupart du temps) qui assure la nébulisation de la solution. Le champ électrique est obtenu par application d'une différence de potentiel de 3 à 6 kV entre ce capillaire et la contre électrode, séparés de 0,3 à 2 cm. Ce champ provoque une accumulation de charges à la surface du solvant, au niveau de l'extrémité du capillaire. Le solvant se déforme alors en prenant l'aspect d'un cône, appelé « cône de Taylor » (Taylor, 1964) qui se rompt et forme des gouttelettes ayant une densité de charge importante (Fenn *et al.*, 1989 ; Bruins, 1998 ; de Hoffmann et Stroobant, 2005). Lorsque les forces de répulsions coulombiennes deviennent plus importantes que la tension de surface du solvant, de petites gouttelettes électriquement chargées quittent l'extrémité du capillaire (Bruins, 1998). A partir du cône de Taylor, ces gouttelettes primaires, avec un rayon de 0,5 à 1,5 μm , sont émises.

Ce processus peut être effectué en mode positif comme en mode négatif et permet, dans un cas, la génération d'ions positifs et, dans l'autre, celle d'ions négatifs. Le passage du mode positif au mode négatif consiste à changer la polarité des électrodes.

La désolvatation et l'ionisation : libération des ions des gouttelettes

Au cours de leur vol à travers la chambre d'ionisation, les gouttelettes sont déformées par des forces de friction et par le champ électrique intense. La combinaison de ces phénomènes a pour effet d'augmenter localement les charges de surface. Les forces de répulsion coulombiennes devenant supérieures à la tension de surface du solvant, chaque gouttelette se déforme localement en un cône de Taylor. A partir de celui-ci, un jet d'une vingtaine de nouvelles gouttelettes est émis de chacune des gouttelettes mères. Ces gouttelettes filles de première génération ont un rayon d'environ 100 nm (figure 9). Ce phénomène, appelé fission coulombienne, semble avoir lieu avant que la limite d'instabilité appelée limite de Rayleigh ait été atteinte (Gomez et Tang, 1994 ; de Hoffmann et Stroobant, 2005). Ces processus se répètent et s'accélèrent avec l'augmentation du rapport surface/volume des gouttelettes due à la diminution de leur taille. De nouvelles générations de gouttelettes filles sont ainsi formées (Kearle et Peschke, 2000). Le phénomène se poursuit jusqu'à ce que les gouttelettes aient un rayon de l'ordre de 10 nm. Les répulsions de charges au niveau de ces dernières deviennent alors suffisantes pour provoquer la désorption d'ions (figure 9 a) (Iribarne et Thomson, 1976 ; Thomson et Iribarne, 1979). Ce processus est appelé évaporation des ions ou « Ion Evaporation Model » (IEM) (Iribarne *et al.*, 1983). Un ion encore solvaté peut aussi se séparer de la gouttelette (nanogouttelette). Le solvant restant s'évapore durant la fin du vol (figure 9 b), ce modèle est appelé « Charge Residue Model » (CRM) (Dole *et al.*, 1968). Il semblerait que la majorité des ions soit formée

selon le modèle IEM. Les ions les plus gros peuvent être formés selon le modèle CRM (Clusters et protéines). Les deux phénomènes conduisent à la formation d'ions non solvatés pouvant pénétrer dans le spectromètre de masse.

Différentes adaptations techniques ont été mises au point afin d'améliorer la désolvatation des ions. Ces derniers pourront finir d'être désolvatés :

- par leur passage à travers un « dry curtain gaz »,
- par leur passage à travers un guide d'ions chauffé,
- ou par des collisions en phase gazeuse faiblement énergétiques.

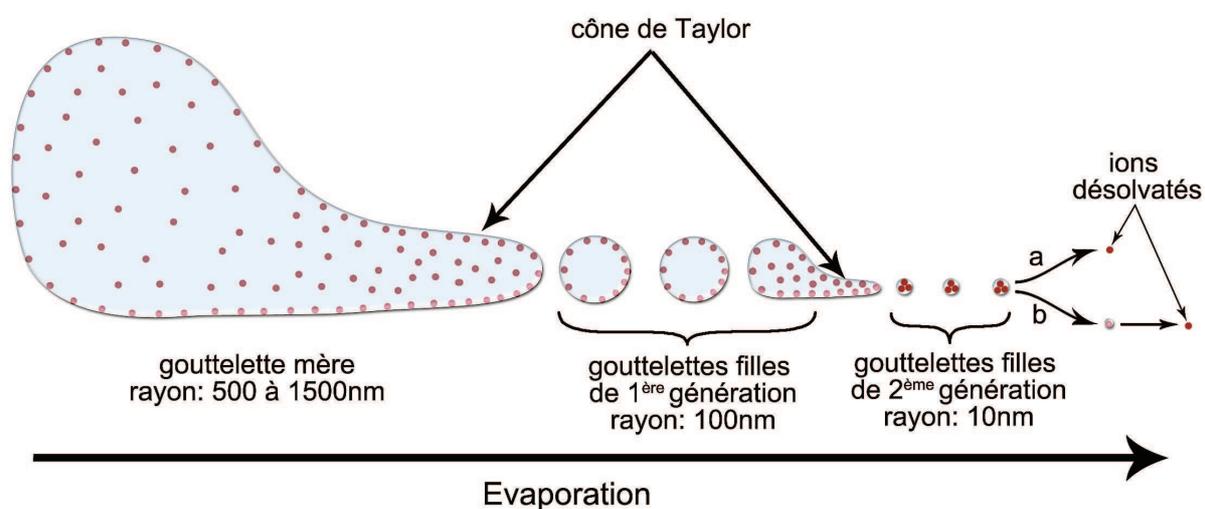


Figure 9. Libération des ions à partir des gouttelettes émises lors de la nébulisation ; a. libération directe des ions à partir de la surface des gouttelettes ; b. éjection d'un ion encore solvaté (nanogouttelette) et dont le solvant résiduel sera évaporé.

Nature des ions : Ions multichargés

Les ions obtenus via une source de type ESI sont multichargés si la molécule mère possède plusieurs sites ionisables. L'obtention d'ions multichargés améliore la sensibilité et permet l'analyse de molécules de poids moléculaire élevé. Ainsi, des analyseurs ayant une faible valeur limite maximale au niveau de la gamme de masse à laquelle ils opèrent peuvent être utilisés pour l'étude de molécules dont la masse est supérieure à cette limite. En effet, les caractéristiques techniques des spectromètres de masse font que la grandeur mesurée ne correspond pas à la masse d'un ion mais à son rapport masse-sur-charge (m/z). Il est ainsi possible d'analyser des protéines de plus de 130 kDa avec une limite de détection de l'ordre de la picomole à l'aide d'un analyseur quadripolaire, dont la limite maximale de la gamme de masse analysable est de 4000 u (de Hoffmann et Stroobant, 2005).

7.1.2. La Désorption et Ionisation par Laser Assistées par Matrice (MALDI)

La MALDI (Matrix-Assisted Laser Desorption Ionisation) est une autre technique d'ionisation douce. Son introduction est due au travail de 2 groupes, celui de Tanaka (prix Nobel de Chimie en 2002) et celui de Karas qui publient séparément en 1988 des travaux sur l'ionisation de molécules ayant des masses moléculaires supérieures à 100 000 Da (Karas et Hillenkamp, 1988 ; Tanaka *et al.*, 1988). Aujourd'hui, la MALDI est devenue l'approche de choix pour l'étude de molécules de hauts poids moléculaire comme les biopolymères (peptides, protéines, oligonucléotides...) et les polymères synthétiques. Elle est ainsi très employée pour l'étude des échantillons biologiques. Les spectres obtenus sont principalement composés d'ions monochargés qui simplifient l'analyse. Quelques espèces di- ou tri-chargées peuvent également être observées. La relative simplicité des spectres MALDI permet l'analyse de mélanges complexes. Cette technique a également l'avantage de tolérer certains sels, tampons ou détergents.

L'ionisation de type MALDI se déroule en deux étapes :

- 1°) Formation d'une matrice cristallisée contenant l'échantillon à analyser.
- 2°) Désorption et ionisation des molécules de l'échantillon sous vide à l'aide d'un laser.

Formation d'une matrice cristallisée contenant l'échantillon

Au cours de cette étape, l'échantillon est mélangé à une matrice composée de petites molécules organiques. Ces dernières sont généralement dissoutes dans une solution acidifiée composée d'eau et d'un solvant organique (en général l'acétonitrile (ACN) ou l'éthanol (EtOH)). Les composés généralement utilisés comme matrice ont les caractéristiques suivantes :

- masses moléculaires relativement faibles, mais suffisamment importantes pour ne pas se volatiliser durant la préparation de l'échantillon ou pendant que la cible est sous vide dans la source du spectromètre,
- faible pression de vapeur saturante (Gross, 2004),
- acides, constituant une source de protons facilitant l'ionisation des analytes, même s'il existe des matrices basiques,
- capacité d'absorption à la longueur d'onde du laser utilisé. Cette caractéristique est généralement associée à la présence au niveau de leur structure chimique de noyaux

aromatiques, qui constituent des chromophores adaptés aux longueurs d'onde des lasers (Ultraviolet (UV) ou Infrarouge (IR)) (Gross, 2004).

- fonctionnalisés avec des groupements chimiques polaires permettant leur utilisation en solution aqueuse.

De nombreuses matrices sont disponibles (Dass, 2007). Le choix final de la matrice dépend de l'échantillon et se fait généralement après la réalisation d'essais préliminaires.

Les données bibliographiques rapportent différentes méthodes de préparation du mélange matrice-échantillon (de Hoffmann et Stroobant, 2005 ; Dass, 2007). La méthode de préparation la plus répandue (utilisée au cours de cette thèse) est appelée "*Dried-droplet method*". Cette technique consiste à mélanger une solution saturée de matrice avec l'échantillon (figure 10). Une gouttelette du mélange obtenu est déposée sur une plaque métallique appelée la cible MALDI. La gouttelette est ensuite mise à sécher à température ambiante.

Après le dépôt sur la cible du mélange matrice-échantillon (figure10), l'évaporation du solvant entraîne la co-cristallisation des molécules de matrice avec celles de l'échantillon. Les molécules de matrice sont en très large excès par rapport aux molécules d'analytes présentes. L'ensemble formé est appelé phase condensée. Lorsque le solvant s'est complètement évaporé, la cible est placée dans la source du spectromètre de masse.

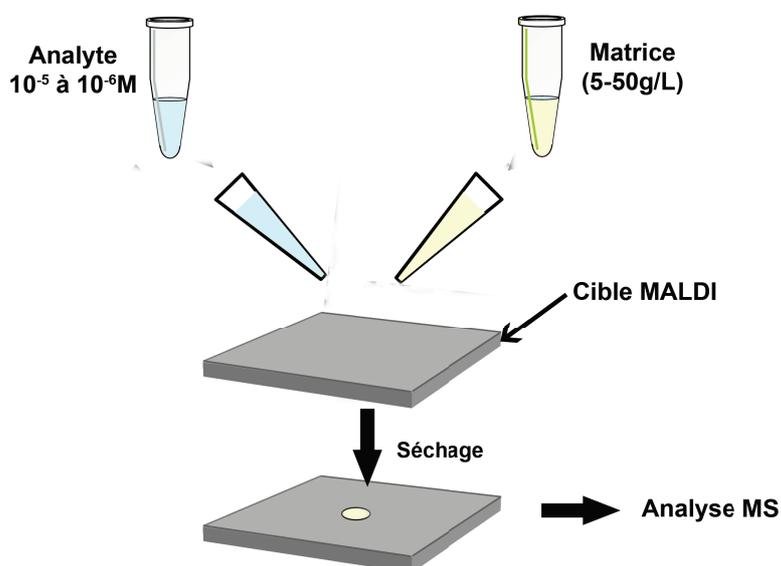


Figure 10. Exemple de dépôt : « *Dried-droplet method* ».

Désorption et ionisation des molécules de l'échantillon à l'aide d'un laser

La deuxième étape s'effectue sous vide à l'intérieur de la source du spectromètre de masse. Au cours de cette étape, l'énergie apportée par les impulsions laser entraîne l'évaporation et l'ionisation des molécules de l'échantillon (figure 11). L'irradiation par le laser conduit à l'accumulation d'une grande quantité d'énergie au sein de la phase condensée par l'excitation des molécules de la matrice. Les « tirs » laser sont focalisés sur une faible surface de 10 à 200 μm de diamètre (Dreisewerd *et al.*, 1995). Cet apport d'énergie très localisé provoque la sublimation de la matrice et l'ablation d'une partie du dépôt (de Hoffmann et Stroobant, 2005). La variation de l'intensité des impulsions laser met en évidence un seuil d'énergie à partir duquel des ions sont détectés. Ce seuil dépend à la fois de la nature de la matrice et du rapport molaire matrice/analyte (Medina *et al.*, 1994). Lorsque le seuil d'énergie est atteint, il se produit une désorption des molécules ionisées et des molécules neutres dans le vide sous la forme d'un panache moléculaire, qui prend la forme d'une petite plume chaude (figure 12).

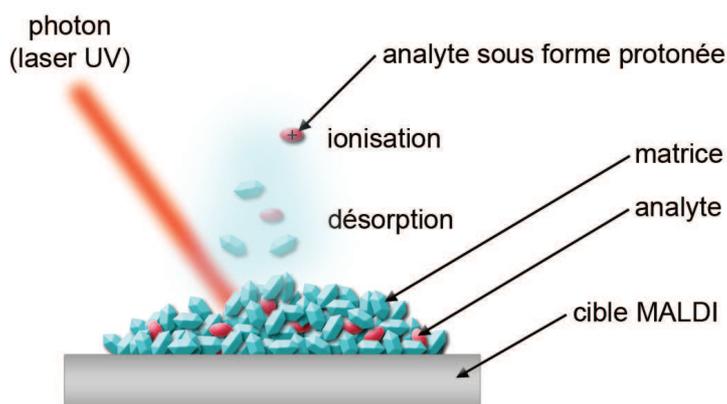


Figure 11. Principe de l'ionisation de type MALDI.

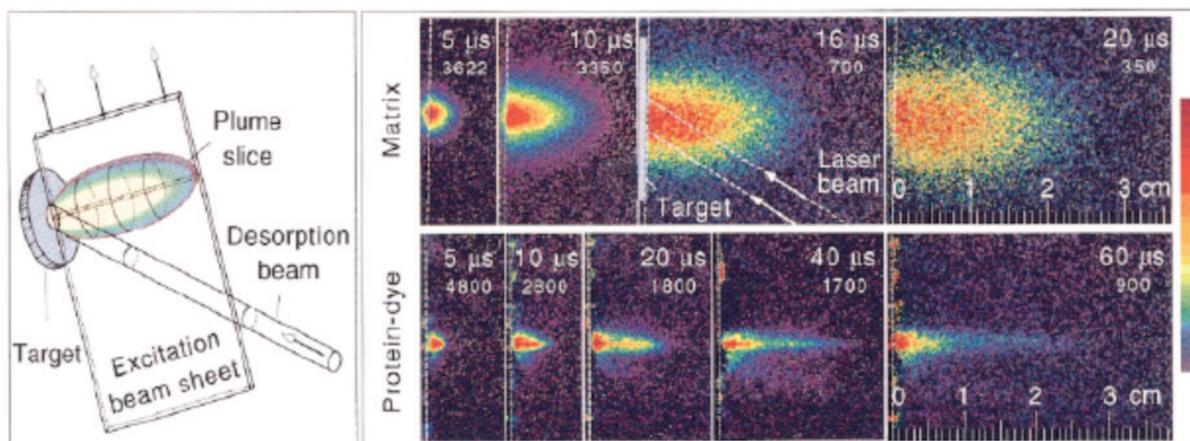


Figure 12. Formation de la plume MALDI, tiré de (Poretzky *et al.*, 1999).

L'utilisation d'une matrice fait qu'il n'est pas nécessaire d'ajuster la longueur d'onde du laser à l'absorption de chaque analyte. En outre, le processus étant indépendant des propriétés d'absorption et de la taille de l'analyte, l'ionisation de type MALDI permet l'analyse de composés de très grande masse moléculaire pouvant dépasser les 100 000 Da. Ainsi, des protéines de masse moléculaire supérieure à 300 000 Da avec des sensibilités de l'ordre de la femtomole ont été étudiés par MALDI-MS (Spengler et Cotter, 1990).

Il existe deux catégories de lasers définies par rapport à leur longueur d'onde d'émission :

- les lasers UV qui émettent des photons dans la gamme de longueur d'onde de l'ultraviolet (UV),
- les lasers IR qui émettent dans la gamme de longueur d'onde de l'infrarouge (IR).

Le mécanisme exact de formation des ions en MALDI n'est pas complètement connu. Il semble que la localisation et les processus de formation des ions en MALDI soient multiples (Glückmann *et al.*, 2001). Le mécanisme le plus largement admis pour la formation des ions implique le transfert de protons, soit avant désorption en phase solide, soit dans le panache d'expansion en phase gazeuse entre la matrice photoexcitée et l'analyte (de Hoffmann et Stroobant, 2005). L'effet favorable de l'acidification de la solution analyte-matrice indique que la désorption d'ions $[M+H]^+$ préformés joue également un rôle. D'autres mécanismes de formation des ions en phase gazeuse tels que la photoionisation (en UV), la capture d'électron et l'échange de charge semblent également impliqués (Liao et Allison, 1995 ; Dass, 2007).

Les ions formés sont ensuite extraits de la source par l'application d'un champ électrostatique et envoyés vers l'analyseur.

7.2. Les analyseurs

Les ions produits sont séparés d'après leur masse et celle-ci est simultanément déterminée au moyen de l'analyseur. De nombreux analyseurs sont disponibles. Toutefois comme pour les sources d'ionisation, l'analyseur parfait n'existe pas. Chaque analyseur a ses avantages et ses inconvénients. En fonction du type d'analyse à réaliser et du type de source utilisée, le choix de l'analyseur va dépendre principalement :

- de sa gamme de masse (valeur limite du rapport m/z mesurable et gamme dynamique),

- de sa précision dans la mesure de la masse,
- de sa résolution, qui correspond à sa capacité à différencier des ions de rapport m/z voisins,
- de sa capacité à focaliser les ions, même après collision.

Les analyseurs hybrides qui combinent les spécificités intéressantes de différents types d'analyseurs équipent les appareils les plus performants. Les analyseurs que nous avons utilisés sont de type trappe ionique, temps de vol, FT-ICR, couplage quadripôle/temps de vol et couplage temps de vol/temps de vol.

A ce niveau, il est important de définir les termes de résolution en masse et de précision de masse.

Résolution en masse : On admet que deux pics sont résolus quand l'intensité de la vallée entre les deux pics est égale à 10% de l'intensité du pic le plus faible (on parle de résolution à 10%) quand on parle des appareils magnétiques ou FT-ICR, et à 50% pour les appareils quadripôle et IT (figure 13).

La résolution est donnée par l'équation $R=m/\delta m$ où δm est la plus petite différence de masse pour deux pic résolus de masse m et $m+\delta m$.

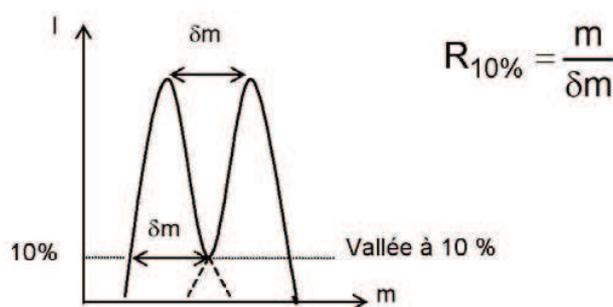


Figure 13. La résolution en masse.

Pour le calcul, on détermine δm comme la largeur du pic à $x\%$ de son maximum. Le plus souvent $x=50$ et on parle alors de largeur à mi-hauteur (FWHM, *Full Width at Half Maximum*).

L'exactitude en masse : elle indique la précision, ou plus correctement la justesse, des rapports m/z mesurés. Elle dépend fortement de la stabilité et du pouvoir de résolution de l'analyseur. Elle est exprimée en pourcent (%) ou en partie par million (ppm) du rapport m/z théorique (de Hoffmann et Stroobant, 2005).

7.2.1. Les analyseurs de type temps de vol « TOF »

Les analyseurs à temps de vol (Time of Flight : TOF) séparent les ions en fonction de leur vitesse de déplacement dans une région libre de champ appelée le tube de vol. Ils mesurent le temps que mettent les ions à franchir une distance fixe après avoir été accélérés par un champ électrique. Les analyseurs de type temps de vol peuvent être séparés en deux catégories, ceux avec réflectron, TOF à réflectron, et ceux sans, appelés TOF linéaire.

Les temps de vol de type linéaire :

Ils sont constitués de deux parties principales situées en aval de la source d'ionisation (figure 14):

- 1- A la sortie de la source d'ions (1) se trouve une **région d'accélération** (2), où les ions sont soumis à une différence de potentiel permettant aux ions ayant un m/z identique d'acquérir la même énergie cinétique.
- 2- **une zone libre de champ ou "Drift region"** (3) où, durant leur vol, les ions se séparent en fonction de leur vitesse qui dépend de leur rapport m/z .

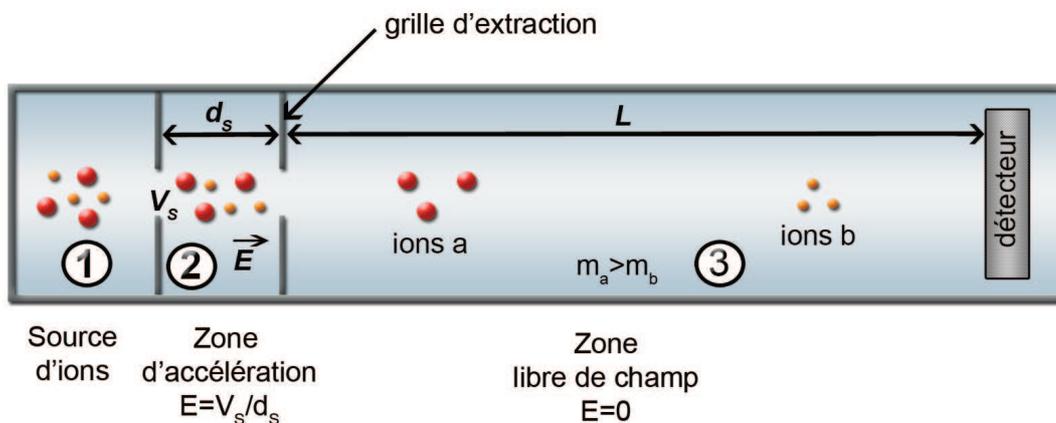


Figure 14. Principe d'un analyseur de type TOF linéaire.

Les ions sont accélérés grâce à la différence de potentiel appliquée entre une électrode et la grille d'extraction (2). Le potentiel électrique (V_s) en un point S du champ électrique (\vec{E}) est égal à :

$$V_s = \frac{E_{pelec}}{q}$$

E_{pelec} étant l'énergie potentielle électrique de l'ion et q sa charge telle que $q=ze$ avec z le nombre de charges de l'ion considéré et e la charge de l'électron.

La variation de l'énergie potentielle électrique d'un ion de charge q dans un champ électrique uniforme \vec{E} ($\Delta E_{p\acute{e}lec}$) étant égale à :

$$\Delta E_{p\acute{e}lec} = q E d_s$$

où d_s est la distance à parcourir par l'ion dans ce champ électrique,

l'équation du potentiel électrique (V_s) en un point S du champ électrique peut s'écrire :

$$V_s = E d_s \text{ ou } E = \frac{V_s}{d_s}$$

et la variation d'énergie potentielle de l'ion de charge q correspond à :

$$\Delta E_{p\acute{e}lec} = q V_s$$

Supposons qu'un ion de masse m et de charge totale q soit accéléré par un potentiel V_s , à la sortie de la zone d'accélération, il aura une énergie cinétique E_c :

$$E_c = \frac{m v^2}{2} = q V_s = e z V_s$$

Il parcourra la distance de vol L en un temps t tel que:

$$t = \frac{L}{v} \quad \text{où } v \text{ est sa vitesse:} \quad v = \left(\frac{2 e V_s}{m} \right)^{1/2}$$

Cette équation équivalente à

$$t^2 = \frac{m}{z} \left(\frac{L^2}{2 e V_s} \right)$$

permet, à partir de la mesure de t , de déterminer le rapport m/z , le terme entre parenthèses étant constant. Cette équation montre qu'un ion ayant un rapport m/z élevé mettra plus de temps à traverser le tube de vol. Les ions volent d'autant plus vite qu'ils sont « légers ».

Les avantages des analyseurs TOF de type linéaire sont que :

- leur domaine de masses n'a pas de limite supérieure. Ils sont donc bien adaptés à l'ionisation de type MALDI capable de générer des ions de rapport m/z élevé.

- ils assurent une transmission des ions pouvant avoisiner les 100%. En effet, la séparation temporelle des ions permet de tous les envoyer vers le détecteur. Par conséquent, l'ensemble des ions produits est en principe analysé et détecté contrairement aux analyseurs à balayage qui analysent les ions successivement au cours du temps.

- leur vitesse d'analyse est très rapide et peut atteindre 1000 unités de masse par μs . Il est ainsi possible de produire théoriquement plusieurs milliers de spectres par seconde sur une très large gamme de masses.

Les limites de ce type d'appareil sont liées au fait que la détermination du rapport m/z des ions est fonction de leur énergie cinétique. Ainsi, les variations d'énergie cinétique initiales entre les ions de même rapport m/z au sortir de la source modifient leur temps de vol et donc le calcul de leur rapport m/z . Cela se traduit par une diminution de la résolution (distribution du temps de vol d'une population d'ions de même masse). Afin, de palier ce problème, les appareils de type TOF les plus récents peuvent être équipés de deux technologies : un extracteur retardé et/ou un réflectron.

Les temps de vols avec système réflectron

Le système réflectron, également appelé miroir électrostatique, permet d'améliorer la résolution des analyseurs de type TOF. Des ions de même m/z , mais ayant des énergies cinétiques initiales différentes, arriveront légèrement décalés au niveau du détecteur. Le réflectron permet de focaliser le temps d'arrivée sur le détecteur des ions d'un même rapport m/z .

Dans ce type d'analyseur, les ions sont émis en direction du réflectron. Ce dernier est composé d'une série d'électrodes portées à des potentiels croissants qui définissent un champ électrique repousseur homogène. Le champ électrique freine les ions jusqu'à ce qu'ils atteignent une vitesse nulle. Il les réaccélère ensuite dans la direction opposée. La trajectoire des ions est ainsi inversée (figure 15). Finalement, les ions ressortent du réflectron à la même vitesse mais avec une direction opposée. Les ions les plus énergétiques arrivent les premiers au niveau du réflectron, mais y pénètrent plus profondément. Ils passent ainsi un temps plus important dans le réflectron.

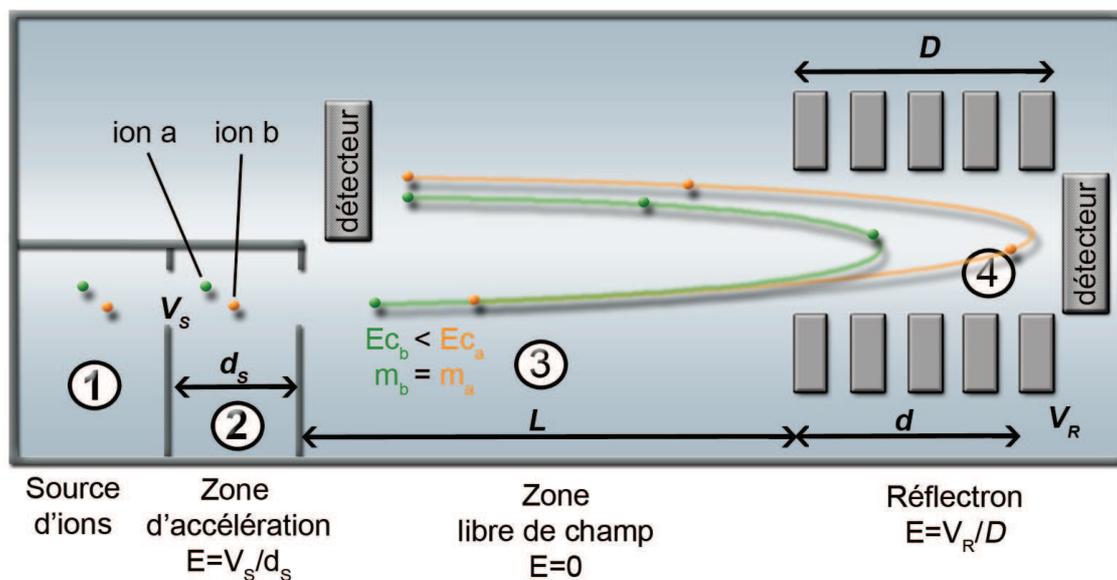


Figure 15. Principe d'un analyseur de type TOF à réflectron.

Si une tension fixe V_R est appliquée au niveau d'un réflectron de longueur D , un champ électrique $E = \frac{V_R}{D}$ est induit. Un ion de charge q possédant une énergie cinétique $E_c = qV_s$ pénètre à une distance d au sein du réflectron telle que :

$$d = \frac{E_c}{qE}$$

Ainsi, plus les ions ont une énergie cinétique importante, plus ils pénètrent profondément dans le réflectron et plus leur temps de parcours à l'intérieur de ce dernier est important. Au sein d'une population d'ions de même rapport m/z , ceux possédant l'énergie cinétique la plus importante passent un temps plus court dans la zone libre de champs (3) mais un temps plus long dans le réflectron (4). Au final, des ions de même rapport m/z mais de différentes énergies cinétiques arrivent en même temps au niveau du détecteur. Ceci a pour conséquence d'augmenter la résolution.

En revanche, l'inconvénient de cette technique est la plus faible transmission des ions, en particulier de ceux de haut rapport m/z . C'est pourquoi, les appareils équipés d'un réflectron possèdent généralement deux détecteurs. Un détecteur est adapté à une utilisation avec le mode réflectron. Le deuxième détecteur est réservé à une utilisation en mode linéaire, lorsque le champ électrique du réflectron est désactivé. L'analyse des molécules de faible masse sera plutôt réalisée en mode réflectron, tandis que celle des grosses molécules se fera préférentiellement en mode linéaire.

L'extraction retardée

L'extraction retardée est un autre système permettant de focaliser les ions. Ce système prend place entre la source et la zone d'accélération des ions. Il peut donc être installé sur les appareils avec et sans réflectron. L'extraction retardée a pour but de mieux focaliser dans le temps les ions de même rapport m/z au niveau du détecteur. Elle consiste à homogénéiser les énergies cinétiques des ions dans la source avant leur injection dans la zone d'accélération. Les ions de même m/z peuvent être décalés soit en raison d'un délai entre la formation des premiers et celle des derniers soit parce qu'ils acquièrent une énergie cinétique différente au moment de leur formation.

Les ions sont émis dans une zone libre de champ. A l'issue d'un délai qui peut aller de la centaine de nanosecondes à plusieurs microsecondes, l'application d'un potentiel d'extraction les dirige dans la zone d'accélération (figure 16). Les premiers ions formés ou ceux ayant acquis le plus d'énergie cinétique sont les plus proches du détecteur au moment de l'application de ce potentiel. Le potentiel électrique V_x en un point x d'un champ électrique E est égal à $V_x = E d_x$, avec d_x la distance à parcourir par les ions dans la zone du champ d'extraction. Pour un même champ électrique E , les ions ayant la plus grande valeur d_x sont soumis à un potentiel électrique V_s plus important. Ainsi lors de l'accélération, une énergie plus importante est transmise aux ions se trouvant le plus loin du détecteur (ceux formés en dernier ou ceux possédant le moins d'énergie cinétique). Ceci permet aux ions "retardataires" de rattraper les premiers car ils ont une vitesse de vol plus rapide.

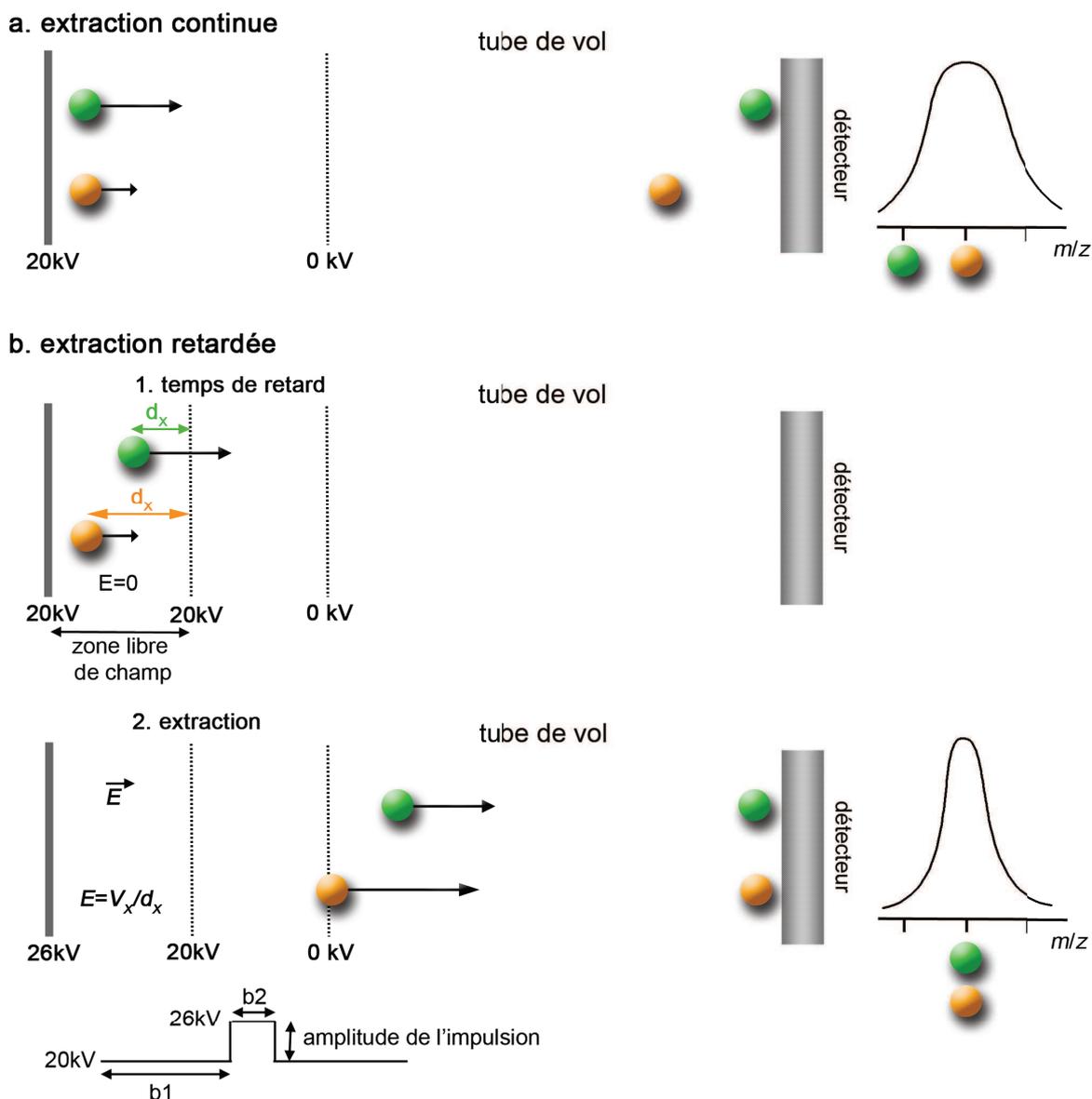


Figure 16. Principe de a. l'extraction continue et de b. l'extraction retardée.

7.2.2. Les analyseurs de type quadripôle

L'analyseur quadripolaire utilise la stabilité de la trajectoire des ions dans un champ électrique oscillant pour les séparer en fonction de leur masse. Les analyseurs de type trappe ionique sont basés sur le même principe.

7.2.2.1. Structure des quadripôles

Les analyseurs quadripolaires sont constitués de 4 barres parallèles sur lesquelles est appliqué un potentiel électrique. Les barres diamétralement opposées reçoivent le même

potentiel électrique (Φ_0 ou $-\Phi_0$) qui est de signe opposé à celui appliqué sur l'autre paire de barre (figure 17.a). Cela forme un champ électrique à deux dimensions.

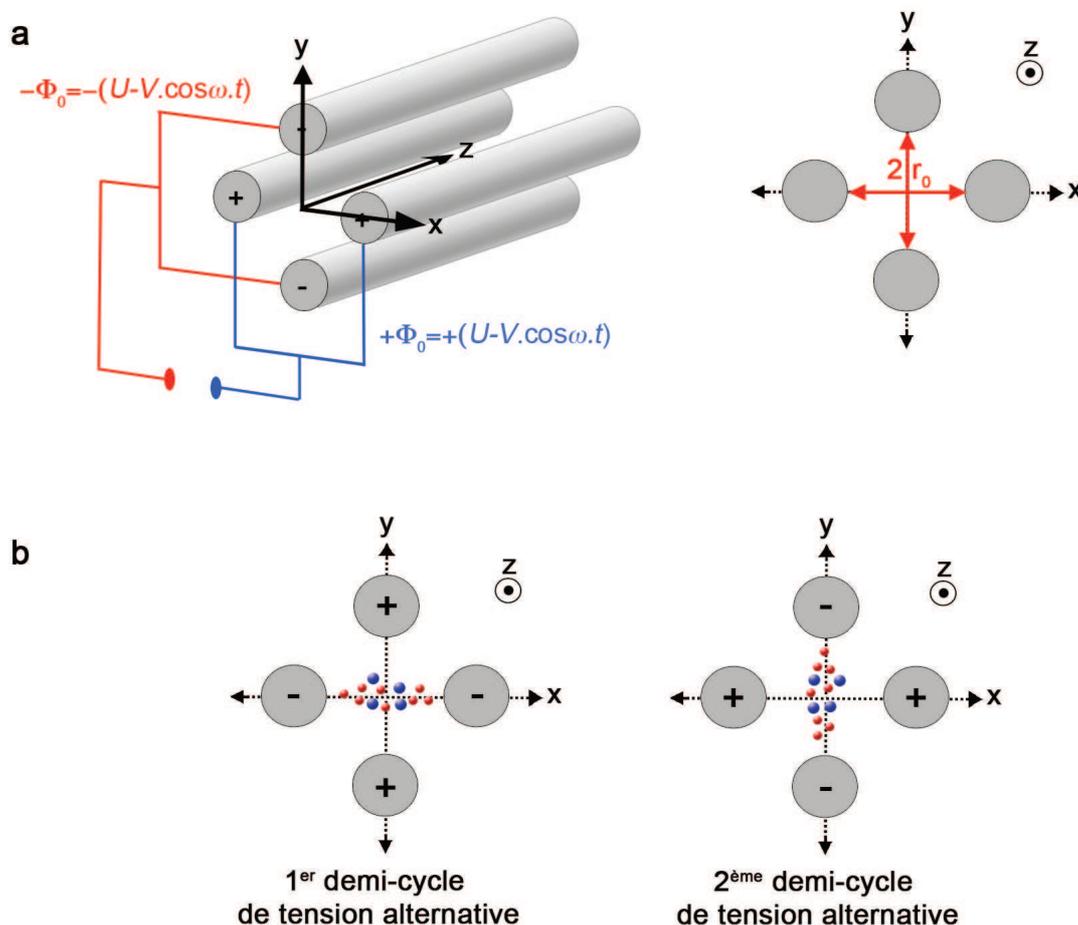


Figure 17. Principe des analyseurs de type quadripôle : a. schéma d'un analyseur quadripolaire ; b. évolution des ions chargés positivement dans un quadripôle au cours d'un cycle de tension alternative.

L'application simultanée d'une tension continue et d'une tension radiofréquence (RF) impose aux ions qui passent entre les barres un chemin oscillant. En effet, un ion positif, injecté entre les barres, est attiré vers une barre de polarité négative. Si la polarité de cette dernière change avant que l'ion n'entre en contact avec elle, l'ion changera de direction (figure 17.b).

7.2.2.2. Stabilité des ions dans un quadripôle

Le champ électrique total est composé d'un champ alternatif quadripolaire ($V \cos(\omega t)$; avec ω : fréquence angulaire et V : amplitude du voltage RF) superposé à un

champ constant U . Les potentiels, respectivement appliqués sur les deux paires de barres, sont définis par les équations suivantes :

$$\Phi_0 = +(U - V \cos \omega t) \text{ et } -\Phi_0 = -(U - V \cos \omega t)$$

ou Φ_0 est la tension appliquée aux barres.

Les ions préalablement accélérés suivant l'axe des z conservent leur vitesse suivant cet axe entre les barres du quadripôle. Mais ils sont soumis suivant x et y aux accélérations résultantes des forces dues aux champs électriques. Pour qu'un ion reste stable au sein du quadripôle, il est nécessaire que ces forces le maintiennent sur des coordonnées x, y se trouvant dans le cercle de rayon r_0 inscrit entre les barres du quadripôle. Si les oscillations d'un ion dépassent r_0 en x ou y , alors celui-ci se déchargera sur les barres du quadripôle.

Les équations des forces dues aux champs électriques permettent d'obtenir les équations du mouvement des ions en x et y . Les équations obtenues, appelées équations de Mathieu en référence au physicien qui les a établies, régissent la trajectoire des ions.

$$\frac{d^2 u}{d\xi^2} + (a_u - 2q_u \cos 2\xi)u = 0$$

$$\text{avec } u=x \text{ ou } y, \quad \xi = \frac{\omega t}{2}, \quad a_u = \frac{8zeU}{mr_0^2 \omega^2} \text{ et } q_u = \frac{4zeU}{mr_0^2 \omega^2} \text{ soit } \frac{a}{q} = \frac{2U}{V}$$

avec r_0 : le rayon du cercle inscrit entre les 4 barres ; ω : la fréquence angulaire telle que $\omega = 2\pi f$ où f est la fréquence du champ alternatif ; m : la masse de l'ion ; ze : la charge de l'ion.

La résolution de ces équations définit les zones de stabilité des ions (valeurs de a et de q pour lesquelles $u < r_0$) pour qu'ils soient transmis. A ces valeurs de a et de q correspondent des valeurs d'amplitude de U et V , et de fréquence ω . Ces valeurs varient également en fonction du rapport m/z des ions. Cette caractéristique permet d'utiliser les quadripôles comme des « filtres ioniques ».

Pour un quadripôle donné, r_0 est constant et ω est maintenu constant. U et V sont les variables. Pour un ion de rapport m/z donné, on peut calculer les valeurs de x et y en fonction de U et de V . Ainsi, dans un diagramme $a_u = f(q_u)$, on peut représenter les zones de stabilité ou zones de valeurs U et V de cet ion, où x et y restent inférieurs à r_0 (figure 18). En superposant les diagrammes de stabilité selon x et y , quatre zones possibles de travail apparaissent. Elles sont marquées de A à D. De manière générale, les analyseurs utilisent la zone A.

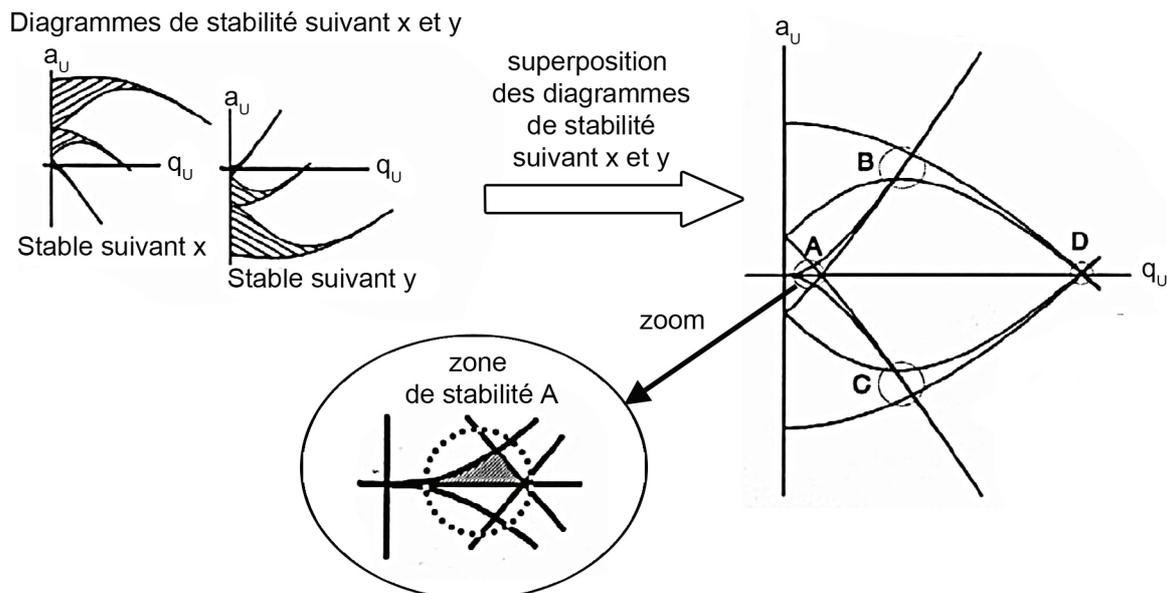


Figure 18. Zones de stabilité pour un ion de masse m donnée suivant x ou y, adapté de (de Hoffmann et Stroobant, 2005).

D'après les équations de a_u et de q_u , la modification du rapport m/z revient à diviser de manière proportionnelle a_u et q_u , ce qui correspond à changer les proportions des zones de stabilité et donc de la zone A. En d'autres termes, des ions de différents rapports m/z ont des zones de stabilité A de même aspect : elles décrivent une sorte de triangle. Par contre, les dimensions du triangle changent d'un rapport m/z à l'autre. Ces zones A peuvent également être représentées dans un diagramme U fonction de V (figure 19). Dans ce diagramme, la superposition des zones A de trois ions de rapport m/z différent (m/z_1 , m/z_2 et m/z_3) montre que ces ions ont des zones de stabilité communes et d'autres qui leur sont propres. Une droite maintenant le rapport U/V constant et traversant successivement les zones de stabilité propres à chaque rapport m/z peut être tracée (figure 19). Un balayage de la tension appliquée selon cette droite permet de faire passer successivement les ions de différents m/z à travers le quadripôle. Les ions parviennent ainsi jusqu'au détecteur en fonction de leur rapport m/z . Comme l'illustre la figure 19, l'augmentation de la pente de la droite améliore la résolution, mais diminue la sensibilité. A l'inverse, la diminution de la pente de la droite diminue la résolution. Ainsi, pour une valeur de tension continue U nulle, la pente et la résolution seront nulles, tous les ions seront transmis, c'est le mode radiofréquence uniquement (« *RF Only* »). Ce mode est très utilisé notamment dans le cadre du couplage de plusieurs analyseurs. Cependant, la valeur de V impose un minimum aux valeurs des rapports m/z stables. Si de $V=0$ on passe à une valeur de V légèrement supérieure au domaine de stabilité des ions de rapport m/z_1 (zone rouge), tous les ions d'un rapport m/z

égal ou inférieur auront une trajectoire instable au sein du quadripôle. La gamme de rapport m/z maximale qui puisse être atteinte est de l'ordre de 4000 u. Cependant, au delà de 3000 u, les massifs isotopiques ne sont plus nettement résolus.

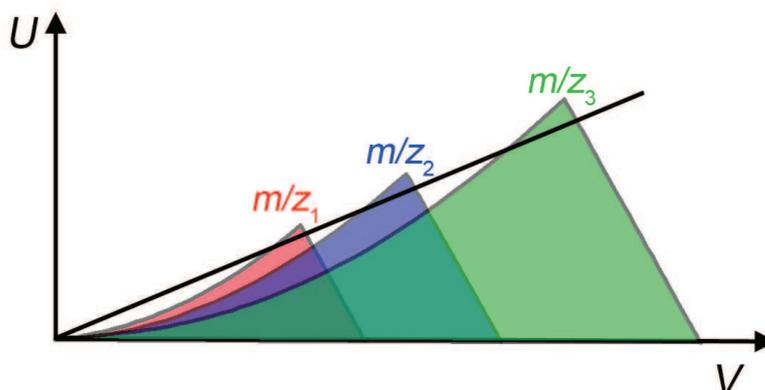


Figure 19. Zones de stabilité A en fonction de U et V pour des ions de différents rapports m/z ($m/z_1 < m/z_2 < m/z_3$), adapté de (de Hoffmann et Stroobant, 2005).

Les quadripôles travaillant en « *RF only* » ont la faculté de focaliser la trajectoire des ions en leur centre. Ainsi, ils sont très souvent utilisés comme guide d'ions ou comme cellule de collision.

7.2.3. Les analyseurs de type trappe ionique de Paul ou 3D

La trappe ionique est en fait l'homologue en trois dimensions des analyseurs de type quadripôle. A l'inverse de ces derniers, qui ne laissent passer que les ions d'un rapport m/z donné, la trappe ionique fonctionne comme un piège à ions qui accumule à l'intérieur de sa zone de champ tous les ions sélectionnés. La capacité de confinement des ions est liée à la formation d'un pseudo puits de potentiel. Les ions piégés sont ensuite libérés successivement suivant leur rapport m/z , suite à leur déstabilisation (McLuckey *et al.*, 1994).

7.2.3.1. Structure de la trappe

La trappe ionique est constituée d'une électrode annulaire et de deux électrodes «chapeaux» de section hyperbolique (Figure 20). Elle peut être vue comme un quadripôle circulaire, dont la barre intérieure est réduite à un point fictif (centre de la trappe), la barre extérieure est devenue un cercle (électrode annulaire) et les deux barres supérieure et inférieure une calotte (électrodes chapeau).

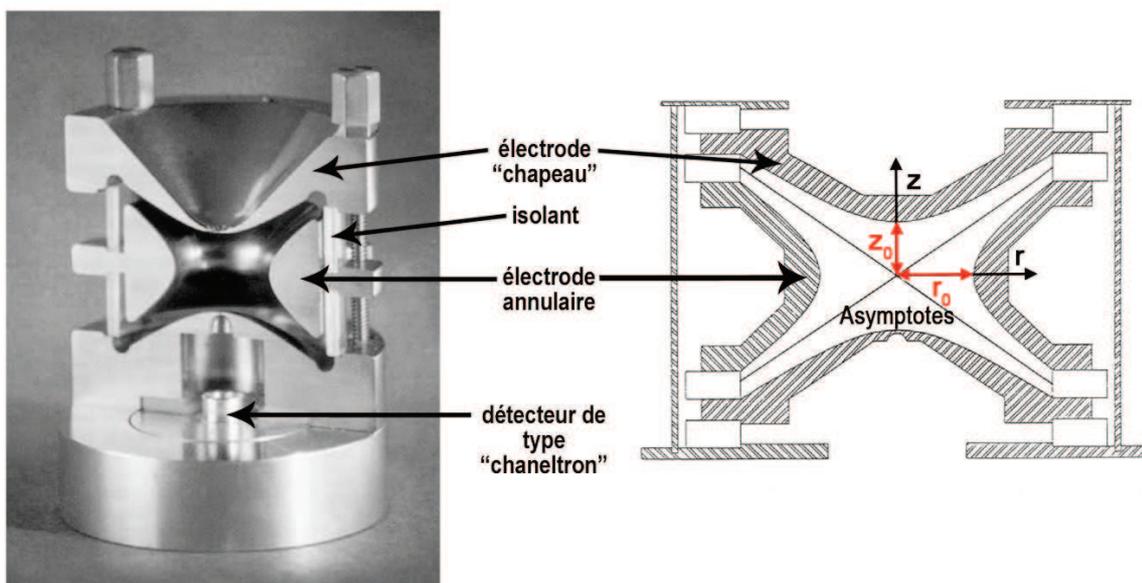


Figure 20. Photographie et schéma d'une trappe ionique 3D, adapté de (March, 1997).

7.2.3.2. Piégeage des ions

Une tension Φ_0 , superposition d'une tension continue et d'une tension alternative, est appliquée sur l'électrode circulaire. Les calottes sont généralement reliées à la terre. Il en résulte un champ électrique à trois dimensions qui piège les ions sur une trajectoire formant une sorte de huit tridimensionnel (figure 21).

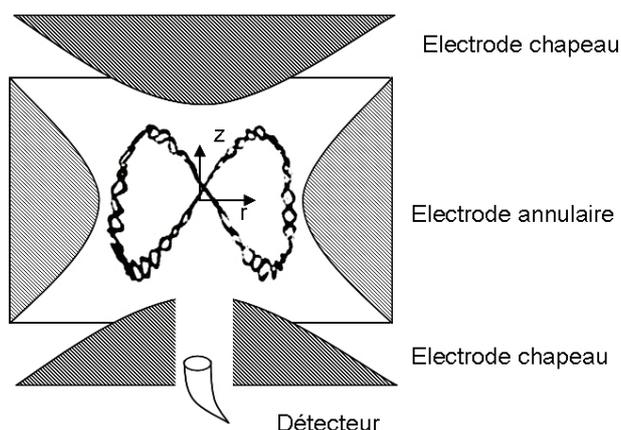


Figure 21. Trajectoire d'un ion sous forme d'un 8 tridimensionnel dans une trappe ionique.

Les ions sont piégés si leur trajectoire est simultanément stable dans le plan axial (suivant z) et dans le plan radial (suivant r). En d'autres termes, les ions sont stables tant que les accélérations qu'ils subissent les maintiennent sur des coordonnées inférieures à r_0 en r et à z_0 en z . A partir des équations des forces résultantes des champs électriques, on

peut déterminer les équations du mouvement des ions. Comme pour les quadripôles, leur analyse mathématique par les équations de Mathieu conduit à la détermination de couples de valeurs a et q pour lesquels les ions d'un rapport m/z donné sont stables soit suivant r soit suivant z (figure 22).

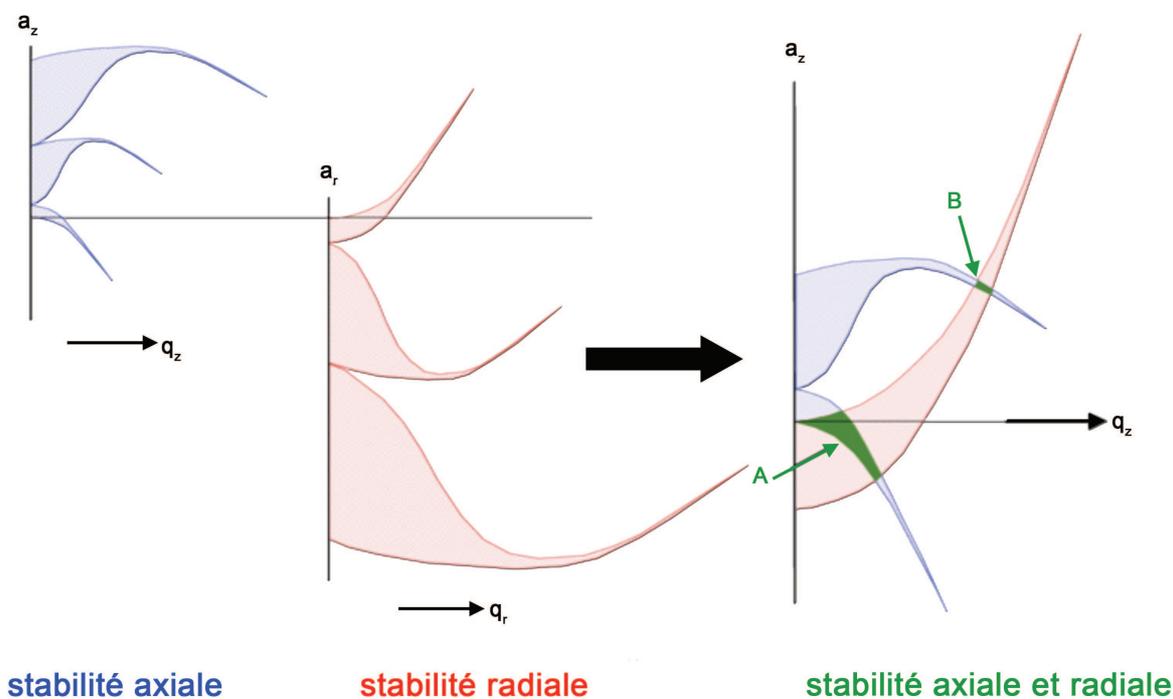


Figure 22. Diagramme de stabilité suivant r et z d'une trappe ionique. Les régions de stabilité simultanée sont marquées A et B, adapté de (March, 1997).

La superposition des domaines de stabilité axiale et radiale met en évidence deux régions de stabilité communes (qui ont les mêmes valeurs de a et de q) appelées A et B. Les trappes ioniques utilisent généralement la zone de stabilité A. Comme pour les quadripôles, il est possible de représenter dans un diagramme en fonction des valeurs de la tension continue U et de l'amplitude de la tension alternative V les domaines dans lesquels les ions sont stables (jamais hors des dimensions z_0 et r_0 de la trappe) (figure 23). Pour obtenir leur spectre de masse on expulsera les ions de la trappe en fonction de leur rapport m/z .

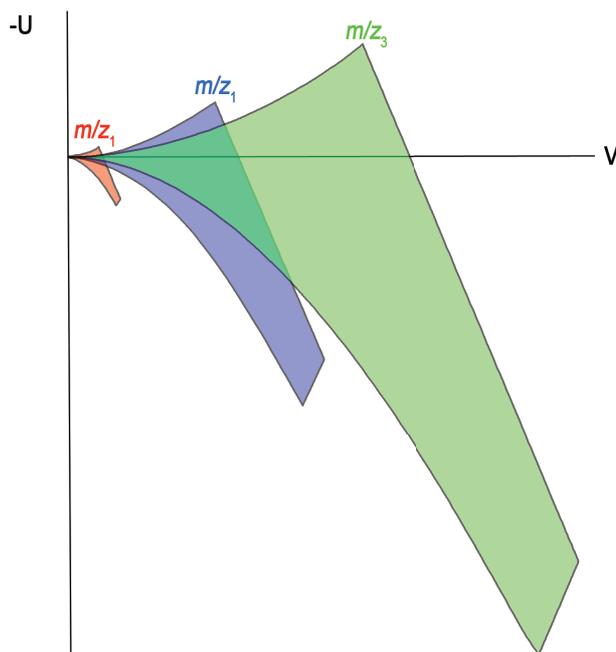


Figure 23. Domaines de stabilité A de trois ions de rapport m/z différents en fonction de U et V dans une trappe ionique, adapté de (de Hoffmann et Stroobant, 2005).

7.2.3.3. Expulsion des ions de la trappe

Dans ce type de piège, les ions d'un rapport m/z donné oscillent à une fréquence qui leur est propre suivant r et z , différente de la radiofréquence appliquée. L'application aux calottes d'une tension alternative suivant z , à la fréquence propre aux ions de ce rapport m/z permet de les expulser hors de la trappe. En effet, cette tension modifie l'accélération de ces ions suivant z , conduisant à leur déstabilisation axiale suivant cet axe. Pour sélectionner des ions d'un rapport m/z donné, on applique une onde couvrant toutes les fréquences à l'exception de la fréquence correspondant aux ions d'intérêt. On peut ensuite irradier ces ions à faible amplitude à leur fréquence propre pour les fragmenter (de Hoffmann et Stroobant, 2005).

Une des limites de ce type d'analyseur provient des effets de charge du paquet d'ions piégés. Les répulsions de charge s'exerçant entre des ions peuvent modifier la zone de stabilité de ces derniers, ce qui provoque une diminution de la résolution.

7.2.4. Les analyseurs de type trappe linéaire ou 2D

La structure des trappes linéaires est basée sur celle d'un quadripôle fermé par des électrodes d'entrée et de sortie. L'application au niveau de ces électrodes d'une tension continue, U_{DC} , de même signe que celui des ions étudiés permet de piéger ces derniers à

entre les barres du quadripôle. Plus la tension appliquée est importante, plus les ions sont comprimés au centre du piège. Les ions réalisent des allers-retours entre les électrodes terminales suivant z et, simultanément, oscillent dans le plan xy par l'application d'une tension RF (V_{RF}) sur les barres du quadripôle.

Un grand avantage des trappes linéaires est d'accepter beaucoup plus d'ions que les trappes 3D avant d'avoir un effet de charge notable.

Deux modes d'expulsion des ions sont utilisés :

1. Ejection axiale

Ejection suivant l'axe des z en utilisant les effets de champ aux extrémités (Londry et Hager, 2003).

2. Ejection radiale

L'expulsion sélective des ions se fait à travers des fentes creusées dans deux barres opposées du quadripôle, sur lesquelles est appliquée une tension alternative adaptée (figure 24).

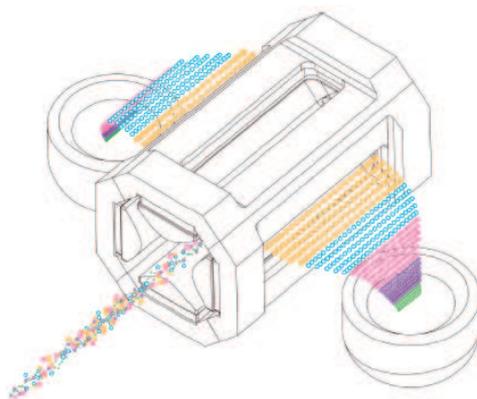


Figure 24. Trappe linéaire ou 2 D avec fentes dans les barres opposées.

Les appareils utilisant le mode d'éjection axiale permettent le couplage de différents analyseurs. Les ions peuvent être transférés d'un analyseur à l'autre suivant l'axe z . Les appareils de type QIT utilisent ce type de technologie (figure 25) et sont constitués d'une trappe linéaire entourée de deux analyseurs quadripolaires. Nous examinerons dans le chapitre consacré à la MS/MS l'intérêt d'avoir plusieurs analyseurs en série.

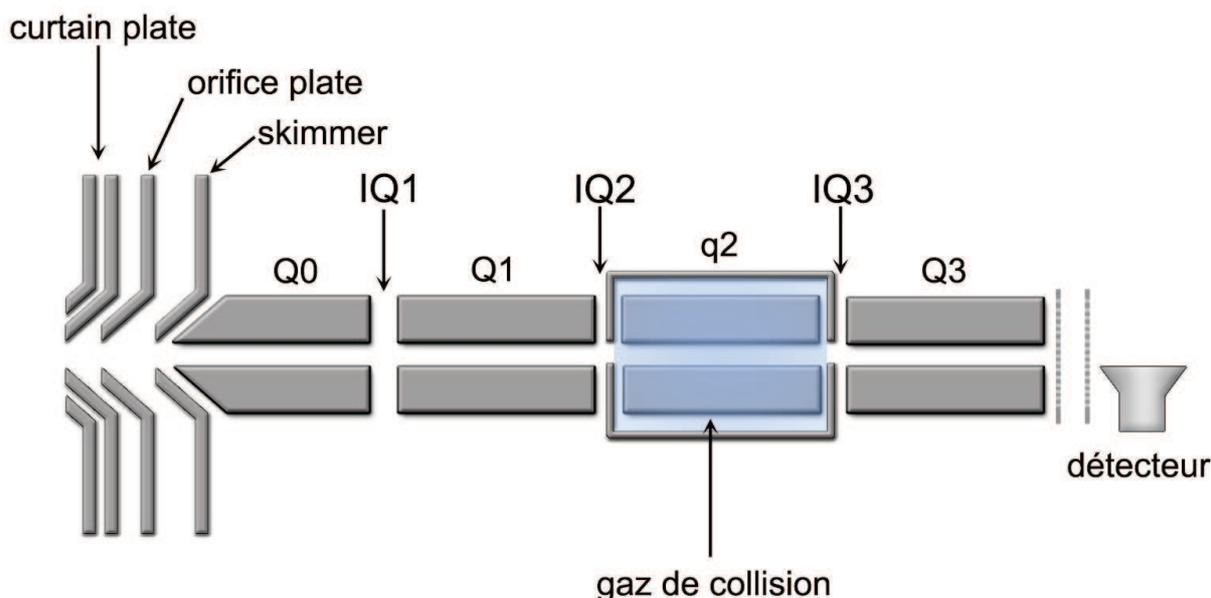


Figure 25. Schéma d'un QIT, spectromètre de masse hybride composé de 2 quadripôles en Q1 et Q3 et d'une trappe linéaire en q2, d'après (Hager, 2002). Les électrodes d'entrée et de sortie de la trappe linéaire sont marquées IQ2 et IQ3. Q0 représente un quadripôle fonctionnant en mode « RF only » (guide d'ion).

Les appareils utilisant le mode d'éjection radiale autorisent les mêmes expériences qu'une trappe 3D, comme la MS^n (expérience de fragmentation à n étapes). La capacité au niveau du nombre d'ions pouvant être piégés est environ 40 fois supérieure à celle d'une trappe 3D. L'efficacité de la capture des ions à l'injection est de 55 à 70%, contre seulement 5% dans une trappe 3D. La résolution peut aller jusqu'à 30 000 FWHM.

7.2.5. Les analyseurs de type FT-ICR

Les spectromètres de masse de type FT-ICR sont des analyseurs permettant d'obtenir des spectres MS d'une très grande résolution ($>100\ 000$) et d'une très grande précision. Un autre de leurs avantages est leur capacité à piéger les ions. Les FT-ICR sont constitués d'une cellule "Ion Cyclotron Resonance" (ICR) composée d'une paire de plaques excitatrices et d'une paire de plaques réceptrices (figure 26). Cette cellule est placée au sein d'un aimant qui génère un champ magnétique uniforme B . Les ions sont piégés au sein de la cellule ICR grâce à la combinaison du champ magnétique B et d'un champ électrique quadripolaire constant. Le champ magnétique piège les ions sur une trajectoire circulaire: c'est le principe du cyclotron. La fréquence de ce mouvement cyclotronique et sa vitesse angulaire dépendent à la fois du champ magnétique et du rapport m/z des ions piégés.

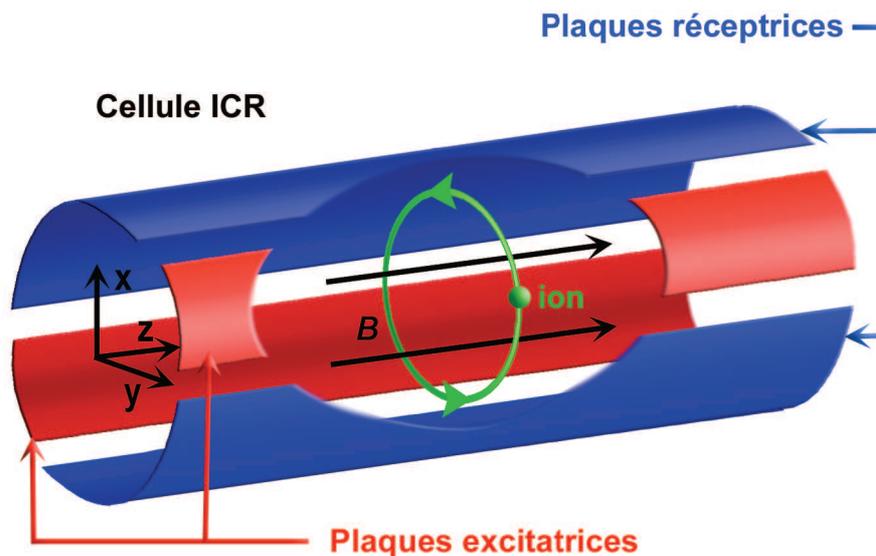


Figure 26. Schéma d'une cellule ICR. Cette cellule est placée à l'intérieur d'un aimant qui génère le champ magnétique B .

Un ion injecté dans un champ magnétique B à une vitesse v se stabilisera sur une trajectoire résultant des forces centripète, force de Lorentz, et centrifuge s'appliquant à lui :

Force centripète: $F = q v B$

Force centrifuge : $F' = \frac{m v^2}{r}$

Ainsi comme $F = F'$: $q v B = \frac{m v^2}{r}$ ou $q B = \frac{m v}{r}$

Avec q la charge de l'ion ($q = z e$), v sa vitesse, m sa masse, r le rayon de sa trajectoire circulaire et B l'intensité du champ magnétique.

L'ion boucle un trajet circulaire de $2\pi r$ à une fréquence ν_c :

$$\nu_c = \frac{v}{2 \pi r} = \frac{q B}{2 \pi m}$$

sa vitesse angulaire ω s'exprime par l'équation :

$$\omega = 2 \pi \nu_c = \frac{v}{r} = \frac{q}{m} B = \frac{e}{m/z} B$$

La fréquence et sa vitesse angulaire ne dépendent que du champ magnétique B et du rapport m/z , suivant la relation $(\frac{q}{m} B)$. Elles sont donc indépendantes de la vitesse de

l'ion. La mesure de cette fréquence ν_C permet de déterminer le rapport m/z de l'ion considéré.

Le champ magnétique piège les ions dans le plan x,y perpendiculaire à la direction du champ magnétique (figure 26). Le piégeage des ions selon l'axe des z peut être réalisé par deux plaques de piégeages placées perpendiculaires à l'axe z et portées à un potentiel répulsif. C'est le principe de la résonance d'ion en cyclotron également appelé piège de Penning. S'il était possible d'appliquer un champ électrique homogène parallèle au champ magnétique avec des plaques de longueurs infinies, le mouvement d'un ion serait hélicoïdal avec des oscillations périodiques suivant l'axe z dans le puits de potentiel électrique. Ce mouvement forme une seconde fréquence, dite fréquence de piégeage ν_T

$$\nu_T = \frac{1}{2\pi} \sqrt{\frac{2qV_{trap}\alpha}{ma^2}}$$

V_{trap} : potentiel de piégeage ; α : facteur géométrique ; a : distance entre les plaques.

On peut noter ici aussi que cette fréquence ne dépend pas de l'énergie cinétique initiales des ions.

Cependant, comme la réalisation d'électrodes de longueur infinie n'est pas possible, le champ électrique généré n'est pas parallèle au champ magnétique. Il présente une composante radiale, qui tend à entraîner les ions à l'extérieur du piège et s'oppose ainsi à la force de Lorentz créée par le champ magnétique.

Le dispositif de génération du champ électrique est généralement constitué de deux groupes d'électrodes: d'une part un anneau, et d'autre part deux coupelles axiales (figure 27). Une différence de potentiel constante est appliquée entre l'anneau et les coupelles. Les coupelles sont soumises à un potentiel de même signe que les ions à stocker, mais de signe opposé à celui du potentiel appliqué sur l'anneau. Le champ électrique ainsi créé produit un col au centre du piège, ce qui assure le confinement des ions dans la direction axiale.

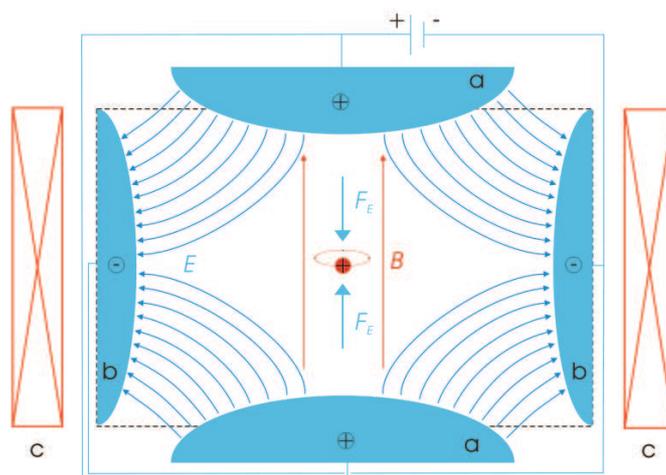


Figure 27. Schéma d'un piège de Penning contenant une particule de charge positive en son centre (en rouge). Le champ électrique E (en bleu) est produit par deux coupelles chargés positivement (a) et par une électrode circulaire chargée négativement (b). Le champ magnétique homogène B (en rouge) est produit par un électro-aimant toroïdal (c).

Au final, les particules piégées subissent trois types d'oscillations :

1) une oscillation axiale,

et deux oscillations dans le plan radial, similaires à celles d'une toupie en mouvement sur le sol (figure 28) :

2) une oscillation cyclotron réduite v_+ correspondant à l'orbite circulaire de la particule au sein du piège,

3) une oscillation magnétron v_- correspondant à la rotation de la particule « sur elle-même »

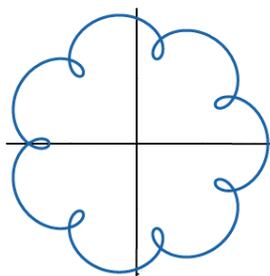


Figure 28. Trajectoire cyclotron classique d'un ion dans le plan radiale d'un piège de Penning. Elle est composée d'un mouvement cyclotron réduit v_+ et d'un mouvement magnétron v_- .

$$\text{Fréquence cyclotron réduite : } \nu_+ = \frac{\nu_c}{2} + \sqrt{\left(\frac{\nu_c}{2}\right)^2 - \frac{\nu_T^2}{2}}$$

$$\text{Fréquence magnétron : } \nu_- = \frac{\nu_c}{2} - \sqrt{\left(\frac{\nu_c}{2}\right)^2 - \frac{\nu_T^2}{2}}$$

Si aux faibles rapports m/z $\nu_+ \approx \nu_c$, pour les rapports m/z élevés (ou des fréquences basses) il existe une région pour laquelle cette approximation n'est plus valable. Ce point sera atteint d'autant plus rapidement que le potentiel de piégeage est élevé, d'où l'intérêt d'utiliser un potentiel le plus faible possible (typiquement 0.5 à 2 V) même si l'efficacité de piégeage en est amoindrie.

Une fois piégés dans la cellule ICR, les ions ont des rayons d'orbite et des positions d'orbite différentes. Ceci rend la mesure de leur fréquence cyclotronique impossible. Pour que cette mesure soit possible, elle doit être effectuée en deux étapes (figure 29).

Première étape : Mise en phase des ions

Durant cette étape, un courant électrique est appliqué entre 2 plaques d'excitation opposées. Lorsque l'irradiation a lieu à la fréquence cyclotronique d'un ion de rapport m/z avec une tension V_0 appliquée pendant un temps T_{exc} , le rayon de l'orbite de l'ion est donné

$$\text{par l'équation : } r = \frac{V_0 T_{exc}}{B_0}$$

Ce rayon dépend de la tension et de la durée de l'excitation. Le contrôle de ces deux paramètres permet d'amener les ions sur une orbite avec un rayon approprié pour les faire passer à proximité des plaques réceptrices. L'excitation à l'aide d'une seule impulsion qui balaye une large gamme de fréquences amènera tous les ions présents dans la cellule sur un même rayon mais à des fréquences propres à chaque rapport m/z . Pour que les ions soient en phase, l'impulsion est très rapide (de l'ordre de la μs).

Deuxième étape : Mesure de l'ensemble des fréquences de rotation des ions

Durant cette étape, l'intensité du courant induit sur les deux plaques réceptrices par la rotation cohérente de l'ensemble des ions est mesurée et digitalisée. Le signal intensité en fonction du temps va être converti par la transformée de fourrier en une relation intensité en fonction de la fréquence cyclotronique, puis en intensité en fonction du rapport m/z .

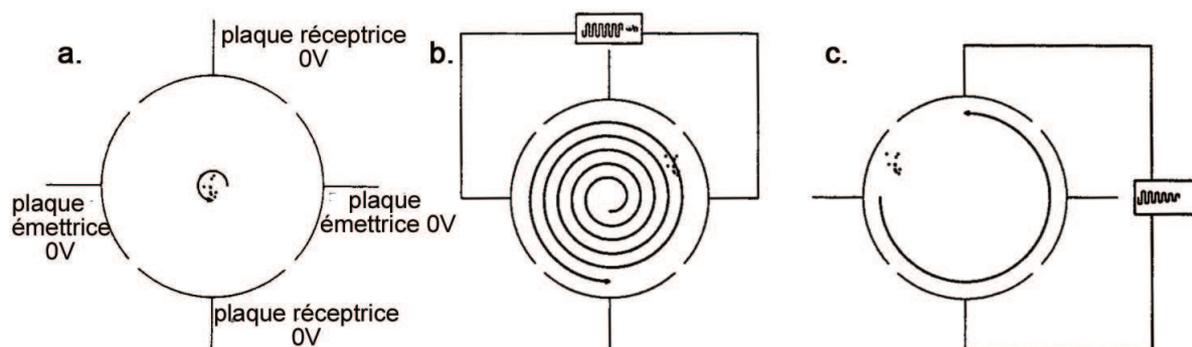


Figure 29. Etapes de la mesure de masse en FT-ICR : a. Ions avant excitation ; b. excitation des ions jusqu'à atteindre une certaine orbite ; c. mouvement cohérent des ions de même m/z créant le courant induit.

La résolution de ce type d'appareil dépend, comme dans toute technique utilisant la transformée de Fourier, du temps d'acquisition. Ce dernier est lié à la décroissance du signal détecté, résultant du ralentissement des ions par collision ion-molécules qui entraîne une diminution du rayon de leur orbite. Ainsi, ce type d'appareil nécessite un vide poussé.

La limite principale de ce type d'appareil concerne leur gamme dynamique. En effet le nombre d'ions présents dans la cellule ICR ne peut excéder 10^5 afin d'éviter un effet de charge trop important.

Les principales caractéristiques de ces analyseurs ont été résumées dans le tableau 4.

Tableau 4. Analyseurs utilisés au cours de cette étude.

Type	Acronyme	Principe	Limites en masse , résolution et précision	Avantages
Time-of-flight	TOF	Les ions sont injectés sous la forme d'un paquet d'ions et se séparent au cours du temps en fonction de leur rapport m/z.	<p><u>Limite en masse:</u></p> <ul style="list-style-type: none"> > 1 000 000 10 000 avec réflectron <p><u>Résolution:</u></p> <ul style="list-style-type: none"> 5 000 20 000 avec réflectron <p>Précision :</p> <ul style="list-style-type: none"> 0.02% 0.001% avec réflectron 	<ul style="list-style-type: none"> Gamme m/z illimitée En MALDI un spectre par tir laser Grande sensibilité Expérience MS/MS possible sur les appareils hybrides Rapidité d'acquisition des spectres large gamme dynamique
Quadrupôle	Q	Les ions circulent sous la forme d'un courant d'ions et sont séparés sur la base de la stabilité de leur trajectoire dans un champ électrique oscillant. Le domaine de stabilité des ions variant en fonction de leur rapport m/z, il est possible de transmettre sélectivement les ions vers le détecteur.	<p><u>Limite en masse:</u></p> <ul style="list-style-type: none"> 4 000 <p><u>Résolution:</u></p> <ul style="list-style-type: none"> 2 000 <p>Précision:</p> <ul style="list-style-type: none"> 0,01% 	<ul style="list-style-type: none"> Analyseur peu sensible à la pression Refocalisation des ions Peut servir de cellule de collision ou de guide d'ion à travers 2 régions du spectromètres Balayage de masse uniforme Vitesse de balayage pouvant être élevée
Linear quadrupole ion trap ou trappe ionique linéaire	LIT	Les ions sont piégés dans une trappe quadrupolaire linéaire dans laquelle ils circulent linéairement et sont séparés par la stabilité de leurs trajectoires dans le champ de radiofréquences du quadrupôle. La stabilité des ions varie en fonction de leur rapport m/z.	<p><u>Limite en masse:</u></p> <ul style="list-style-type: none"> 6 000 <p><u>Résolution:</u></p> <ul style="list-style-type: none"> 4 000 <p>Précision</p> <ul style="list-style-type: none"> 0,01% 	<ul style="list-style-type: none"> Analyseur peu sensible à la pression MSⁿ Piégeage des ions Fragmentation laser possible
Quadrupole ion trap ou trappe ionique 3D ou de Paul	QIT	Les ions sont piégés dans un champ électrique oscillant dans une trappe quadrupolaire 3D et sont séparés par la stabilité de leurs trajectoires dans le champ de radiofréquences du quadrupôle. La stabilité des ions varie en fonction de leur rapport m/z.	<p><u>Limite en masse:</u></p> <ul style="list-style-type: none"> 6 000 <p><u>Résolution:</u></p> <ul style="list-style-type: none"> 4 000 <p>Précision:</p> <ul style="list-style-type: none"> 0,01% 	<ul style="list-style-type: none"> Analyseur peu sensible à la pression MSⁿ Piégeage des ions Fragmentation laser possible
Fourier Transform-ion cyclotron resonance	FT-ICR	Les ions sont piégés dans une cellule ICR puis sont séparés en fonction de leur fréquence cyclotron dans un champ magnétique qui varie en fonction de leur rapport m/z.	<p><u>Limite en masse:</u></p> <ul style="list-style-type: none"> 30 000 <p><u>Résolution:</u></p> <ul style="list-style-type: none"> 1 000 000 <p>Précision:</p> <ul style="list-style-type: none"> <0,0005% 	<ul style="list-style-type: none"> Très grande résolution Très grande précision en masse MSⁿ

7.3. La spectrométrie de masse en tandem ou MS/MS

La spectrométrie de masse en tandem (MS/MS) a introduit une nouvelle dimension dans l'analyse de mélanges complexes en permettant de déterminer la structure d'un ion sur la base de son profil de fragmentation. Elle permet d'isoler puis de fragmenter un ion parent, avant de reconstituer brique par brique sa structure à partir des ions fragments obtenus. Le développement des techniques de MS/MS a permis une multiplication de ses domaines d'applications, en particulier dans les sciences du vivant (McLafferty, 1981 ; McLafferty, 2008).

La spectrométrie de masse en tandem tire son nom de la conception des premiers instruments ayant permis son application, constitués de deux analyseurs de masse en tandem séparés par une zone de réaction (figure 30). Les analyses de type MS/MS sont typiquement constituées de trois étapes. Au cours de la première étape (temps 1), le premier analyseur sépare les ions formés dans la source d'ionisation selon leur rapport m/z . Durant la deuxième étape, ces ions sont fragmentés dans la zone de réaction. Dans la troisième étape, les fragments résultants, s'ils sont chargés, sont séparés en fonction de leur rapport m/z dans le second analyseur de masse. Enfin, les ions sont détectés par un multiplicateur d'électrons comme pour des spectromètres de masse conventionnels. Le spectre de fragmentation de l'ion parent apporte des informations supplémentaires sur celui-ci. Cette technique s'est ainsi révélée très utile dans le domaine de la protéomique car elle permet de déterminer les séquences complètes de peptides ou de protéines à l'issue des étapes de fragmentations (Hunt *et al.*, 1986).

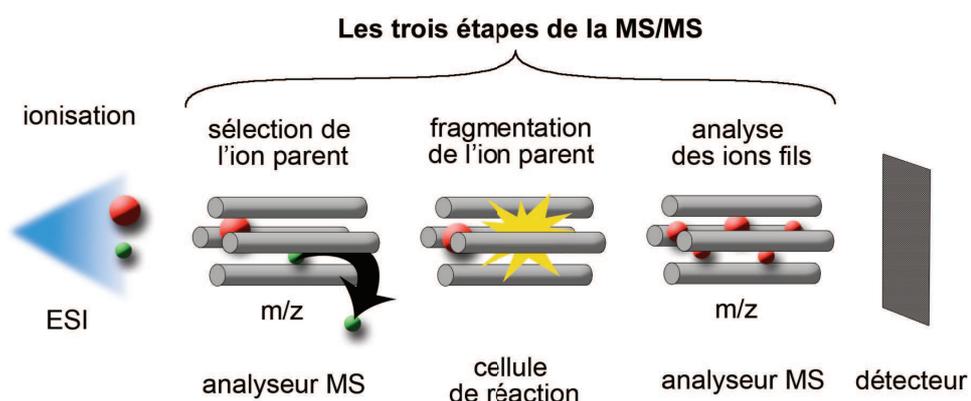


Figure 30. Principe de la spectrométrie de masse en tandem (MS/MS).

A l'heure actuelle, il existe différents types d'appareils capables de réaliser la spectrométrie de masse en tandem. Ils se décomposent en trois catégories (tableau 5) :

- la **MS/MS dans l'espace ou spatiale** (deux analyseurs séparés par une cellule de collision, i.e. triple quadripôle, secteurs magnétiques et spectromètres de masse hybrides dont les plus courants sont actuellement les Q-q-Q, Q-TOF et QIT). Les trois étapes se déroulent dans des analyseurs différents qui sont couplés les uns aux autres ;

- la **MS/MS dans le temps ou temporelle** (un seul analyseur, e.g. IT, FT-ICR). Les trois étapes se déroulent au sein du même analyseur mais à des temps différents (T_1 , T_2 et T_3) ;

- et le cas particulier du **PSD (Post-Source Decay** sur les instruments de type MALDI-TOF) qui est un mélange des deux précédents.

Tableau 5. Tableau de comparaison MS/MS spatiale versus MS/MS temporelle.

	Sélection de l'ion parent	Fragmentation de l'ion parent	Analyse des ions fils
MS/MS temporelle dans un Q-q-Q	Sélection dans le 1 ^{er} quadripôle	Fragmentation dans le 2 ^{ème} quadripôle	Acquisition des spectres MS/MS dans le 3 ^{ème} quadripôle
MS/MS spatiale dans une IT	Sélection de l'ion parent dans le piège ionique (T_1)	Fragmentation de l'ion parent dans le piège ionique (T_2)	Acquisition des spectre MS/MS dans le piège ionique (T_3)

Les instruments réalisant de la MS/MS spatiale permettent en général la répétition des trois étapes MS/MS n fois à la suite, on parle alors de MS^n . Un ion fils de première génération est sélectionné puis fragmenté produisant des ions fils de seconde génération. A leur tour ces derniers peuvent être sélectionnés et fragmentés. Le nombre des informations acquises et disponibles pour l'identification des molécules analysées augmente considérablement.

Différentes méthodes de fragmentation existent. Pour un même échantillon, les profils de fragmentation dépendent de la méthode utilisée. Le mode le plus répandu est appelé dissociation induite par collision (CID, *Collision-Induced Dissociation*) car les ions parents sont activés suite à leur collision contre une ou plusieurs molécules de gaz neutre. La dissociation par capture d'électrons (ECD, *Electron Capture Dissociation*) ainsi que sa variante la dissociation par transfert d'électrons (ETD, *Electron Transfert Dissociation*) reposent sur une activation des ions par la capture d'un ou de plusieurs électrons.

7.3.1. La dissociation induite par collision (CID)

Cette méthode de fragmentation est basée sur une activation des ions parents par transfert d'énergie à la suite de collisions inélastiques contre un gaz neutre (Azote, Hélium...) (de Hoffmann et Stroobant, 2005). Le processus de collision s'effectue en deux étapes (figure 31). La première correspond à la collision entre l'ion et le gaz cible amenant une partie de l'énergie cinétique de l'ion à se convertir en énergie interne et conduisant l'ion à un état d'excitation. La deuxième étape correspond à une décomposition unimoléculaire de l'ion activé : le chauffage vibrationnel global de la molécule induit la fragmentation de la molécule dès que les voies de dissociation nécessitant le moins d'énergie sont accessibles.

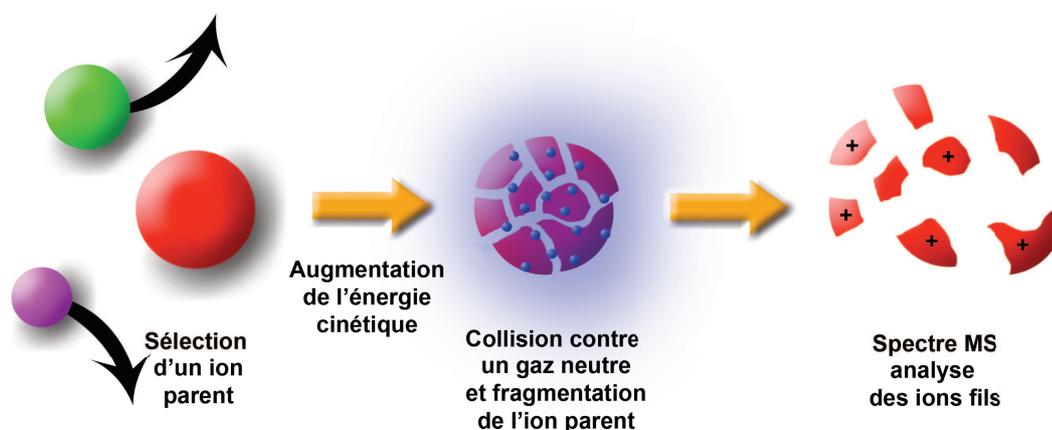


Figure 31. Principe de la MS/MS de type CID.

La MS/MS de type CID permet de séquencer et donc d'identifier les protéines. En effet, leurs fragmentations s'effectuent toujours au niveau des liaisons peptidiques. Les processus de fragmentation de la chaîne peptidique sont influencés par leur environnement et donc par la nature des chaînes latérales. Une nomenclature de la fragmentation des peptides a été établie en fonction de la localisation du clivage et de la partie de la chaîne peptidique portant la charge (partie N-terminale ou C-terminale) (figure 32) (Roepstorff, 1984 ; Biemann, 1988). A partir du type de fragments obtenu, des séries d'ions fragments sont définies. Les séries de fragments de type *b*, *c* et *y* sont généralement les plus abondantes lors des fragmentations MS/MS de type CID de basse énergie. La différence de masse entre des ions consécutifs d'une même série permet de déterminer l'identité et l'enchaînement des acides aminés (a.a.).

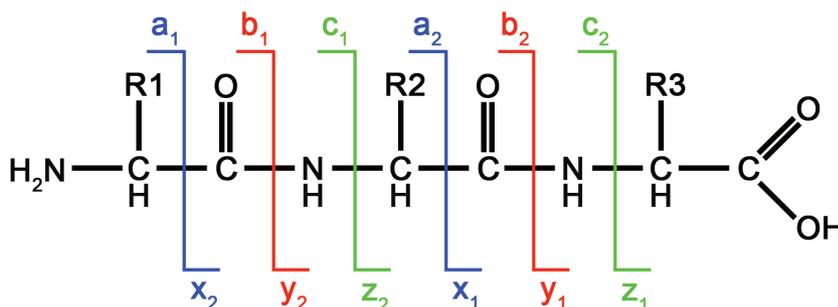


Figure 32. Nomenclature de la fragmentation peptidique en MS/MS.

Une des limites de cette technique est qu'elle n'apporte pas d'information sur la localisation des modifications post-traductionnelles (PTM, *Post-translational Modification*) (glycosylations, phosphorylations..); celles-ci sont rapidement perdues lors du mécanisme de fragmentation mis en jeu.

La MS/MS de type CID peut également être utilisée pour étudier la stabilité des complexes non-covalents. La collision de ces derniers avec les molécules de gaz neutre entraîne la dissociation des différents partenaires (figure 33). La comparaison des énergies internes nécessaires pour rompre les forces de liaison de différents complexes permet d'établir, par exemple, un classement entre plusieurs ligands en fonction de la force de leur interaction avec une même protéine (Daniel *et al.*, 2002 ; Jorgensen *et al.*, 2002).

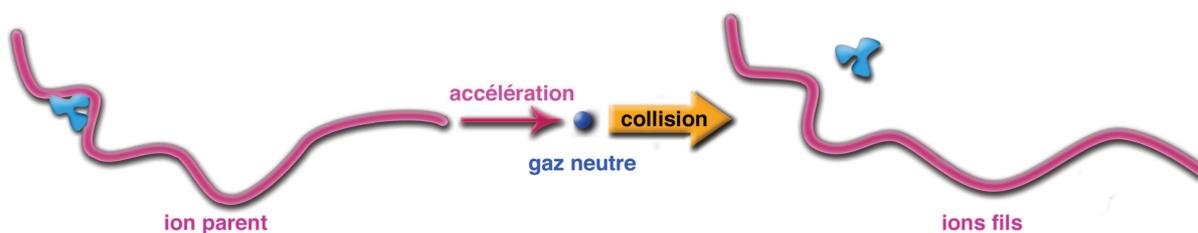


Figure 33. Utilisation de la CID pour l'étude de la dissociation des complexes non-covalents.

L'évaluation de l'énergie interne ajoutée à un ion lors de ce type d'expérience nécessite de prendre en compte le processus de transfert d'énergie qui s'effectue lors des collisions.

Au cours des collisions du complexe ionisé sélectionné avec les molécules de gaz neutre, une partie de l'énergie cinétique de ce complexe est convertie en énergie interne. Il s'agit de collisions inélastiques, au cours desquelles l'énergie cinétique n'est pas conservée.

Les transferts d'énergie cinétique en énergie interne sont régis par les lois des collisions d'un mobile (l'ion) contre une cible statique (le gaz de collision). Dans une collision inélastique, la conservation de l'énergie et de la quantité de mouvement implique que seule une partie de l'énergie cinétique est transformée en énergie interne. Cette fraction d'énergie est donnée par l'équation suivante :

$$E_{CoM} = E_{lab} \frac{m_2}{(m_1 + m_2)}$$

avec m_1 la masse de l'ion, m_2 la masse du gaz cible, E_{lab} l'énergie cinétique de l'ion dans le référentiel du laboratoire et E_{CoM} l'énergie maximale qui peut-être convertie en énergie interne. Ainsi, pour augmenter l'énergie disponible à la conversion, il faut soit augmenter l'énergie cinétique de l'ion sélectionné, soit utiliser un gaz cible ayant une masse plus importante.

Les appareils permettant la MS/MS CID de basse énergie (entre 1 et 100 eV ; 1 eV par ion équivaut à environ 100 kJ/mole) possèdent en général une cellule de collision constituée d'un quadripôle « *RF only* » dont le pouvoir de focalisation des ions est important. La cellule a généralement une longueur assez élevée. Ainsi les ions subissent un grand nombre de collisions avec les molécules de gaz. Ce nombre de collisions peut-être calculé à partir de l'équation suivante :

$$N_{col} = n \sigma l$$

avec n représentant la densité numérique de gaz, σ correspondant à la section efficace de collision de l'ion et l représentant la longueur de la cellule. Finalement, l'énergie interne transférée à l'ion peut être calculée par l'équation suivante :

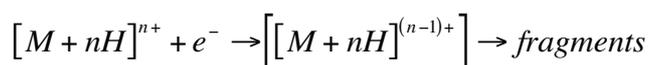
$$E_{Int} = E_{CoM} N_{col}$$

A cause du grand nombre de collisions, l'énergie interne transférée à l'ion est très proche de l'énergie interne maximale communicable.

La comparaison des énergies internes transférées aux complexes se fait généralement à partir de l'énergie de demi-dissociation des complexes. Cette dernière est déterminée en faisant varier les énergies cinétiques de collision. Les intensités relatives des ions parents sont déterminées à partir des intensités des pics présents sur les spectres MS/MS enregistrés pour différentes énergies de collision. Le graphique de l'intensité relative des ions parents en fonction de l'énergie de collision est tracé et l'énergie de demi-dissociation est déterminée graphiquement.

7.3.2. La dissociation par capture d'électrons (ECD)

La dissociation par capture d'électrons (ECD, *Electron Capture Dissociation*) consiste à envoyer des électrons faiblement énergétiques en direction d'ions préalablement capturés dans un piège à ions (figure 34) (Zubarev, 2004). Cette technique implique en règle générale un ion multichargé M, qui va capturer un ou plusieurs électrons libres (McLafferty *et al.*, 2001):



L'ECD est particulièrement employée pour l'étude des protéines. En effet, l'excitation électronique induite conduit à des voies de fragmentation différentes et complémentaires (principalement de type *c* et *z* (Liu et Håkansson, 2006)) de celles observées par CID (principalement *b* et *y*) (Creese et Cooper, 2007). En effet, même si le mécanisme exact de clivage en ECD reste sujet à discussion, il s'agit d'un processus rapide (non-ergodique) ne produisant pas forcément la rupture de la liaison la plus faible (Zubarev *et al.*, 1998).

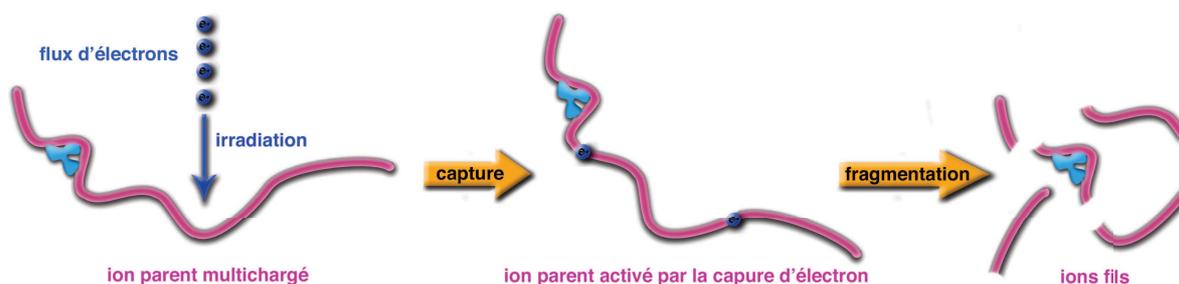


Figure 34. Principe de la fragmentation de type ECD appliquée à l'étude d'un complexe non-covalent.

Ainsi à la différence de la CID, l'ECD peut fragmenter la chaîne peptidique d'une protéine sans avoir rompu les liaisons les plus faibles au préalable. Ainsi, l'ECD est particulièrement adaptée à l'étude de la localisation des modifications post-traductionnelles (Cooper *et al.*, 2005) telles que les glycosylations (Mirgorodskaya *et al.*, 1999 ; Renfrow *et al.*, 2005) ou les phosphorylations (Shi *et al.*, 2000 ; Chalmers *et al.*, 2004 ; Creese et Cooper, 2008). Elle peut également permettre de déterminer les sites d'interactions non-covalentes d'une protéine avec un ligand (Xie *et al.*, 2006) (figure 34).

7.4. La spectrométrie de masse couplée à la spectrométrie de mobilité ionique (IMS-MS)

Le couplage de la spectrométrie de masse avec la spectrométrie de mobilité ionique (IMS-MS, *Ion Mobility Spectrometry – Mass Spectrometry*) ajoute une dimension aux spectres de masse. En effet, la mobilité ionique permet de séparer les ions en fonction de leur structure en phase gazeuse. Un tube de mobilité est inséré au sein d'un spectromètre de masse classique entre la source d'ionisation et l'analyseur (figure 35).

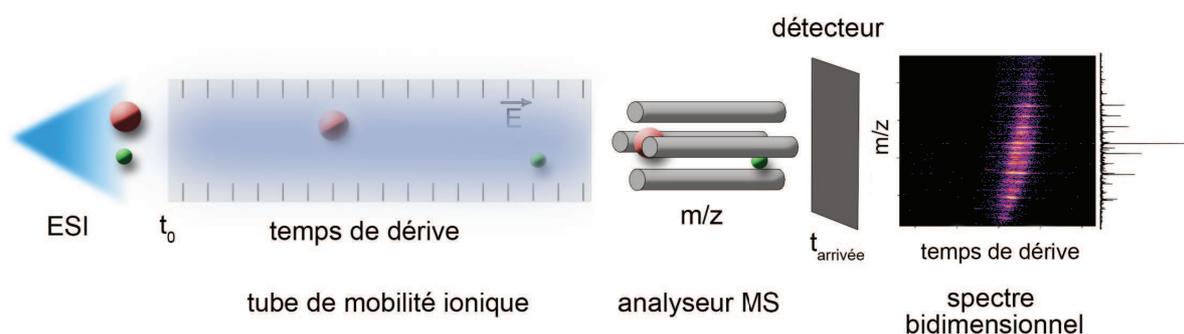


Figure 35. Principe du couplage IMS-MS.

A l'intérieur du tube de mobilité ou cellule de dérive, un champ électrique homogène assure la dérive des ions, alors qu'un gaz tampon à pression élevée permet leur séparation en fonction de leur conformation tridimensionnelle en phase gazeuse. Plus un ion a une conformation étendue, plus il met de temps à traverser le tube de mobilité. L'analyseur de masse sépare ensuite les ions en fonction de leur rapport m/z . Typiquement, les spectres IMS-MS sont à deux dimensions : le rapport m/z des ions et leur temps de dérive. L'intensité des différentes espèces est montrée par des variations de couleurs (figure 36).

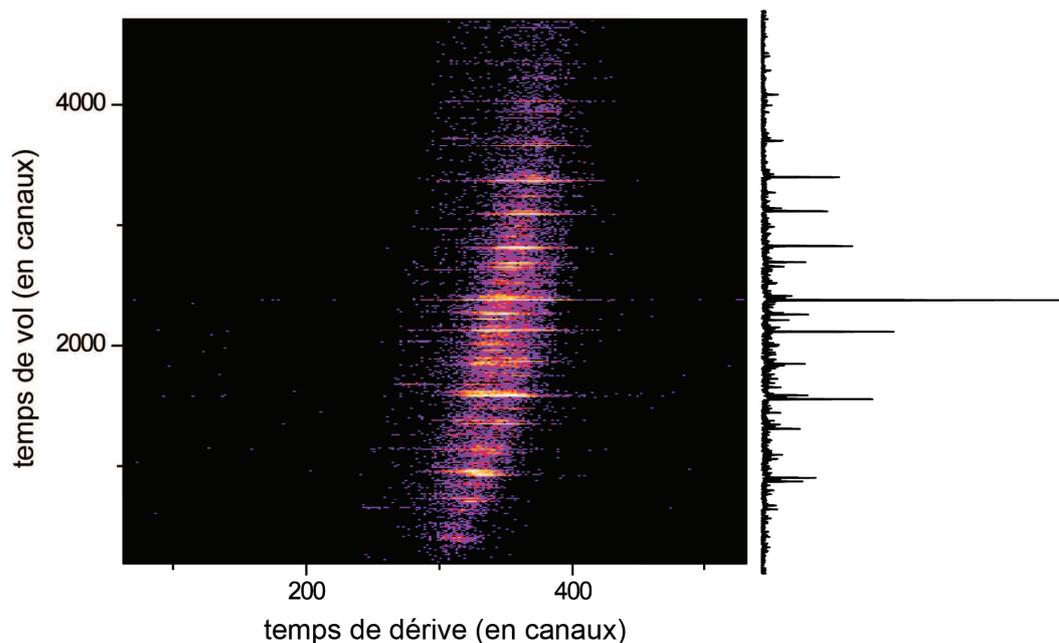


Figure 36. Spectre bidimensionnel IMS-MS. Le spectre de masse est en ordonnée et le spectre de mobilité ionique est en abscisse. L'intensité des signaux est indiquée par les variations de couleur.

7.4.1. Instrumentation

Un appareil de IMS-MS est classiquement composé d'une source d'ionisation (souvent de type ESI), d'une zone de guidage et de piégeage des ions en amont d'une ou de plusieurs « cellules de dérive », suivie par une seconde zone de focalisation et de guidage précédant un analyseur en masse en position terminale (figure 37). Au niveau de la chambre de mobilité ionique, une pression élevée de gaz tampon (de l'ordre de 5 Torr d'hélium) et un champ électrique sont maintenus.

La détection des ions est réalisée au niveau du détecteur de masse. Le temps d'arrivée des ions correspond au temps qu'ils mettent pour parcourir le trajet allant du piège ionique, se trouvant en amont de la (des) « cellule(s) de dérive », au détecteur. La traversée de l'analyseur représente un temps mort mesuré et pris en compte dans la détermination des temps de mobilité.

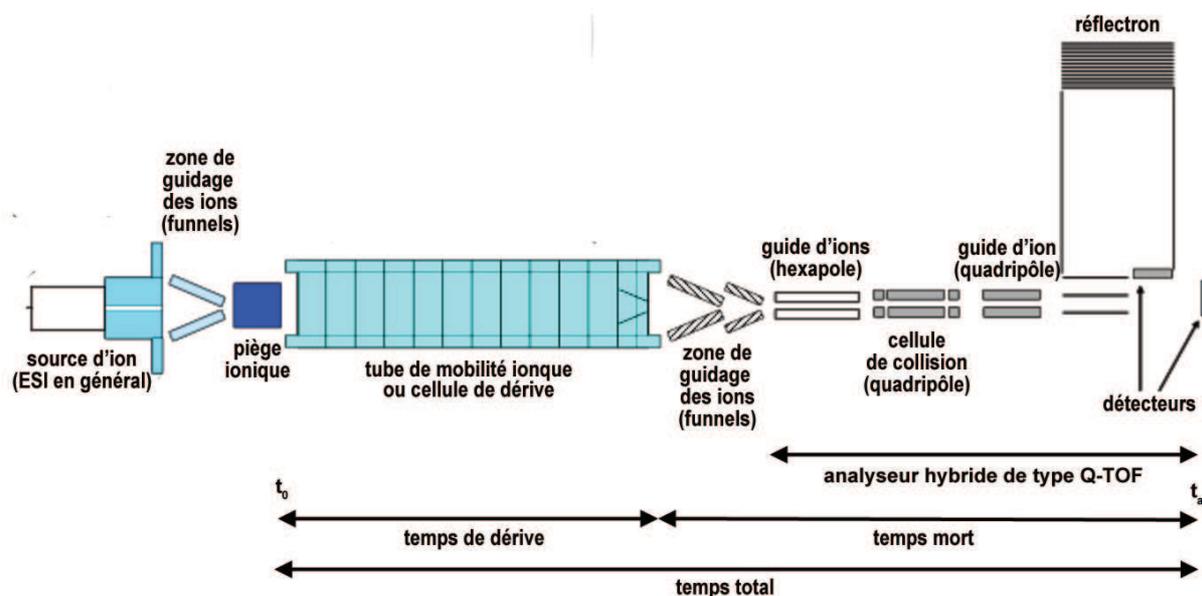


Figure 37. Structure d'un appareil de IMS-MS à une seule cellule de mobilité avec un analyseur de type Q-TOF.

7.4.2. Principe

La mesure du temps de dérive nécessite de définir un temps de départ t_0 . L'ESI est une technique d'ionisation continue qui ne permet pas de déterminer un t_0 . L'insertion d'un piège à ions entre l'ESI et le tube de mobilité permet d'envoyer les ions dans le tube de mobilité sous la forme de paquets d'ions compacts et de définir un t_0 . A l'intérieur du tube de mobilité, les ions sont à la fois accélérés par le champ électrique et freinés par les nombreuses collisions élastiques (non-énergétiques) avec le gaz tampon (figure 38). Le champ électrique E appliqué dépend du voltage appliqué V et de la longueur de la cellule L .

Donc :

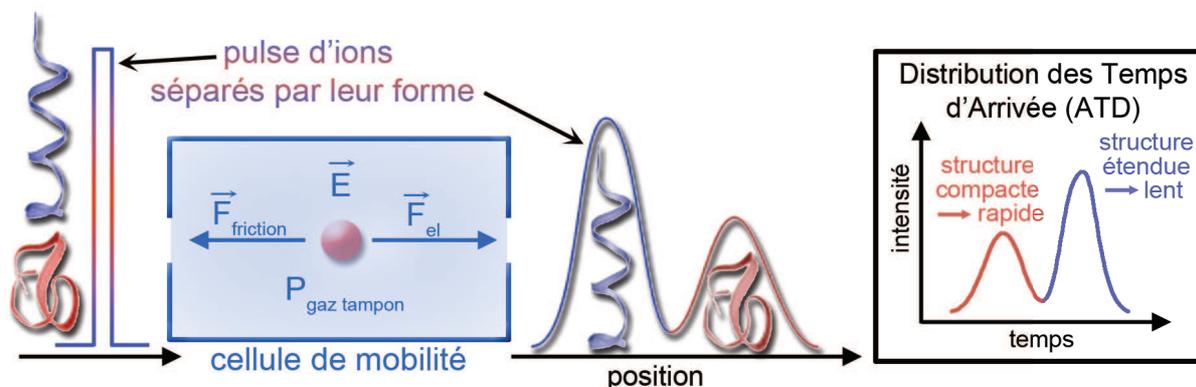
$$E = \frac{V}{L}$$


Figure 38. Distribution des temps d'arrivée en mobilité ionique.

Il s'établit un régime stationnaire : les ions atteignent une vitesse de translation constante. Le rapport entre la vitesse de « drift » (dérive) V_d et le champ électrique appliqué E est appelé la mobilité ionique K . Plus l'ion a une conformation compacte, moins il subit de collisions, moins il est freiné et donc plus sa vitesse de croisière est élevée. Un ion compact est caractérisé par une valeur de K élevée, et arrivera plus rapidement au détecteur qu'un ion de conformation plus étendue. Sa vitesse de dérive est ainsi fonction du champ électrique et de sa mobilité ionique.

$$V_d = K E$$

La mobilité ionique est directement proportionnelle à la pression de gaz. Elle est typiquement convertie en une mobilité réduite K_0 correspondant à la mobilité de l'ion dans les conditions standard pour une température T_0 , égale à 273,15 K, et une pression P_0 , égale à 760Torr.

$$K_0 = K \frac{P}{P_0} \frac{T_0}{T}$$

La combinaison de ces équations permet d'exprimer le temps de dérive t_D en fonction des conditions expérimentales de pression et de température (P , T) au niveau du tube et de la tension appliquée V :

$$t_D = \frac{L}{v_D} = \frac{L}{K E} = \frac{L^2 T_0}{K_0 P_0 T} \frac{P}{V}$$

Les données enregistrées expérimentalement correspondent à l'intensité du signal du courant ionique en fonction du temps d'arrivée. Ces données sont appelées distribution des temps d'arrivée (ATD). Le temps d'arrivée (t_a) correspond au temps mis par les ions pour arriver au détecteur à partir de leur temps d'injection (t_0). Le temps d'arrivée est égal à la somme du temps de dérive des ions à travers la cellule (t_D) et du temps passé par les ions en dehors de la cellule de mobilité t_m (temps mort).

$$t_a = t_D + t_m$$

$$\text{ainsi : } t_a = t_m + \left(\frac{L^2 T_0}{P_0 T} \right) \frac{1}{K_0} \frac{P}{V}$$

La relation t_a en fonction de P/V est une corrélation linéaire dont la pente de la droite est égale à : $\frac{1}{K_0}$ (figure 39). L'interception de cette droite avec l'axe des ordonné permet la

détermination du temps mort, t_m . Connaissant le temps mort, il est alors possible de calculer le temps de dérive à partir de : $t_D = t_a - t_m$. Typiquement, les temps de dérive vont de 100 μ s à 10ms et sont à la fois fonction de la tension appliquée et de la conformation des ions pour une pression et une température donnée.

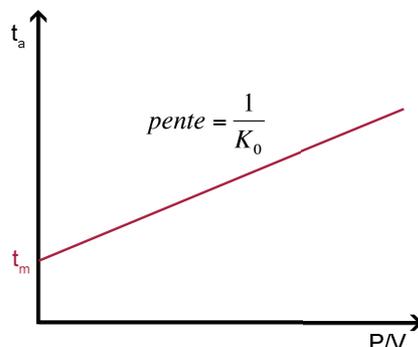


Figure 39. Corrélation entre le temps d'arrivée des ions et le rapport pression de gaz dans le tube de mobilité sur la tension appliquée.

Ainsi, la façon la plus simple de déterminer le temps mort consiste à enregistrer les temps de dérive pour différentes tensions. La droite de corrélation obtenue permettant ensuite la détermination du t_m . La pente de la droite permet de déterminer K_0

La mobilité réduite K_0 d'un ion à travers un gaz tampon est liée à sa section efficace de collision (Ω).

$$K_0 = \frac{3q}{16N_0} \left(\frac{2\Pi}{\mu k_B T} \right)^{1/2} \frac{1}{\Omega}$$

avec, q pour la charge de l'ion, N_0 pour la densité de gaz à P_0 et T_0 , μ la masse réduite du gaz tampon et de l'ion, k_B la constante de Boltzmann, T la température ambiante et Ω la section efficace de collision.

Connaissant K_0 , la section efficace de collision de l'ion étudié peut être calculée par :

$$\Omega = \frac{3q}{16N_0} \left(\frac{2\Pi}{\mu k_B T} \right)^{1/2} \frac{1}{K_0}$$

Une des limites actuelles de la mobilité ionique concerne sa résolution. Celle-ci étant définie comme le temps de dérive de l'ion sur la largeur à mi-hauteur du pic de mobilité.

$$R = \frac{t_D}{\Delta t}$$

Selon le type de montage, la résolution varie considérablement mais reste de trois ordres de grandeur inférieure à celle de la MS. Cela peut poser des problèmes dans l'attribution précise de différences de sections efficaces observées d'ions de m/z proche.

Cette technique couplée à la MS permet à la fois d'identifier les molécules présentes en solution et de déterminer leur conformation (Clemmer et Jarrold, 1997 ; McLean *et al.*, 2005). Il est ainsi possible de mettre en évidence les différents états de conformation d'une protéine ou d'un édifice supramoléculaire non-covalent (Ruotolo et Robinson, 2006). Le couplage IMS-MS est donc particulièrement bien adapté à l'étude des conformations individuelles d'objets présents dans des mélanges complexes. Elle est également capable de mettre en évidence les modifications de conformation d'une protéine induites par la liaison de son ligand (Ruotolo *et al.*, 2005 ; Benesch *et al.*, 2007).

Matériel et méthodes

1. Le matériel biologique

1.1. Les PRP modèles choisies : IB5 et II-1

Dans le cas des protéines salivaires riches en proline, la purification à partir de salive ne permet pas d'obtenir une quantité suffisante de protéines pures pour une étude physicochimique du fait de leur faible concentration dans la salive et du polymorphisme de cette famille de protéine. La production par voie hétérologue permet de produire toutes sortes de protéines. Son schéma global consiste à cloner le gène ou l'ADNc codant pour la protéine sélectionnée dans un plasmide qui permet ensuite son expression dans une cellule hôte.

Pour cette étude, le gène *PRB4S* a été choisi car il code pour une proprotéine salivaire humaine appelée PRB4S qui est clivée en deux protéines : II-1, une PRP glycosylée, et IB5, une PRP basique, (Chan et Bennick, 2001) qui correspondent respectivement aux parties N-terminale et C-terminale de la proprotéine PRB4S (figure 40). L'ADNc correspondant au gène *PRB4S* a été fourni par le Professeur Anders Bennick (Université de Toronto, Canada).

```
Met RFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDL
EGDFDVAVLPFSNSTNGLLFINTTASIAAKEEGVSLEKREAEA
SSEDVSQEESLFLISGKPVGRRPQGGNQPQRPPPPGKPPGPP
PQGGNQSQGGPPPPGKPEGRPPQGRNQSQGPPPHGKPERPP
PQGGNQSQGTPPPPGKPERPPQGGNQSHRPPPPGKPERPP
PQGGNQSRGPPPHRGKPEGPPQEGNKSRSARSPPGKPPGPP
QQEGNKPPGPPPPGKPPGPPAGGNPQQPQAPPAGKPPGPP
PPQGGRPPRPAQGQPPQ
```

EA : deux derniers acides aminés du peptide signal de pPIC9K

AR: deux derniers acides aminés de II-1 (site de clivage de la furine)

N: sites potentiels de glycosylation

Figure 40. Séquence primaire de la proprotéine PRB4S.

La proprotéine PRB4S est clivée dans l'appareil de Golgi par une protéase similaire à la furine (Chan et Bennick, 2001). Cette proprotéine comporte six sites potentiels de N-glycosylation, tous situés en aval du site de clivage de la furine (figure 40). Le clivage de la

proprotéine conduit à la présence de deux protéines dans la salive totale suite à l'expression du gène *PRB4S* (figure 41) :

- II-1, protéine basique (pI = 11,9) glycosylée de 163 a.a. (figure 42.a)
- IB5, protéine basique (pI = 11,2) non glycosylée de 70 a.a. (figure 42.b).



Figure 41. Clivage de la proprotéine PRB4S par les protéases similaires à la furine.

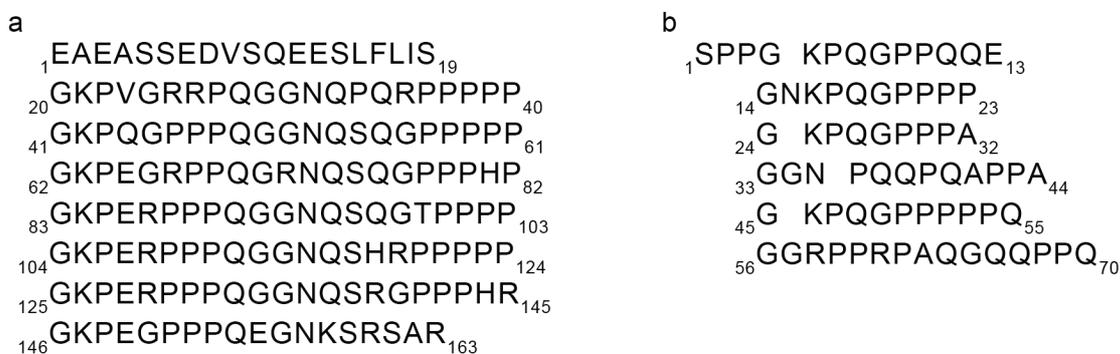


Figure 42. Séquence primaire des protéines a. II-1 et b. IB5 issues du clivage de PRB4S.

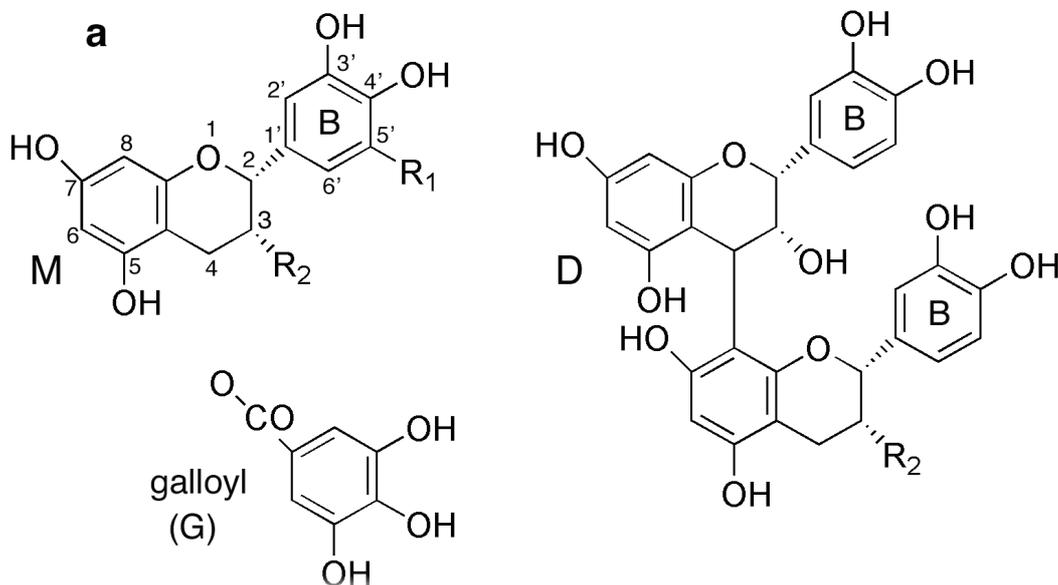
1.2. Les tanins modèles

Au cours de cette étude, cinq tanins modèles ont été mis en interaction avec la protéine IB5 : 3 monomères et 2 dimères (figure 43.a). Les monomères étudiés sont l'épigallocatechine (EgC), l'épicatéchine gallate (ECG) et l'épigallocatechine gallate (EgCG). Les deux dimères sont le B2 ((-)-épicatéchine (4- β -8)-(-)-épicatéchine) et le B2 3'O gallate (B2 3'OG, (-)-épicatéchine (4- β -8)-(-)-épicatéchine-3-O-gallate). Les trois monomères et le dimère B2 ont été fournis par Sigma-Aldrich (Saint Louis, Missouri, Etats-Unis). Le dimère B2 3'OG a été purifié au laboratoire à partir de poudre de pépins de raisin selon la méthode décrite par (Ricardo da Silva *et al.*, 1991b).

1.3. Le « ligand » témoin

Durant l'étape d'ionisation par électrospray, des agrégats non spécifiques sont susceptibles de se former. Afin de s'assurer que les complexes IB5•tanin observés par ESI-MS résultent d'interactions spécifiques, une molécule témoin a été utilisée. La réserpine a

été choisie (figure 43.b) car elle présente une masse moléculaire (MM) proche de celle des tanins étudiés et une structure incluant un noyau phénolique méthoxylé. La réserpine a été fournie par Sigma-Aldrich (Saint Louis, Missouri, Etats-Unis).



Tanins	Squelette	R ₁	R ₂
Epigallocatechine	M	OH	OH
Epicatechine gallate	M	H	galloyl
Epigallocatechine gallate	M	OH	galloyl
B2	D		OH
B2 3'OG	D		galloyl

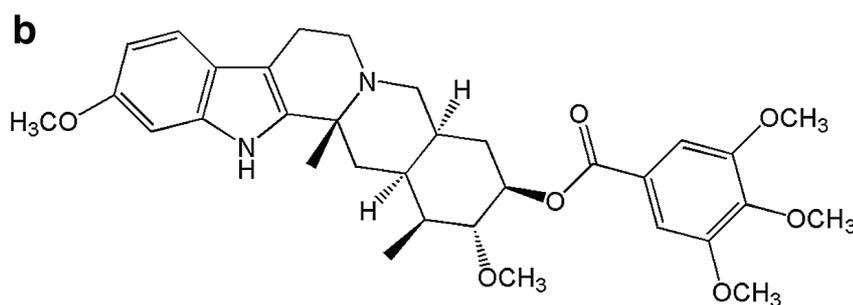
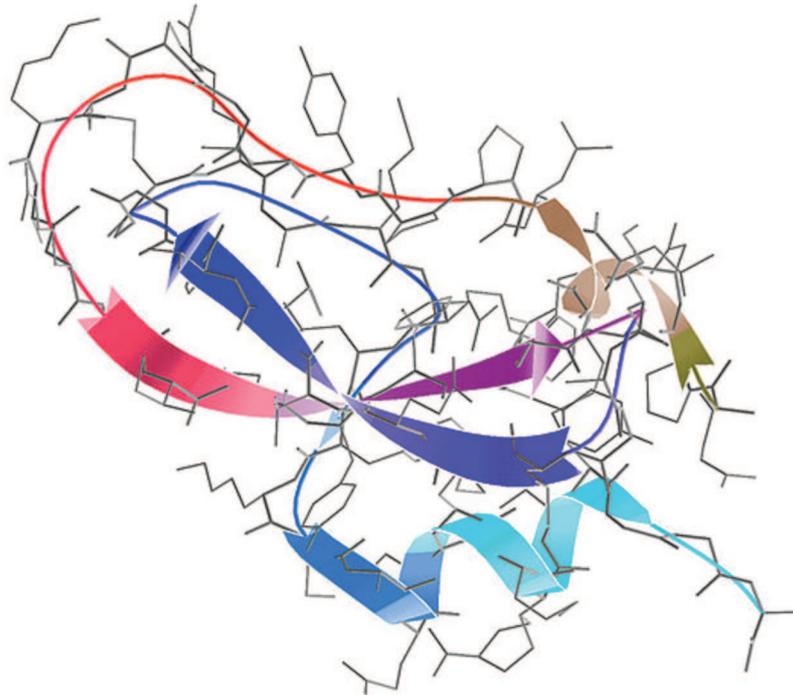


Figure 43. a. Structure des tanins modèles (M : monomères, D : dimères, G : galloyl) ; b. structure de la réserpine.

1.4. La protéine témoin BPTI

Dans le but de comparer les interactions IB5-EgCG avec celles entre une protéine structurée et l'EgCG, nous avons sélectionné une protéine témoin : l'inhibiteur pancréatique

bovin de la trypsine (BPTI, Bovine Pancreatic Trypsin Inhibitor). Elle a été choisie car elle présente une masse atomique et une valeur de pI proches de celles d'IB5 (BPTI MM : 6511,51 Da et pI : 10,5 (Trautschold *et al.*, 1967)). Contrairement à IB5, il s'agit d'une protéine structurée fortement stabilisée par la présence de trois ponts disulfures (Wlodawer *et al.*, 1984 ; Amir et Haas, 1988 ; Dadlez et Kim, 1995) (figure 44).



RPDFCLEPPYTGPCKARIIRYFYNAKAGLC
QTFVYGGCRAKRNNFKSAEDCMRTC GGA

Figure 44. Structure et séquence primaire de BPTI.

BPTI a été fournie par Sigma-Aldrich (Saint Louis, Missouri, Etats-Unis).

2. Production des PRP modèles

La souche de levure *Pichia pastoris* a été retenue comme organisme hôte, car elle présente l'avantage de produire les protéines recombinantes en quantité importante (du mg au g par litre de culture). De plus, en tant qu'organisme eucaryote, *P. pastoris* est capable d'assurer les PTM de type glycosylation (N-oligosaccharides constitués principalement de mannoses (man) α -liés) (Gemmill et Trimble, 1999). Les protéines produites sont excrétées dans le milieu de culture.

La souche recombinante *Pichia pastoris-PRB4S* (Pascal *et al.*, 2006) permet la surexpression des différentes protéines issues de *PRB4S*. Elle a été cultivée en fermenteur dans les conditions décrites par (Pascal *et al.*, 2006) jusqu'à l'obtention d'une quantité suffisante de culture pour réaliser la purification de Il-1 et de IB5.

L'évolution de la population cellulaire de levure a été suivie par mesure de la densité optique à 600 nm. La production des protéines d'intérêt a été suivie par l'analyse du surnageant du fermenteur par électrophorèse en conditions dénaturantes (SDS-PAGE) selon la méthode de Laemmli (Laemmli, 1970) et la coloration au bleu de coomassie décrite par (Beeley *et al.*, 1991a) qui utilise la particularité des PRP de présenter une coloration rose.

Chaque étape de la purification des protéines a également été suivie de cette manière.

SDS-PAGE :

L'analyse des protéines en SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) permet de séparer les protéines en fonction de leur mobilité électrophorétique, qui varie en fonction de leur MM. En présence de SDS, les protéines ont une densité de charge uniforme et conservent une structure dépliée en raison des répulsions de charges. La migration électrophorétique est alors théoriquement uniquement fonction de la MM. Toutefois, les protéines intrinsèquement déstructurées présentent une migration anormale en SDS-PAGE (Hu et Ghabrial, 1995 ; Iakoucheva *et al.*, 2001 ; Receveur-Brechot *et al.*, 2006 ; Hebrard *et al.*, 2009). En effet, la composition en acides aminés atypique des PID est moins favorable à la fixation du SDS. Leur masse apparente sur gel est ainsi 1,2 à 1,8 fois plus élevée que leur masse réelle (Tompa, 2002). Nous distinguerons donc leur masse apparente sur gel de leur MM réelle déduite de leur séquence ou mesurée par MS.

Préparation des gels et migration :

Les échantillons sont mélangés à un volume équivalent de tampon de dépôt (Tris-HCl 125 mM à pH 6,8, SDS 4% (m/v), glycérol 20% (v/v), DTT 10% (v/v), bleu de bromophénol 0,1% (m/v)) et soumis à une électrophorèse sur gel de polyacrylamide à 12,5% en présence de SDS (SDS-PAGE 12,5%) après dénaturation (90°C pendant 5 min.). Le gel de séparation est composé de 12,6% d'acrylamide / bis-acrylamide (29/1) (v/v), Tris-HCl 0,375 M à pH 8,8, SDS 0,1% (v/v), TEMED 0,05% (v/v) et persulfate d'ammonium 0,07% (m/v). Le gel de concentration est composé de 3,5% d'acrylamide / bis-acrylamide (29/1) (v/v), Tris-HCl 0,125 M à pH 6,8, SDS 0,1% (v/v), TEMED 0,1% (v/v) et persulfate d'ammonium 0,05% (m/v). La

migration est effectuée à l'aide du tampon d'électrode (Tris 25mM, glycine 192 mM, SDS 0,1% (m/v)). Les marqueurs de masse moléculaire (MqM) utilisés sont les Low Molecular Weight (LMW) de chez Pharmacia Amersham Biosciences GE Healthcare (Uppsala, Sweden). Leurs MM sont comprises entre 97 000 à 14 400 Da. La migration s'effectue à 40 mA pendant 1 h dans le gel de concentration, et à 100 mA pendant 2 h dans le gel de séparation. Après migration, les protéines sont colorées au bleu de Coomassie R250 0,1% dans une solution d'éthanol 40% et 10% d'acide acétique (3 heures), puis décolorées par une solution d'acide acétique à 10% sans solvant organique afin d'obtenir la coloration rose caractéristique des PRP (Beeley *et al.*, 1996).

Au cours des différentes étapes, les protéines sont quantifiées par densitométrie de leur coloration (TotalLab 120, Amersham Biosciences, GE Healthcare, Uppsala, Sweden).

3. Purification des PRP modèles

Les protéines IB5 et II-1 ont été purifiées par deux protocoles différents, le deuxième résultant d'évolutions expérimentales.

La purification des protéines débute par la centrifugation de la culture qui sépare un précipité de cellules de levure d'un surnageant. Le volume total de surnageant (3 L) est aliquoté (10 lots de 280 mL et 2 lots de 100 mL) et conservé à -20°C .

Les étapes de purification des PRP sont suivies par la mesure de leur absorbance à la longueur d'onde de 230 nm, l'absence de résidu aromatique au sein des PRP n'autorisant qu'un suivi de l'absorbance de la liaison peptidique.

3.1. Protocole de purification n°1

Les premiers lots de surnageant ont été traités suivant les étapes de purification décrites par (Pascal *et al.*, 2006) et résumées dans la figure 45.

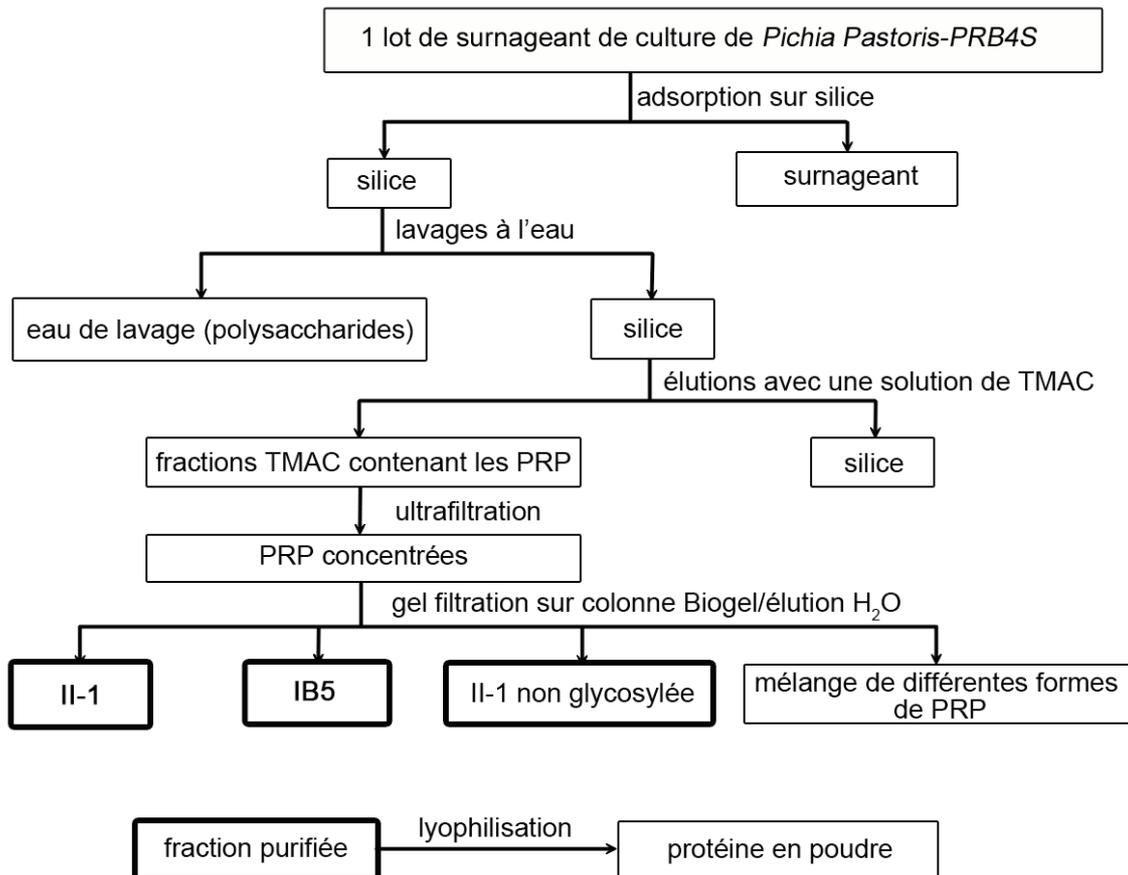


Figure 45. Schéma simplifié des étapes de purification des protéines II-1 et IB5.

La purification se déroule en trois étapes.

1- chromatographie d'adsorption : elle sépare les molécules de l'échantillon en fonction de la différence d'affinité qu'elles ont pour la phase stationnaire. Après dépôt de l'échantillon, le passage d'un solvant de lavage élimine les molécules non fixées à la phase stationnaire, puis celui du solvant de désorption permet de décrocher les molécules fixées à la phase stationnaire. Dans notre cas, cette étape sépare les PRP de l'ensemble des composés résiduels du milieu nutritif (sels) et des produits du métabolisme de *P. pastoris* (protéines propres à la levure, polysaccharides).

2- ultrafiltration : elle permet de concentrer une solution par l'application d'une pression hydrostatique. La pression force l'écoulement du solvant à travers une membrane semi-perméable qui retient les molécules d'intérêt. Elle a permis de concentrer les fractions de PRP obtenues.

3- chromatographie d'exclusion stérique : la séparation des molécules résulte de l'existence de pores dans la phase stationnaire. Les molécules les plus grosses sont exclues

des pores alors que les molécules de plus faible taille pénètrent au sein des billes de la phase stationnaire. Ainsi les molécules de haute masse sont éluées rapidement et celles de plus faible taille ont un trajet allongé au sein de la phase stationnaire. Cette troisième étape a conduit à l'obtention de fractions « pures » d'IB5 et de II-1.

3.2. Protocole de purification n°2

Si l'efficacité du protocole n°1 est réelle, sa mise en œuvre requiert plusieurs mois pour traiter un lot de 280 mL de surnageant de la culture. Dans le but de raccourcir la durée de la purification des PRP mais aussi d'améliorer les rendements de purification, une deuxième procédure a été mise au point et appliquée aux autres lots de surnageant.

Ce deuxième protocole de purification repose sur deux étapes, une première étape de chromatographie d'échange d'ions et une deuxième étape de chromatographie d'exclusion stérique.

3.2.1. Adsorption/désorption des PRP sur une colonne échangeuse d'ions

La chromatographie d'échange d'ions sépare plusieurs composés sur la base de leurs propriétés électriques. La phase d'échange d'ions (phase stationnaire) peut comporter des groupements fonctionnels anioniques (échange de cations) ou cationiques (échange d'anions). Les molécules ont plus ou moins d'affinité pour la phase stationnaire en fonction de leur charge. L'augmentation de la force ionique du solvant d'éluion permet la désorption des molécules fixées sur la phase stationnaire. La chromatographie en lit expansé est une technique de purification des protéines qui permet de combiner clarification, concentration et purification en injectant directement des solutions contenant des particules en suspension (colloïdes, débris cellulaires) dans la colonne par un flux ascendant. Le protocole expérimental de la chromatographie en lit expansé diffère peu de celui de la chromatographie en lit fixe, la principale différence étant la direction du flux de liquide. Cette chromatographie comporte cinq étapes successives : équilibration, application de l'échantillon, lavage, éluion et nettoyage (figure 46).

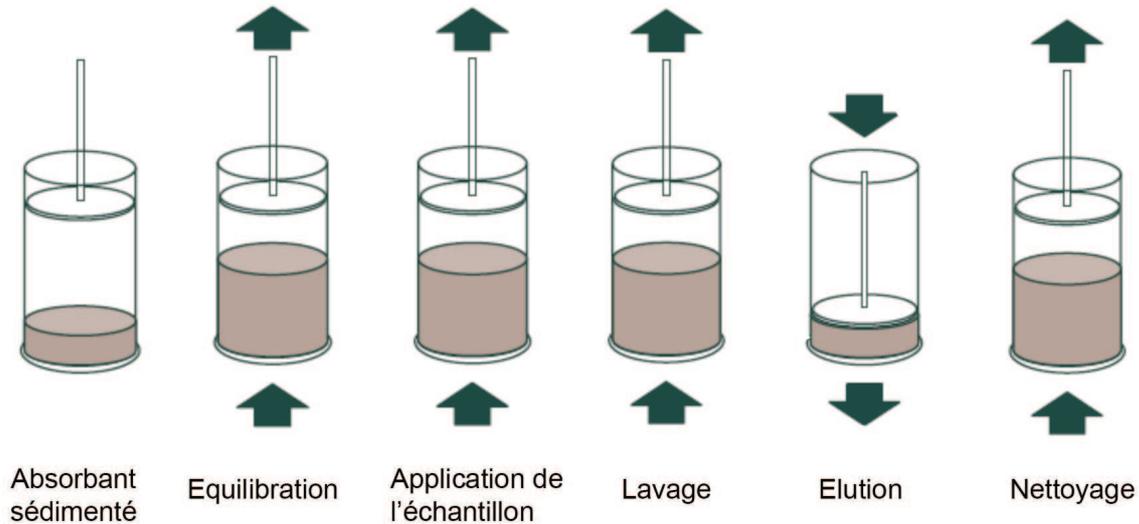


Figure 46. Schéma des étapes de la chromatographie en lit expansé.

Pour cette étape de purification, le système FPLC Biocad sprint (Perseptive Biosystems, Massachusetts, USA) a été utilisé. Le gel sulphopropyl – XL (Amersham Pharmacia Biotech) est équilibré avec un tampon TrisHCl 50 mM, pH 8,0 (débit 13 mL.min⁻¹) au sein d'une colonne Pharmacia Streamline 50 (id 2 cm/ L 40 cm). Après équilibration, le surnageant de culture dilué dans le même tampon est injecté dans le lit stabilisé. Le système fonctionne en circuit fermé pendant une durée proportionnelle au volume de surnageant traité (75 min pour 280 mL), ceci afin d'optimiser le rendement de liaison entre les protéines d'intérêt et les particules adsorbantes. Un lavage, également effectué en flux ascendant, permet ensuite d'éliminer les espèces résiduelles faiblement liées. Pour l'étape d'éluion, l'adsorbant est compacté et la direction du flux de solvant est inversée avec un débit de 5 mL.min⁻¹. Le solvant utilisé est un tampon TrisHCl 50 mM, pH 8,0 avec un gradient de NaCl de 0 à 1 M pendant 7 min puis un palier de 1 M de NaCl de 15 min. Ainsi, les PRP sont éluées en flux descendant de la résine sédimentée. L'éluion est suivie en temps réel par une mesure en parallèle à 230 et 280 nm permettant d'identifier les fractions contenant les PRP. Cette étape est répétée si toutes les PRP n'ont pas été adsorbées après contrôle en SDS-PAGE. Le lit est régénéré par un premier lavage, en utilisant des tampons spécifiques. Ce nettoyage est réalisé directement dans la colonne en mode expansé ou sédimenté.

3.2.2. Séparation des PRP par gel filtration

Les fractions de PRP récupérées à l'issue de l'étape d'échange d'ions sont ensuite séparées par chromatographie d'exclusion stérique pilotée par le système FPLC Biocad sprint (Perseptive Biosystems, Massachusetts, USA). La fraction PRP est injectée sur une colonne (Ø : 1,6 cm / L : 60 cm) HiLoad Superdex 75 PC 16/60 (seuil de coupure : 3 à 70

kDa ; Amersham Biosciences, GE Healthcare, Uppsala) équilibrée par un tampon d'acétate d'ammonium 50 mM pH 5,5. L'élution des protéines est assurée par le passage du même tampon à un débit de 0,8 mL.min⁻¹. Des fractions de 1 mL sont récupérées et analysées sur gel de polyacrylamide en conditions dénaturantes. Les fractions contenant respectivement les protéines IB5 et II-1 isolées sont assemblées. Elles sont ensuite systématiquement lyophilisées afin, d'une part, de les concentrer et, d'autre part, de les conserver. Les lyophilisats obtenus sont pratiquement exempts de sel, l'acétate d'ammonium étant un tampon volatil.

Les protéines ainsi obtenues sont conservées à -20°C.

3.3. Caractérisation des PRP

3.3.1. Caractérisation des PRP par SDS-PAGE

Les fractions de protéines ont été analysées par SDS-PAGE en conditions dénaturantes (SDS 0,1%) avec des gels à 12,5% d'acrylamide.

3.3.2. Caractérisation de la séquence N-terminale des PRP

Le microséquençage N-terminal des protéines purifiées a été réalisé selon la méthode d'Edman par la Plate-forme de Protéomique Analytique et Fonctionnelle du laboratoire d'analyses séparatives et de microséquençage du Centre de Recherches INRA de Tours, Nouzilly, France.

3.3.3. Caractérisation des PRP par ESI-MS

Les PRP purifiées ont été analysées par ESI-MS en infusion directe. Ces analyses ont été effectuées en mode positif à l'aide d'une trappe ionique (LCQ, Thermo Finnigan, Bremen, Germany) et d'un TOF (AccuTOF, JEOL, Tokyo, Japan). Les spectres de masse ont été acquis sur des gammes de masse comprises entre 300 et 2000 *m/z*. Les appareils ont été calibrés via une calibration externe.

Les spectres ont été traités à l'aide des logiciels Excalibur (Thermo Finnigan, Bremen, Allemagne), pour les données provenant de la trappe ionique, et Mass Center (JEOL, Tokyo, Japon), pour celles issues du TOF. Les données converties à partir de ces logiciels en format ASCII ont été déconvoluées à l'aide du logiciel MagTran (Zhongqi Zhang, Amgen, Thousand Oaks, CA). Cette déconvolution permet de convertir un spectre

présentant l'intensité en fonction du rapport m/z en spectre présentant l'intensité en fonction de la masse. La masse des différentes espèces est alors directement déterminée sur le spectre déconvolué.

3.4. Déglycosylation de II-1

Les protéines enzymatiques sont un outil inestimable de par la spécificité des réactions qu'elles catalysent. Elles offrent également des avantages évidents de reproductibilité et de « contrôle » par rapport aux méthodes chimiques. L'utilisation de N-glycannases (O'Neill, 1996), autorisant le clivage spécifique de structures oligosaccharidiques N-liées à des résidus d'asparagine, a véritablement révolutionné l'analyse des N-glycoprotéines. Ainsi, nous avons envisagé l'utilisation de telles enzymes dans la caractérisation de la protéine II-1 qui comporte 6 sites potentiels de N-glycosylation.

Les fractions purifiées de II-1 ont été soumises, d'une part, à une déglycosylation à l'aide de l'endo- β -N-acetylglucosaminidase H (EndoH) et, d'autre part, à une déglycosylation à l'aide de la peptidyl N-glycosidase F (PNGaseF). Selon l'enzyme utilisée, EndoH ou PNGaseF, la protéine conserve respectivement un ou aucun résidu GlcNAc du noyau pentasaccharidique de base lié au niveau du résidu Asn (N) (figure 47).

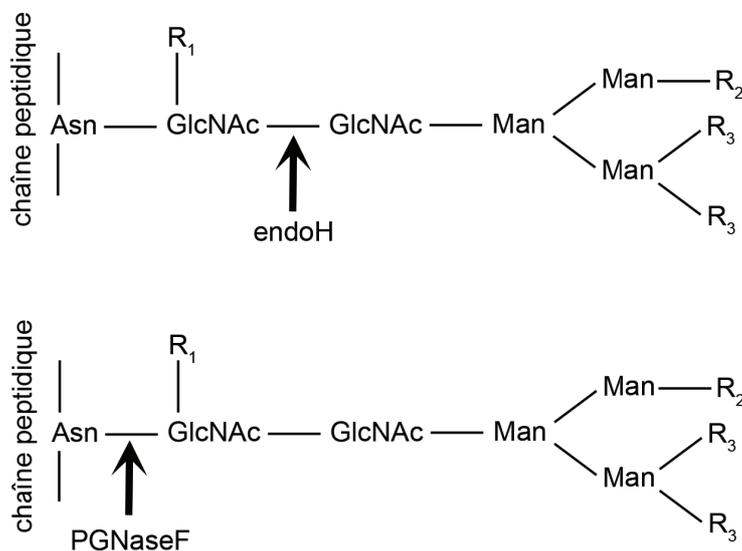


Figure 47. Représentation schématique de l'action des N-glycannases utilisées.

Les protéines glycosylées purifiées (500 μ g) ont été digérées sans dénaturation préalable :

- par 0,1 unité d'endoH (Roche, Bâle, Suisse)

ou

- par 10 unités de PNGaseF (Roche, Bâle, Suisse)

Les réactions ont été réalisées dans de l'eau à 37 °C pendant 17 h.

L'incidence de la déglycosylation sur la masse moléculaire apparente a été vérifiée par électrophorèse en présence de SDS et coloration au bleu de coomassie.

Les produits issus des réactions ont été analysés par ESI-MS et MALDI-MS.

4. Solutions d'interaction PRP - tanins

4.1. Les milieux utilisés

L'eau qui a été utilisée est de l'eau MilliQ filtrée et désionisée avec un système de filtration millipore (Millipore, Milford, USA).

L'éthanol (HPLC grade), l'acide acétique (HPLC grade), le TFA et l'acétate d'ammonium ont été fournis par Merck (Whitehouse Station, New Jersey, Etats-Unis).

L'ensemble des expériences a été réalisé dans les milieux suivants :

- eau pure,
- eau/éthanol (EtOH) 88/12 (v/v) acidifiée avec de l'acide acétique (valeurs de pH allant de 2,9 à 5,7),
- eau/éthanol 88/12 contenant 20 mM d'acétate d'ammonium, pH 6,8.

Les protéines ont également été analysées dans un solvant eau/acétonitrile (ACN) (v/v ; 50/50) TFA 0,01% qui est le milieu classique d'analyse de ces bio-macromolécules en ESI-MS (Mouls *et al.*, 2007).

4.2. Les solutions mères

Des solutions mères de protéines et de tanins ont été préparées. Elles ont été immédiatement congelées et conservées à -20°C afin d'éviter les phénomènes de protéolyse des protéines ou d'oxydation des tanins.

4.3. Les solutions d'interaction

Les solutions d'interactions ont été préparées à partir des solutions mères et de solvant. Les mélanges d'interaction ont été systématiquement réalisés dans l'ordre suivant :

solvant, protéine puis tanin. Seuls les mélanges relatifs aux expériences de compétition ont fait exception.

Des expériences de compétition entre deux tanins pour la protéine ont été effectuées. Il existe deux types d'expériences de compétition en MS (figure 48) :

- le déplacement d'équilibre par compétition de ligand en solution,
- la compétition directe en solution.

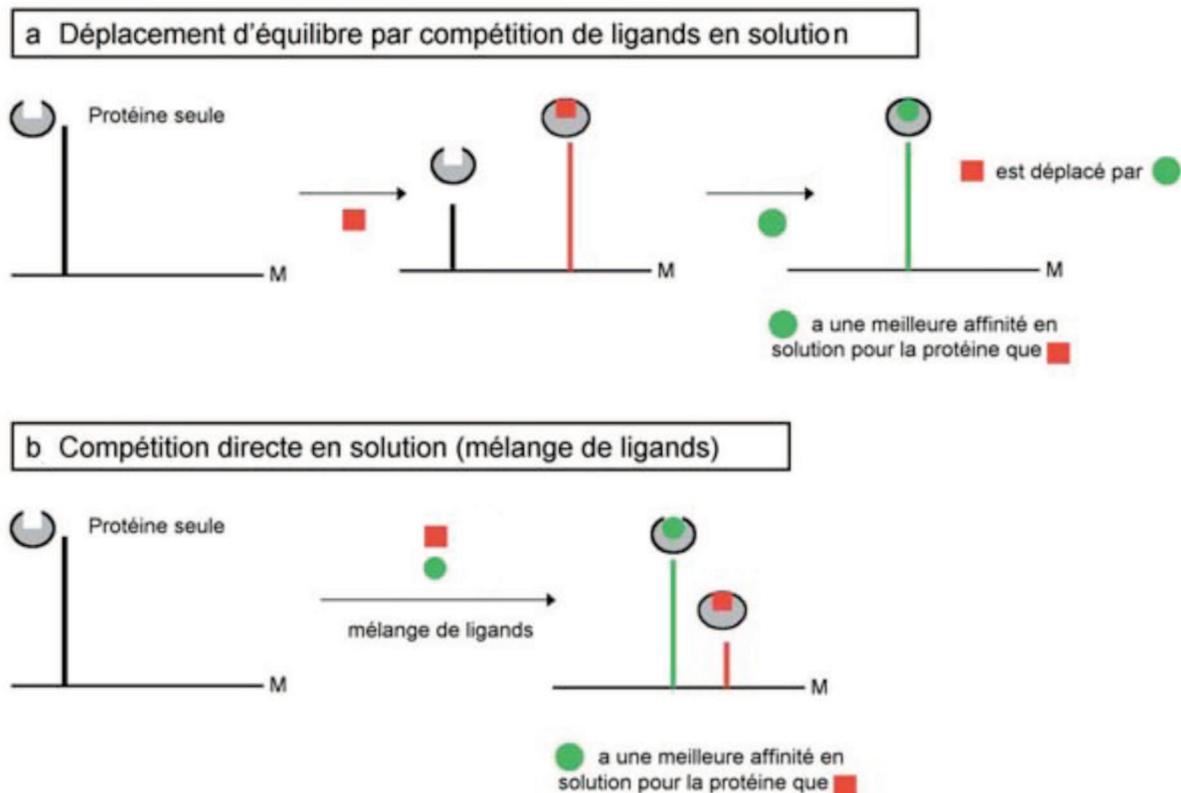


Figure 48. Détails des deux expériences de compétition.

L'ordre d'introduction des molécules est un paramètre important suivant le type d'expérience que l'on souhaite réaliser. Au cours de cette étude, des expériences de compétition directe en solution ont été réalisées : les compétiteurs (les tanins) ont été introduits dans la solution avant le partenaire d'interaction (IB5), afin que les deux ligands soient en présence d'IB5 en même temps.

5. Etude par spectrométrie de masse

5.1. Spectromètres de masse utilisés avec source ESI

Au cours de cette étude, différents spectromètres de masse avec une source de type électrospray ont été utilisés :

- une trappe ionique 3D (LCQ, Thermo Finnigan, Bremen, Allemagne)
- un temps de vol (TOF) (AccuTOF, JEOL, Tokyo, Japon)
- un appareil hybride de type Q-TOF, (Qstar Pulsar i mass spectrometer, Applied Biosystems, Forster City, CA)
- une trappe ionique linéaire, (LTQ, Thermo Fisher Scientific, Bremen, Allemagne)
- une cellule ICR à transformée de Fourier, (FT-ICR, Bruker APEX III, Bremen, Allemagne)
- Un appareil hybride constitué du couplage d'un tube de mobilité ionique, développé au sein du laboratoire de spectrométrie ionique et moléculaire (LASIM) par l'équipe de Philippe Dugourd, et d'un spectromètre de masse de type Q-TOF (microQ-TOF, Bruker, Bremen, Germany).

5.2. Spectromètres de masse utilisés avec source MALDI

Les spectres réalisés avec une source MALDI ont été acquis sur deux spectromètres de masse différents.

- MALDI-TOF (BiFlex III, Bruker Daltonik GmbH, Bremen, Allemagne)
- MALDI-TOF/TOF (Ultraflex mass spectrometer, Bruker Daltonik GmbH, Bremen, Allemagne)

5.3. Préparation des échantillons pour les analyses MALDI-MS

Le choix de la matrice utilisée a été réalisé à partir d'essais préliminaires. Il s'est porté sur l'acide dihydroxybenzoïque (DHB).

La solution de matrice à une concentration de 10 mg.mL⁻¹ dans une solution aqueuse d'acétonitrile à 50% acidifiée avec 0,1% d'acide trifluoroacétique (TFA) a été mélangée à

l'analyte selon la méthode « dry-droplet » (50%, v/v). Après une brève agitation, 1 μ L de la solution matrice-protéine a été déposé sur la cible et séché à l'air ambiant.

Une calibration de masse externe a été réalisée avec la BSA pour les spectres acquis sur une gamme de masses comprises entre 10 000 et 80 000 et avec un mélange de peptides de masses connues pour les gammes de masses inférieures à 10 000.

Chapitre I. Caractérisation des protéines II-1, II-1NG et IB5

Afin de disposer d'une quantité suffisante de protéines pures pour développer l'ensemble de ce travail (et plus largement du travail mené avec les partenaires de l'ANR PROTANIN), nous avons réalisé plusieurs cultures de la souche *Pichia pastoris-PRB4S*. Nous avons également amélioré et répété les purifications. Ces étapes ont donné lieu à la constitution d'une « réserve » d'IB5, de II-1 et de II-1NG (forme de II-1 ne portant pas de glycosylation mais produite par la souche *Pichia pastoris-PRB4S*) constituée de divers lots de purification qui ont tous été caractérisés. Les caractéristiques des principaux lots utilisés dans le cadre de ce travail sont présentées ici.

Un schéma général de caractérisation a été suivi pour tous les lots, y compris pour ceux comprenant les protéines doublement marquées ^{13}C et ^{15}N destinées aux études RMN. La confirmation de l'identité des protéines purifiées passe par la caractérisation la plus fine possible des chaînes peptidiques mais aussi par celle de leurs modifications post-traductionnelles (PTM). En effet, ces dernières sont des éléments essentiels de la modulation de l'activité biologique des protéines (Kukuruzinska et Lennon-Hopkins, 1999). La glycosylation est considérée comme la PTM la plus importante dans les systèmes eucaryotes. Elle intervient en particulier dans la structuration (« *folding* ») des protéines (Helenius et Aebi, 2001).

Parmi les différentes techniques de caractérisation des protéines, la spectrométrie de masse est rapidement devenue une technique de choix, suite à l'apparition des sources d'ionisation douce de type MALDI et ESI. La MS permet d'accéder rapidement à des informations précises sur leur masse moléculaire, leur séquence primaire et leur identité. Elle permet également d'obtenir des informations concernant les éventuelles PTM, telles que leur pattern de glycolysation.

L'objectif du travail décrit dans ce chapitre est d'identifier précisément les séquences des protéines co-produites par *Pichia pastoris-PRB4S* et purifiées. Nous nous sommes également attachés à déterminer le nombre et la position des glycosylations portées par la protéine II-1. Ce chapitre est divisé en trois parties, la première fait le bilan de la production des protéines, la deuxième concerne la caractérisation de la protéine II-1 et sa forme non glycosylée (II-1 NG) et la troisième traite de la caractérisation de la protéine IB5.

I.1. Production et purification des PRP

La souche *Pichia pastoris-PRB4S* a été cultivée en fermenteur pendant 146,5 heures selon la méthode décrite par Pascal *et al.* (Pascal *et al.*, 2006).

La culture est arrêtée lorsque son volume atteint le volume utile du fermenteur. Un volume final de culture de 4,2 L a permis d'obtenir 3L de surnageant de culture après centrifugation.

L'analyse de ce surnageant par gel d'électrophorèse en conditions dénaturantes montre la présence de cinq polypeptides majoritaires ayant des MM apparentes sur gel de 40, 20, 19, 17 et 14 kDa (Figure 49, piste 2). La comparaison avec les données de (Pascal *et al.*, 2006) a permis d'identifier ces protéines comme étant respectivement les protéines II-1, II-1 non-glycosylée, une protéine « fragment a » de II-1 non-glycosylée (II-1_{fa-ng}) une deuxième protéine « fragment b » de II-1 (II-1_{fb-ng}) et IB5. Elle a également révélé l'absence de la bande correspondant à la proprotéine PRB4S présente dans l'étude de Pascal *et al.* . Cette observation indique que PRB4S a été complètement maturée alors que sa maturation était partielle dans le fermenteur décrit par (Pascal *et al.*, 2006). Le volume total de surnageant (3 L) a été partagé en 10 lots de 280 mL et 2 lots de 100 mL, conservés à – 20°C.

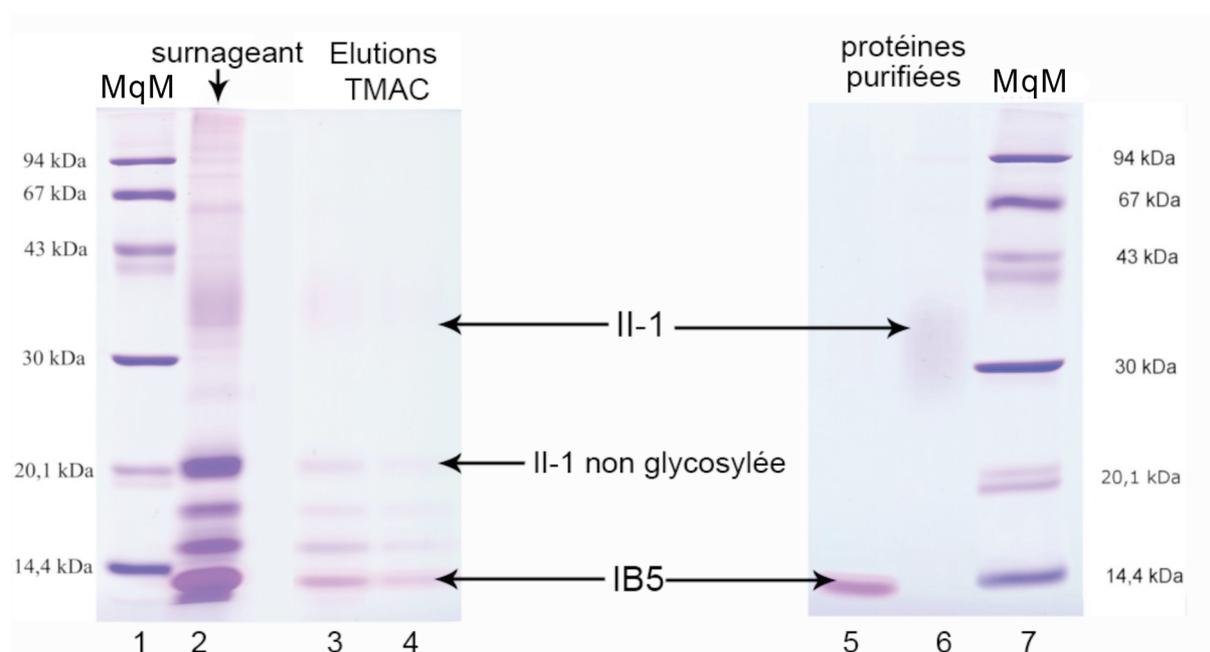


Figure 49. Profils SDS-PAGE (12,5% acrylamide, 0,1% SDS) des différentes étapes de la production et de la purification des PRP excrétées par la souche recombinante *Pichia pastoris-PRB4S* (coloration au bleu de Coomassie. MqM : marqueurs de masse moléculaire en kDa).

Le protocole de purification n°1 a été appliqué à 280 mL de surnageant. La première étape de purification (silice-TMAC) a permis de séparer les PRP des autres molécules du surnageant (figure 49, pistes 3 et 4 notées éluions TMAC). Ces fractions ont ensuite été concentrées par ultrafiltration et les protéines IB5 et II-1 ont été isolées au cours de la troisième étape de chromatographie d'exclusion.

IB5 présente une mobilité électrophorétique légèrement supérieure à celle du marqueur de 14,4 kDa, avec une masse apparente aux alentours de 14 kDa (figure 49, piste 5). II-1 présente une mobilité électrophorétique plus diffuse, avec une large bande s'étendant de 35 kDa à 40 kDa (figure 49, piste 6). Cette migration diffuse s'explique par la présence de glycosylations hétérogènes. La glycosylation s'effectue en des positions particulières de la séquence, en nombre et en structure très variables d'une protéine à l'autre. Ainsi, une même protéine peut exister sous plusieurs isoformes ("glycoformes") (Rademacher *et al.*, 1988) qui diffèrent seulement par leur composition en glycanes : on parle de macro- et de microhétérogénéité de la protéine (Cunningham *et al.*, 1963).

Ce protocole a permis d'obtenir après lyophilisation 31,5 mg de la protéine IB5 et 7 mg de la protéine II-1.

L'analyse par séquençage N-terminal d'Edman des protéines purifiées a révélé que la protéine IB5 est un mélange de différentes formes protéiques (tableau 6). Parmi ces formes, la forme attendue d'IB5 dont l'extrémité N-terminale est « SPPXK » est présente. Le résultat obtenu pour la protéine II-1 ne permet pas une identification précise de l'extrémité N-terminale des chaînes polypeptidiques présentes.

Tableau 6. Résultats du séquençage N-terminal des protéines IB5 et II-1. Les acides aminés en gras sont majoritaires. La lettre X indique que le signal était trop faible ou trop diffus pour permettre l'identification.

	IB5	II-1
Acide aminé n°1	S et Q	X
Acide aminé n°2	P et K	X
Acide aminé n°3	P, K et G	P
Acide aminé n°4	X	P
Acide aminé n°5	G, P et K	X

La première étape du protocole n°2, qui utilise la chromatographie d'échange d'ions, a séparé les PRP des autres molécules présentes dans le surnageant de la culture (figure 50).

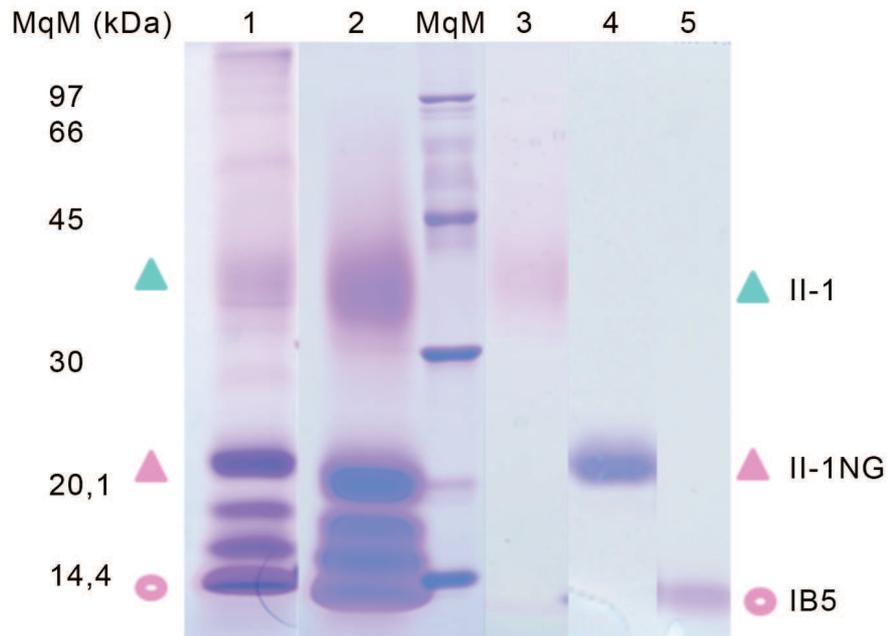


Figure 50. Espèces protéiques sécrétées par *P.pastoris-PRB4S*. Les protéines ont été séparées sur un gel SDS-PAGE à 12,5% d'acrylamide et colorées au bleu de Coomassie. Piste 1 : surnageant brut ; piste 2 : fraction de PRP après adsorption/désorption sur une colonne échangeuse d'ion ; les pistes 3 à 5 correspondent aux fractions purifiées par chromatographie d'exclusion sur une colonne HiLoad Superdex 75 HR ; piste 3 : II-1 lyophilisée ; piste 4 : II-1 non-glycosylée lyophilisée ; piste 5 : IB5 lyophilisée ; piste MqM : marqueurs moléculaires de faible masse moléculaire (LMW, Pharmacia).

La figure 51 montre le profil d'élution de la colonne HiLoad Superdex 75 qui a permis de séparer les PRP les unes des autres. Les fractions ont été repérées sur la base de leur migration apparente sur SDS-PAGE, regroupées et lyophilisées (figure 50).

La caractérisation ESI-MS des protéines pures sera décrite dans les chapitres suivants.

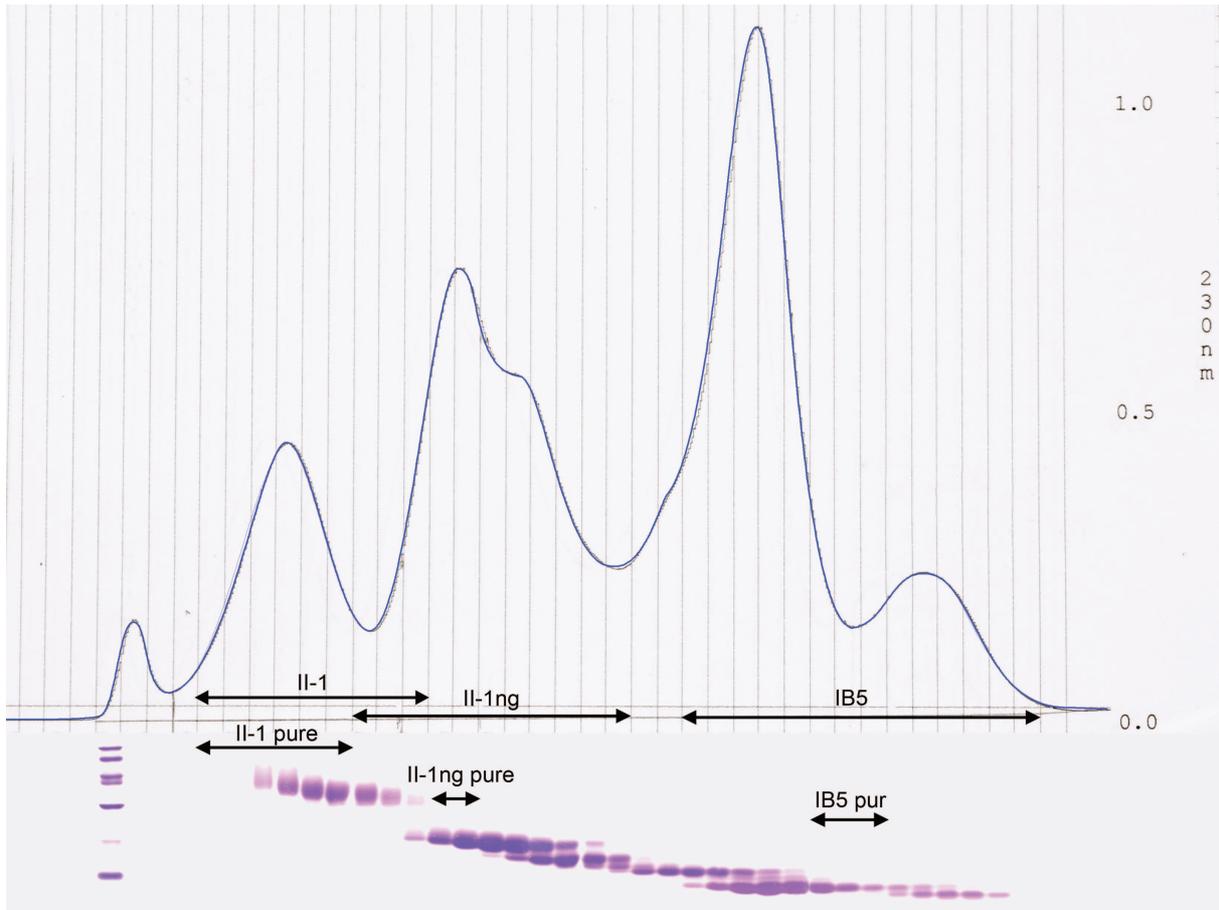


Figure 51. Profil d'élution de la colonne HiLoad Superdex 75.

I.2. Caractérisation des protéines II-1 et II-1NG par ESI-MS et MALDI-MS

De très nombreuses protéines sont des glycoprotéines et la variété des chaînes glucidiques (glycanes) liées à la chaîne polypeptidique accroît leur diversité. La plupart des glycoprotéines sont des protéines sécrétées à l'extérieur des cellules, ce qui est le cas de la PRP salivaire II-1 dans son contexte physiologique. La glycosylation des protéines se décline en deux catégories selon la nature de la liaison chimique entre les glycanes et la protéine. On parle de N-glycosylation lorsque les sucres sont attachés à l'azote de la fonction amide de la chaîne latérale d'une asparagine (Asn ou N), et de O-glycosylation lorsque les sucres sont reliés à l'atome d'oxygène de la chaîne latérale d'une sérine (Ser ou S) ou d'une thréonine (Thr ou T) (figure 52).

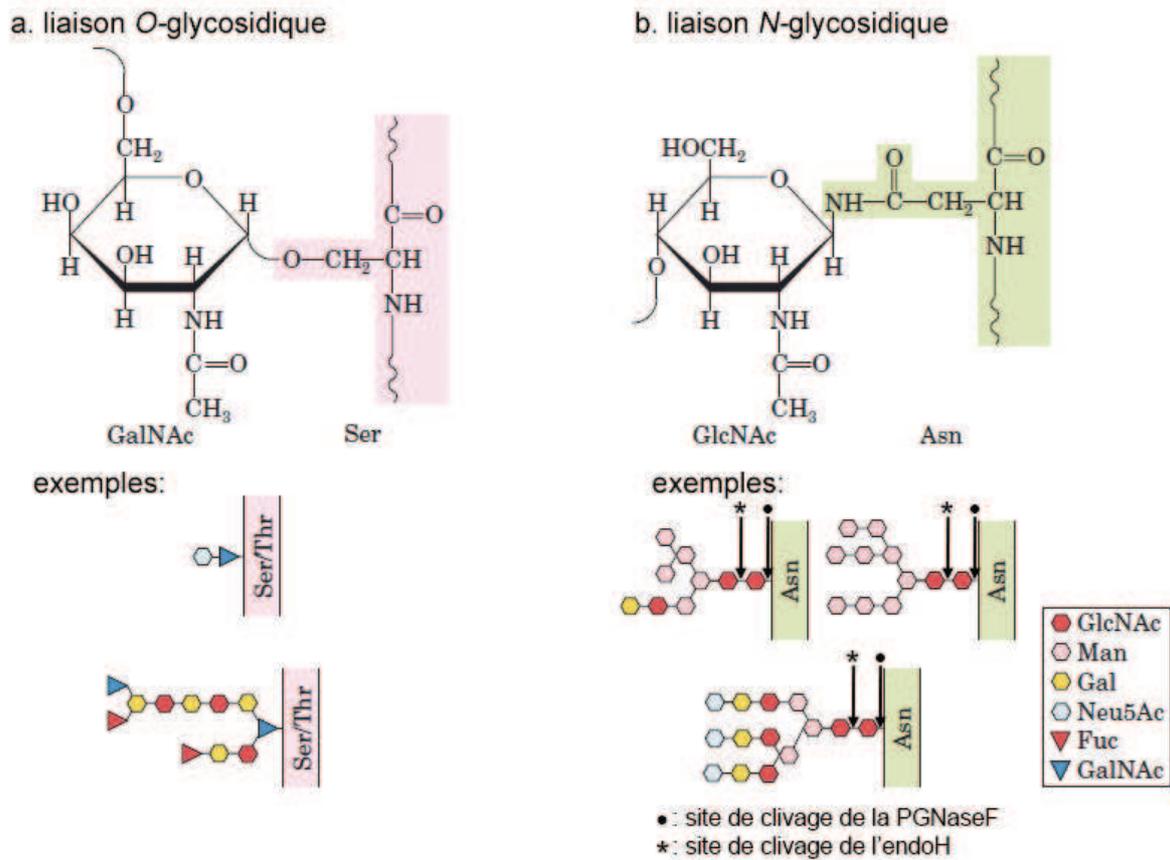


Figure 52. Les deux types de glycosylation des protéines : a. liaison de type O-glycosidique ; b. liaison de type N-glycosidique.

La biosynthèse des glycanes met en jeu des glycosyltransférases et des nucléotides-oses. Il existe sur la chaîne peptidique des signaux que reconnaît la glycosyltransférase fixant la N-acétylglucosamine sur la protéine. Ainsi, dans le cas d'une liaison « asparaginyln-acétylglucosamine », l'ose n'est fixé sur l'Asn qu'en présence de la séquence consensus Asn-X-Ser/Thr. La glycosylation est plus facilement réalisée si la séquence est Asn-X-Thr que Asn-X-Ser. La nature du résidu X de la séquence consensus Asn-X-Ser/Thr est également importante. La glycosylation est inhibée si X est une proline (Shelikoff *et al.*, 1996) et dans le cas où X est un résidu hydrophobe, comme le tryptophane (Yan *et al.*, 1999) ou la phénylalanine (Shakin-Eshleman *et al.*, 1996). Les acides aminés chargés négativement tels que glutamate et l'aspartate, sont moins favorables à l'occupation du site potentiel de N-glycosylation que les acides aminés chargés positivement (lysine, histidine et arginine) (Yan *et al.*, 1999). L'hyperglycosylation est favorisée lorsque deux séquences de sites potentiels sont en tandem dans la séquence primaire de la protéine. Il a ainsi été suggéré par différents auteurs que la séquence Asn-Asn-Thr-Thr est celle de l'hyperglycosylation chez *P. pastoris* (Bretthauer et Castellino, 1999). Les sites de N-

glycosylation d'une glycoprotéine restent des sites potentiels qui peuvent être occupés ou non.

La protéine II-1 présente six séquences consensus de type Asn-X-Ser, pour lesquelles X n'est jamais ni un résidu proline, ni phénylalanine, ni tryptophane. Elle est donc susceptible de porter six *N*-glycanes.

Les méthodes de déglycosylation enzymatique/chimique et de marquage de site sont une aide indispensable pour la localisation du point d'attache peptidique d'une *N*-glycosylation. Deux types de *N*-glycannases dont l'action a été détaillée dans la figure 48 ont été utilisées en parallèle. La PNGaseF permet de caractériser la chaîne peptidique qui est alors totalement déglycosylée. L'endoH permet de déterminer le nombre de glycanes liés en mesurant la différence entre la masse des polypeptides totalement déglycosylés générés par la PNGaseF et celle des polypeptides issus de l'action de l'endoH qui ont conservé un GlcNAc par glycanes liés. La présence d'un GlcNAc sur un peptide confère à ce dernier un incrément de masse de 203 u.

L'utilisation d'endopeptidases spécifiques permet d'établir la carte peptidique de la protéine. La comparaison de la carte peptidique massique de la protéine déglycosylée par la PNGaseF avec celle de la protéine déglycosylée par l'endoH permet de localiser les sites de glycosylation. Le choix de l'endopeptidase se fait à partir du nombre de clivages théoriques qu'elle est susceptible d'effectuer au niveau de la séquence primaire. La localisation des nombreux résidus proline présents au niveau de la séquence de II-1 exclut l'utilisation de nombreuses endopeptidases spécifiques, ce qui limite le nombre d'informations accessibles par cette méthode. Toutefois, l'endoprotéinase GluC qui clive de façon spécifique les liaisons peptidiques situées en position C-terminale (C-term) des résidus Glu (E) est théoriquement capable de fournir le plus d'informations. Le tableau 7 présente les peptides théoriques obtenus après action de la GluC. Les fragments peptidiques que la GluC génère comportent un seul site de *N*-glycosylation permettant la distinction et la détermination des sites potentiels de glycosylation portant un glycanes ou pas.

Tableau 7. Séquences théoriques obtenues après clivage de la protéine II-1 par l'endoprotéinase GluC et leurs masses moléculaires correspondantes.

Séquences théoriques de II-1 obtenues après action de la GluC	Position	Masse (Da)
GPPPQE	150 - 155	624,67
GNKSRSAR	156 - 163	875 ,96
RPPPQGGNQSQGTPPPPGKPE	87 - 107	2124,33
GRPPQGRNQSQGPPPHPGKPE	66 - 86	2219,44
RPPPQGGNQSHRPPPPGKPE	108 - 128	2228,49
RPPPQGGNQSRGPPPHRGKPE	129 - 149	2247,49
SLFLISGKPVGRRPQGGNQPRPPPPGKPPQGGNQSQGGPPPPGKPE	14 - 65	5292,02

La caractérisation de la protéine II-1 a pour objectifs :

- 1- la détermination de la masse moléculaire globale de II-1 ;
- 2- la caractérisation de la chaîne peptidique de II-1 ;
- 3- la détermination du nombre de glycosylations ;
- 4- l'identification des sites glycosylés.

La stratégie de caractérisation retenue est basée sur des analyses MALDI-MS et ESI-MS réalisées sur des échantillons de protéines pures initiales ou préalablement soumises à différents traitements enzymatiques. Cette approche est résumée sur la figure 53 et les informations recherchées sont décrites dans le tableau 8.

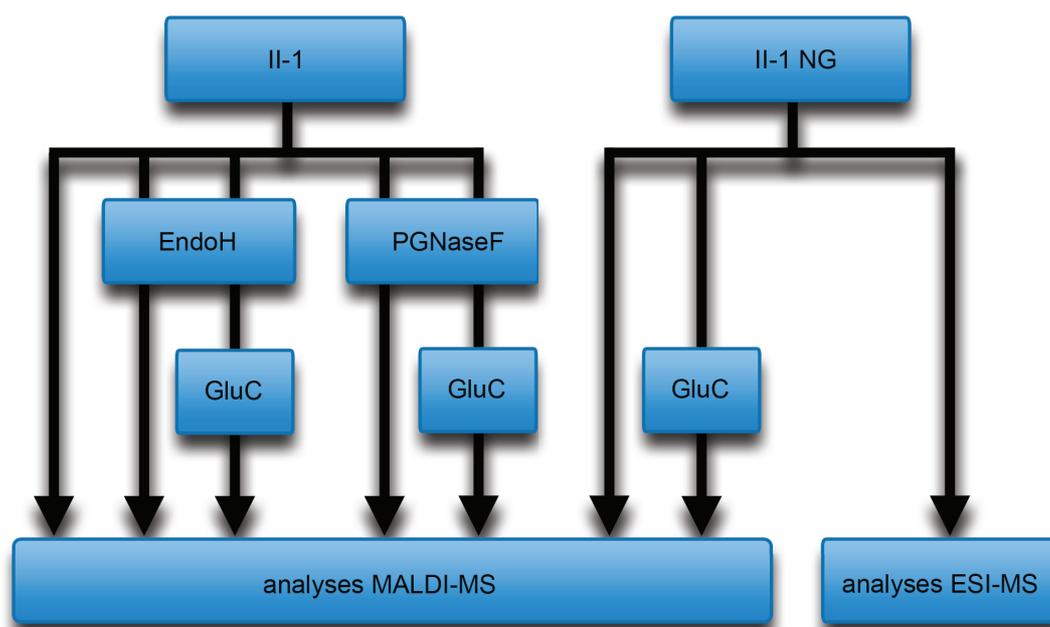


Figure 53. Approche développée pour l'étude de II1.

Tableau 8. Informations attendues à partir des différentes digestions enzymatiques.

Enzyme	Activité	Informations recherchées
PNGaseF	Coupe la liaison N-glycopeptidique entre l'Asn et le premier GlcNAc de la chaîne glycosidique	Caractérisation de la chaîne peptidique
EndoH	Coupe la liaison entre le premier et le deuxième GlcNAc de la chaîne glycosidique	Détermination du nombre de N-glycanes liés
GluC	Coupe la liaison peptidique en position C-term des résidus Glu (E)	Détermination de la position des chaînes glycosidiques

I.2.1. Caractérisation des chaînes peptidiques des protéines II-1NG et II-1

a. Caractérisation de la chaîne peptidique de II-1NG

La caractérisation de la chaîne peptidique de la protéine II-1 a été effectuée en parallèle à celle de la chaîne peptidique de II-1NG. Ces caractérisations ont été réalisées par ESI-MS (figure 54.a) et MALDI-MS (figure 54.c). Pour décrire les différentes chaînes peptidiques identifiées, un code composé de deux lettres est utilisé (ex : B-x). La première lettre correspond à la position du début de la séquence et la seconde à la position de la fin de la séquence sur la figure 55.

L'analyse du spectre ESI-MS de II-1NG révèle que neuf chaînes polypeptidiques sont co-purifiées et qu'elles correspondent toutes à II-1NG (figures 54.a et 54.b). Les masses moléculaires de chacune de ces chaînes polypeptidiques sont données dans le tableau 9 avec l'identification de la séquence correspondante.

Le spectre MALDI-TOF de II-1NG est présenté sur la figure 54.c. En accord avec le spectre ESI-MS, son analyse confirme la présence de plusieurs chaînes polypeptidiques co-purifiées. Les six chaînes identifiées par cette analyse correspondent à celles déterminées par ESI-MS. Toutefois, la résolution du spectre MALDI-TOF ne permet pas de mettre en évidence la présence des trois autres formes de II-1NG révélées par ESI-MS. En effet, la majorité des ions obtenus par MALDI-TOF est sous forme monochargée et l'analyse des ions ayant une masse moléculaire supérieure à 5000 u requiert l'utilisation du mode linéaire de ces analyseurs qui a une résolution inférieure au mode avec réflectron, comme indiqué dans l'introduction. A l'inverse, l'étude ESI-TOF de II-1NG, formant des ions multichargés dont les rapports m/z sont inférieurs à 3000 Da, permet l'utilisation du mode réflectron.

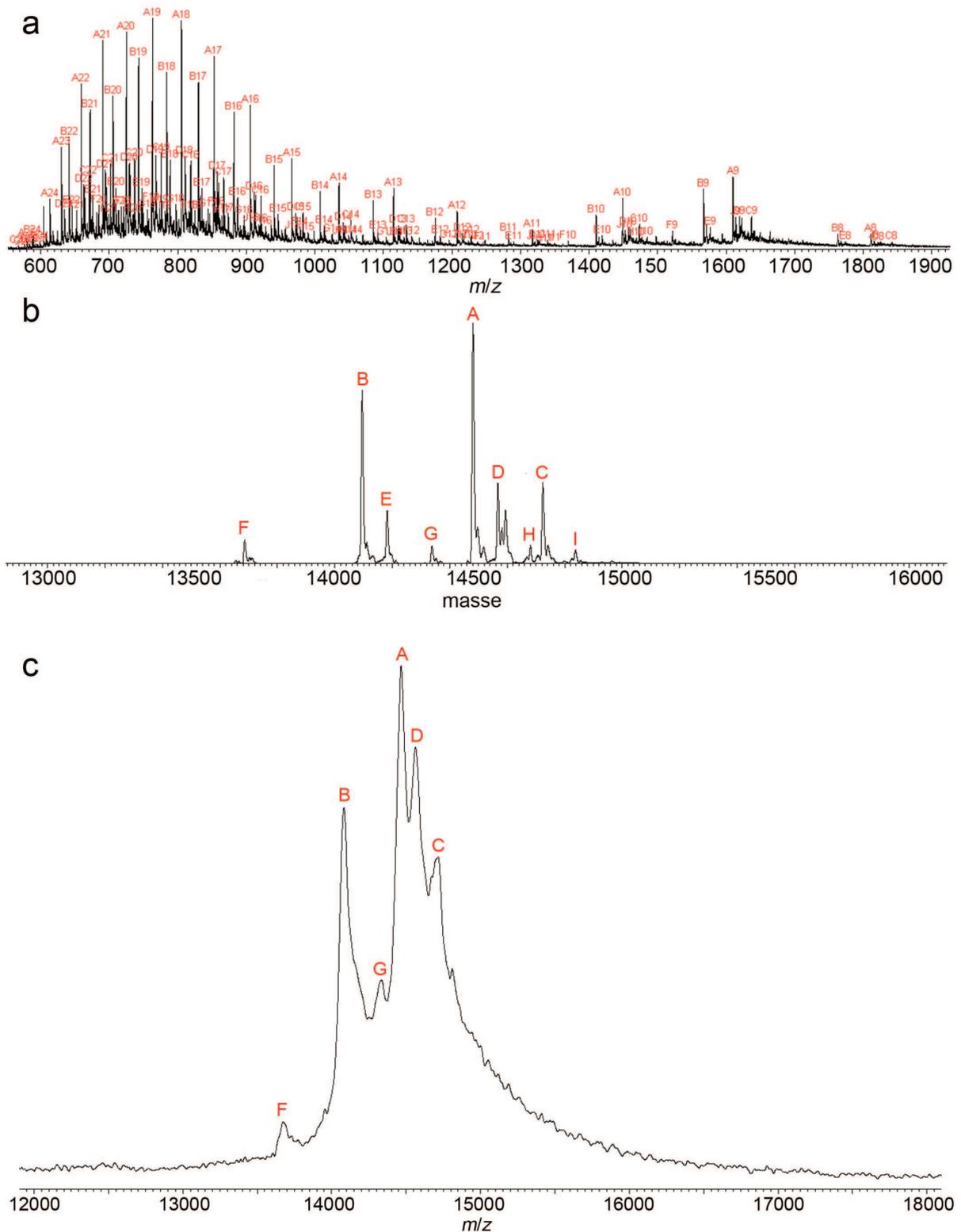


Figure 54. a. spectre ESI-MS de la protéine II-1NG ; b. déconvolution du spectre ESI-MS de la protéine II-1NG ; c. spectre MALDI-MS de la protéine II-1NG. Les 9 formes présentes sur ces spectres sont nommées de A à I (N.B. ce code de lettre est différent de celui des figures 55 à 63). Sur le spectre ESI-MS, les lettres d'identification sont accompagnées d'un chiffre correspondant à l'état de charge de l'ion correspondant.

Ainsi des chaînes polypeptidiques identiques ont été identifiées par ESI-MS et MALDI-MS. Elles correspondent à la séquence primaire théorique de II-1. Elles diffèrent entre elles de quelques acides aminés au niveau de leurs extrémités C- et N-terminales. Ces variations sont probablement dues à des phénomènes de protéolyse qui ont pu avoir lieu au sein du surnageant du fermenteur.

Tableau 9. Tableau récapitulatif des différentes formes de chaînes polypeptidiques de la protéine II-1NG et des masses correspondantes (attention le code de lettre de la figure 54 (code à une seule lettre) est différent de celui des figures 55 à 63 (code à deux lettres)).

Formes présentes sur la figure 54	Masse moléculaire observée (Da)	Identification (figure 55)	Masse moléculaire moyenne théorique (Da)
A	14480	C-y	14480.11
B	14095	F-y	14094.64
C	14723	C- α	14723.37
D	14566	C-z	14567.19
E	14182	F-z	14181.21
F	13685	H-y	13685.15
G	14338	F- α	14337.91
H	14680	B-z	14679.53
I	14836	B- α	14835.63

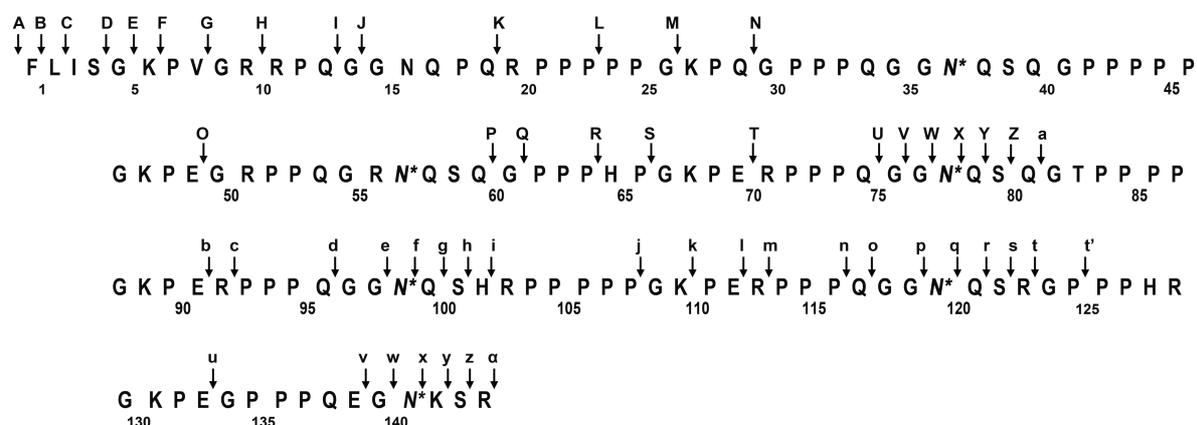


Figure 55. Position des débuts et des fins de séquence des différents peptides de II-1 et de II-1NG identifiés.

b. Caractérisation de la chaîne peptidique de II-1

La caractérisation de la chaîne peptidique de la protéine II-1 a été réalisée par l'analyse MALDI-MS de sa forme totalement déglycosylée (PNGaseF). La masse moléculaire apparente de II-1 sur gel est comprise entre 30 et 43 kDa (figure 56, piste 2). Après action de la PNGaseF, l'électrophorèse du milieu de déglycosylation a révélé la présence de plusieurs chaînes peptidiques dont les masses moléculaires apparentes sur gel se situent entre 12 et 20 kDa (figure 56, piste 3) confirmant l'activité de l'enzyme sur II-1 et le caractère glycosylé de cette dernière. Sur la même piste la bande aux alentours de 30 kDa correspond à l'enzyme PNGaseF.

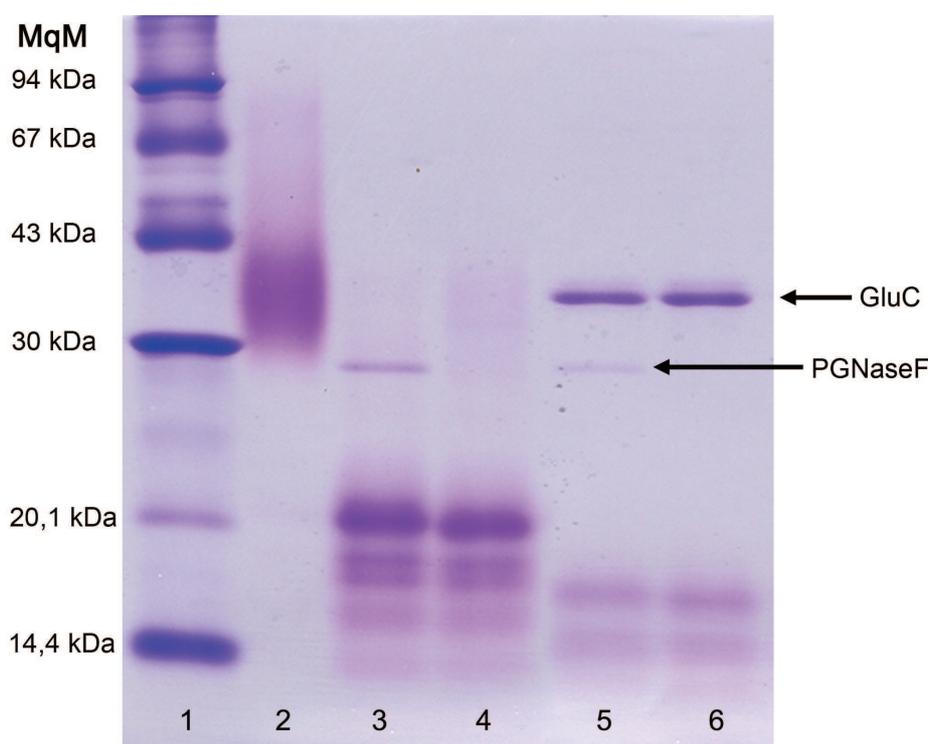


Figure 56. SDS-PAGE de la protéine II-1 avant et après traitement enzymatique. Piste 1 : MqM LWM ; piste 2 : II-1 purifiée ; piste 3 : II-1 déglycosylée par l'enzyme PNGaseF ; piste 4 : II-1 déglycosylée par l'enzyme endoH ; piste 5 : II-1 après action des enzymes PNGaseF et GluC ; piste 6 : II-1 après action des enzymes endoH et GluC.

Le spectre MALDI-MS correspondant est présenté sur la figure 57. En accord avec l'analyse électrophorétique, ce spectre révèle la présence de nombreux polypeptides de différentes masses moléculaires comprises entre 5 et 14 kDa. La plupart des polypeptides ont été identifiés sur la base de leur MM mesurée par MS. Ils correspondent à des fragments plus ou moins grands de la chaîne peptidique théorique de II-1, leur identification est reportée sur le spectre de masse (le code utilisé renvoie à la figure 55). Les peptides non identifiés sont nommés X1, X2, X3, ... Tous les polypeptides qui ont été identifiés sont plus

courts que la chaîne polypeptidique théorique de II-1. Les fragments polypeptidiques présents suggèrent une protéolyse importante de la protéine II-1. Il est probable que cette protéolyse ait eu lieu dans le milieu de culture du fermenteur où sont excrétées les protéines recombinantes d'intérêt et où peuvent être sécrétées des protéases alcalines et déversées des protéases cytoplasmiques lors de la lyse des levures. La séquence GGNQP décrite comme un site préférentiel de protéolyse (Argos, 1990 ; Yan *et al.*, 1999) est présente en position C-term de nombreux fragments détectés. Cette protéolyse de II-1 est en partie due à sa structure dépliée qui facilite l'accès aux sites reconnus par différentes protéases dont les sites de clivage ne sont pas gênés par la présence de proline (Receveur-Brechot *et al.*, 2006).

Cette protéolyse plus marquée que pour II-1NG provient probablement d'une purification moins résolutive lors de l'étape de séparation par gel filtration due aux glycanes sur sa chaîne polypeptidique qui augmentent l'encombrement stérique de la protéine et sa diversité. Ainsi, les chaînes polypeptidiques glycosylées courtes et longues n'ont pas été séparées lors de la purification de II-1 tandis qu'au cours de la purification de II-1NG, l'absence de glycosylation permet leur séparation.

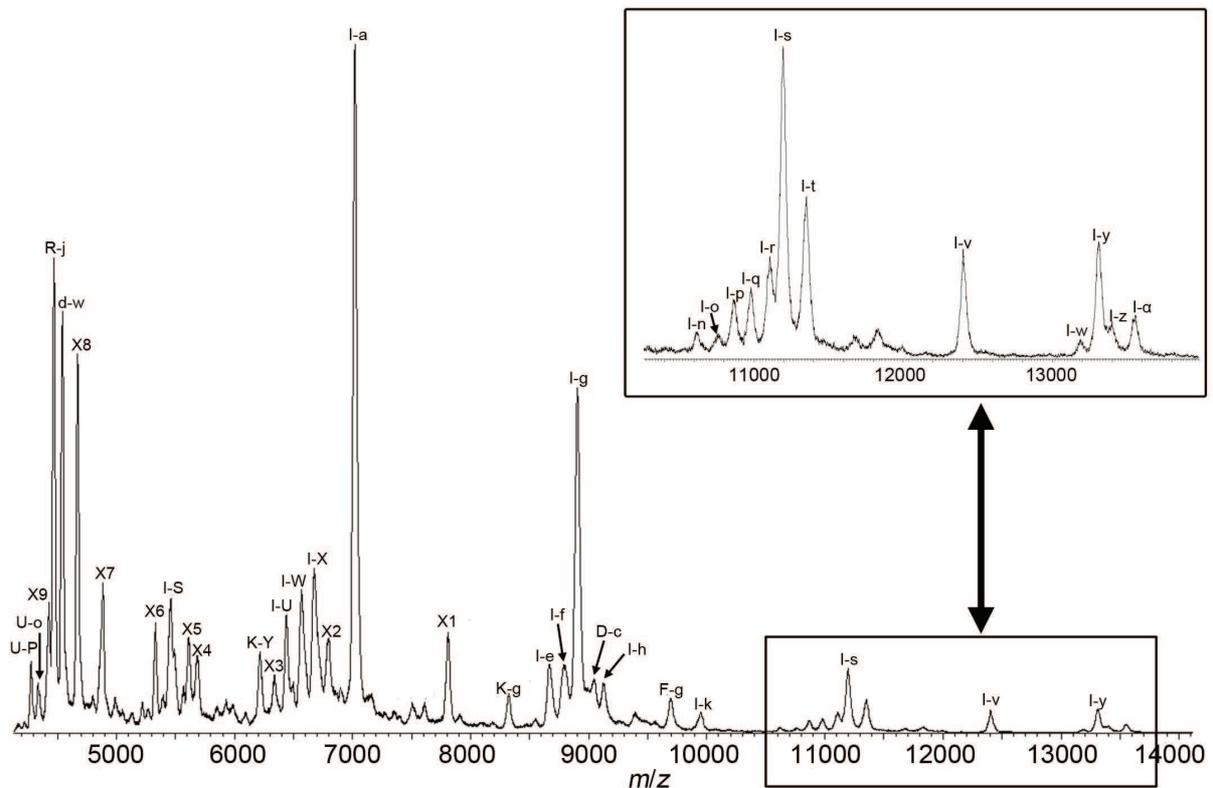


Figure 57. Spectre MALDI-MS de la protéine II-1 déglycosylée par la PNGaseF (se reporter à la figure 55 pour le code lettre utilisé).

I.2.2. Caractérisation de la protéine II-1

La protéine entière II-1 a été caractérisée par MALDI-TOF en mode linéaire. Le spectre de masse obtenu est présenté sur la figure 58.

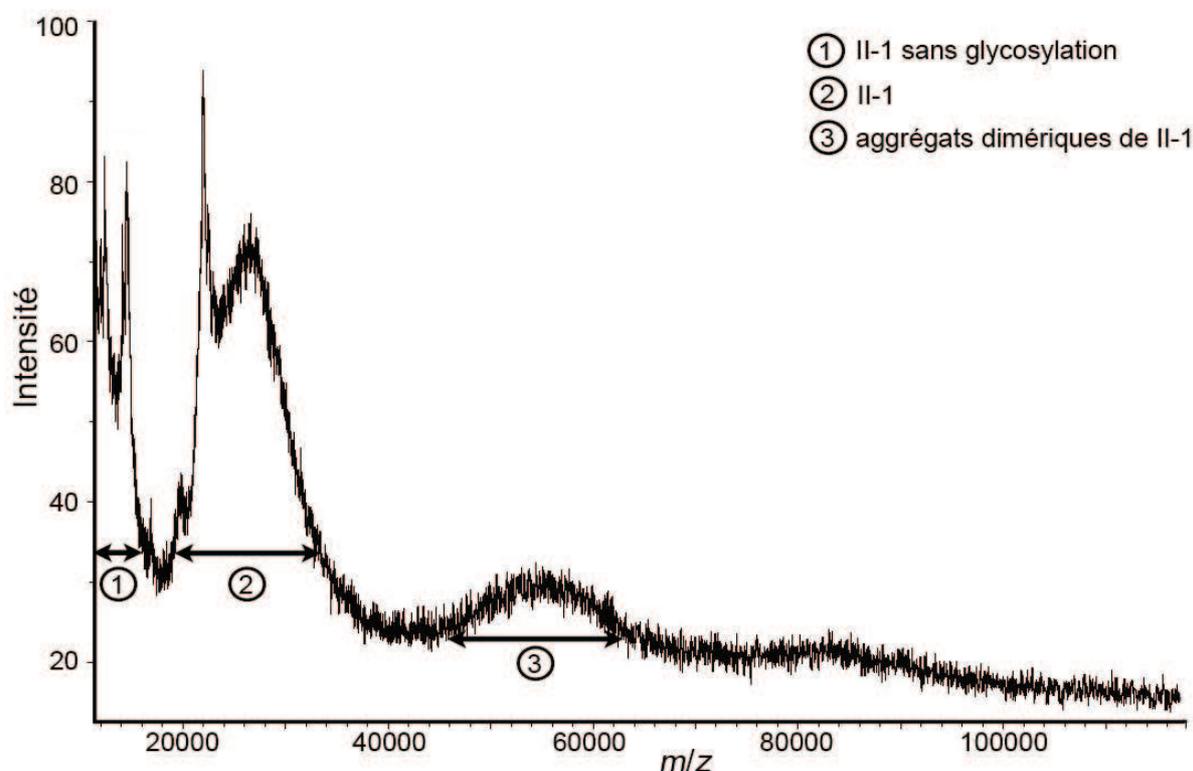


Figure 58. Spectre MALDI-TOF de la protéine II-1 acquis en mode positif.

Sur ce spectre, des signaux sont présents à des valeurs de m/z inférieures à 20 kDa. Ces pics peuvent être interprétés comme les différentes formes de chaînes polypeptidiques de II-1 ayant perdu leur(s) glycosylation(s) lors des processus de désorption et d'ionisation mis en jeu. Deux « massifs » se trouvant côte à côte sont situés entre 20 et 40 kDa. La plus forte intensité est observée pour un rapport m/z moyen centré sur 21,9 kDa, le deuxième rapport m/z moyen est centré sur 26,6 kDa. Enfin, un troisième « massif » d'intensité plus faible est présent entre 45 et 65 kDa avec un rapport m/z moyen centré sur 55,1 kDa.

Ces trois « massifs » peuvent être interprétés comme les ions des formes glycosylées de II-1. Leur manque de résolution et de définition est dû à la fois à l'importante hétérogénéité des chaînes polypeptidiques et des chaînes glycosidiques et à la résolution relativement faible des analyseurs de type TOF utilisés en mode linéaire.

L'hétérogénéité observée au niveau des glycosylations de II-1 est une caractéristique courante des glycoprotéines produites par voie recombinante. En effet, dans un système d'expression et pour des conditions de culture définies, les glycosyltransférases, enzymes précisément localisées et réalisant l'attachement des sucres aux protéines naissantes, sont

variées. Il en existe plus de 500 différentes (Colley, 1997 ; Drickamer et Taylor, 1998). Ces enzymes sont hautement spécifiques, présentes en faible quantité et chacune créant un seul type de liaison. Il en résulte une microhétérogénéité de la partie oligosaccharidique entre et sur chaque site de glycosylation pour une glycoprotéine donnée. Le qualificatif le mieux adapté pour définir la glycosylation est la « diversité », diversité à la fois de composition et de fonction de la partie oligosaccharidique. Cette variété contraste avec la conservation absolue de l'oligosaccharide précurseur lié au dolichol ($\text{Glu}_3\text{Man}_9\text{NAcGlu}_2$) utilisé pour initier la N-glycosylation dans le RE. La pléthore de structures générées reflète la maturation, c'est à dire les allongements multiples des N-oligosaccharides réalisés dans l'appareil de Golgi.

Les ions avec les rapports m/z les plus élevés peuvent être interprétés comme étant les formes glycosylées des chaînes polypeptiques les plus longues. Ainsi, la masse de la (des) chaîne(s) glycosique(s) liée(s) peut-être estimée aux environs de 13 kDa. *Pichia pastoris* produit des N-oligosaccharides constitués principalement de mannoses α -liés. Ces oligosaccharides majoritaires ont généralement pour formule brute $\text{Man}_{8-11}\text{NAcGlu}_2$. Enfin, le massif situé entre 45 et 65 kDa présente des valeurs de m/z qui sont environ le double de celles des deux autres. Il peut ainsi être interprétés comme le résultat de phénomènes d'aggrégation entre deux protéines. Ces phénomènes de « dimérisation » ont pu avoir lieu lors de la cristallisation de la matrice contenant l'échantillon.

Ainsi, la masse moléculaire de II-1 est comprise entre 20 et 30 kDa, en fonction de la variation de la longueur de sa chaîne peptidique et surtout de celle(s) de son (ses) glycane(s) qui représente(nt) environ 40% de la masse moléculaire de la protéine entière, en accord avec les résultats précédemment obtenus (Pascal *et al.*, 2006).

I.2.3. Détermination du nombre de glycanes portés par II-1

La détermination du nombre de N-glycanes portés par II-1 a été réalisée à partir de sa forme partiellement déglycosylée par l'enzyme endoH. Après action de l'endoH, l'électrophorèse du milieu de déglycosylation montre la présence de plusieurs chaînes peptidiques dont les masses moléculaires apparentes sur gel se situent entre 12 et 20 kDa (figure 56, piste 4) confirmant l'activité de l'enzyme sur II-1 et son caractère glycosylé.

L'analyse des spectres MALDI-MS révèle des différences de masse entre le spectre de masse de II-1 après action de la PNGaseF (figure 57) et le spectre de II-1 après action de l'endoH (figure 59) et permet de déterminer le nombre de glycosylations portées par chacun des polypeptides.

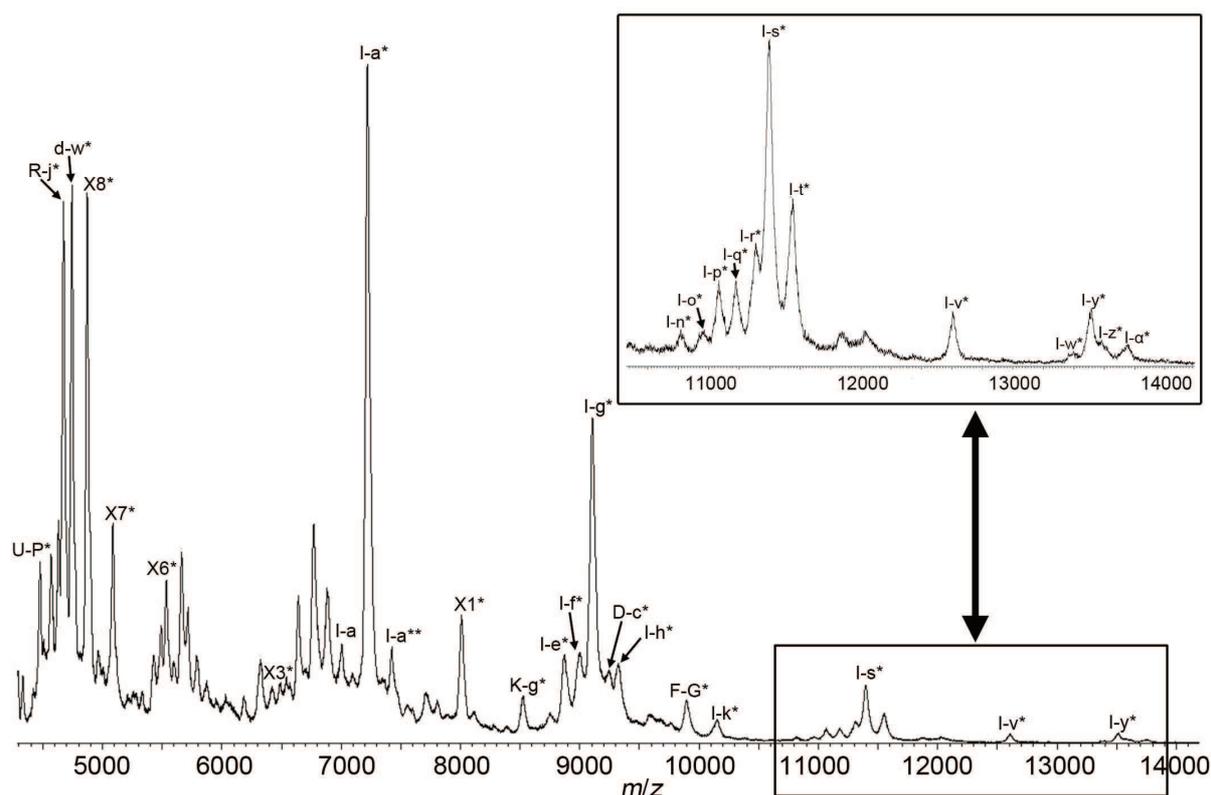


Figure 59. Spectre MALDI-MS de la protéine II-1 déglycosylée par l'endoH. Les espèces marquées * présentent une seule glycosylation et celle marquées ** en possèdent deux (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).

Cette étude montre que les chaînes peptidiques de II-1 ont en général une seule glycosylation. Ce nombre est relativement faible par rapport aux six possibilités existantes. Cela pourrait provenir du fait que la séquence consensus Asn-X-Ser est moins favorable à la glycosylation que la séquence Asn-X-Thr (Yan *et al.*, 1999).

Dans l'hypothèse où un seul site est effectivement glycosylé, la masse de sucres estimée de 13 kDa indique la présence d'une hyperglycosylation d'environ 75 mannoses. Même si l'hyperglycosylation (supérieure à 20 mannoses) est nettement moins fréquente chez *Pichia pastoris* que chez *Saccharomyces cerevisiae*, elle a déjà été observée à différentes reprises (Martinet *et al.*, 1997 ; Ruitenbergh *et al.*, 2001 ; Jiang *et al.*, 2002 ; Lawson *et al.*, 2002 ; Paramasivam *et al.*, 2002)

I.2.4. Détermination de la localisation des glycanes

L'étape suivante a consisté en la détermination de la localisation du (des) site(s) porteurs des N-glycosylations. Cette étude a été réalisée à l'aide de deux doubles digestions de II-1: PNGaseF + GluC et EndoH + GluC (figure 53). Après action des

enzymes, l'électrophorèse des deux milieux de déglycosylation a indiqué la présence de plusieurs peptides dont les masses moléculaires apparentes sur gel sont inférieures à 16 kDa (figure 56, pistes 5 et 6) confirmant l'activité des enzymes. La bande aux environs de 35 kDa sur chacune des deux pistes correspond à l'enzyme GluC. Sur la piste 5, la bande aux environs de 30 kDa correspond à la PNGaseF.

Pour chacune de ces doubles digestions, deux spectres MALDI-MS ont été enregistrés, l'un en mode linéaire sur une gamme de rapport m/z allant de 4 000 à 10 000 (figure 60) et l'autre en mode réflectron sur une gamme de rapport m/z allant de 1 000 à 5 800 (figure 61).

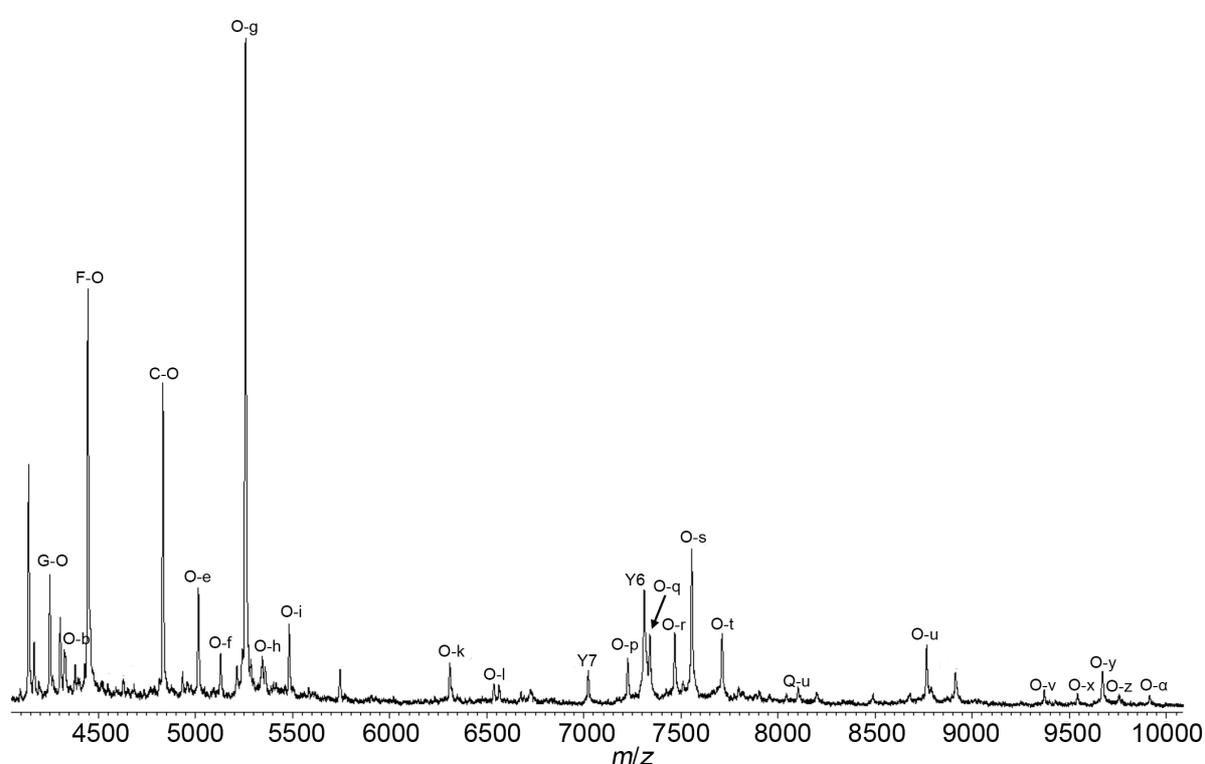


Figure 60. Spectre MALDI-MS en mode linéaire de la protéine II-1 déglycosylée par l'enzyme PNGaseF et clivée par l'enzyme GluC (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).

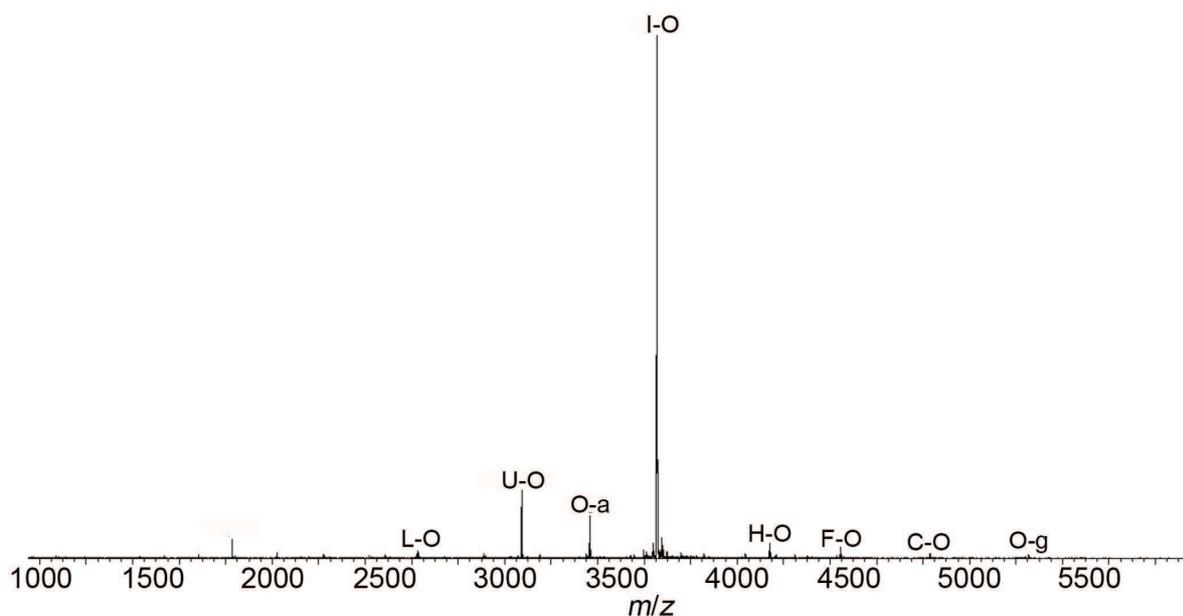


Figure 61. Spectre MALDI-MS en mode réflectron de la protéine II-1 déglycosylée par l'enzyme PNGaseF et clivée par l'enzyme GluC (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).

Les spectres MALDI-MS de la double digestion PNGaseF + GluC ont permis d'identifier les fragments peptidiques obtenus et d'établir la carte peptidique de l'hydrolysate issu de GluC. L'identité du peptide majoritaire sur le spectre MALDI-TOF en mode réflectron a été confirmée par analyse MS/MS. Il apparaît que l'enzyme GluC a clivé principalement la chaîne peptidique au niveau du premier site de clivage potentiel en partant de l'extrémité N-term. Ce clivage a été effectué au niveau du point marqué O sur la figure 56 et a généré deux types de fragments :

- des fragments comportant les extrémités N-term et incluant un seul site potentiel de glycosylation ;
- des fragments comportant les extrémités C-term et présentant 4 ou 5 sites potentiels de glycosylation.

Les fragments obtenus confirment l'identification proposée pour de nombreux polypeptides présents sur le spectre MALDI de II-1 déglycosylée par la PNGaseF. Cependant, si les fragments N-terminaux peuvent nous renseigner précisément sur la présence d'un N-glycane sur le premier site de N-glycosylation, les fragments C-term ne permettent pas de discriminer les autres sites.

L'analyse des spectres MALDI-MS de la double digestion EndoH + GluC révèle la présence de fragments glycosylés et non-glycosylés (figures 62 et 63). Elle montre que de

nombreux fragments sont présents à la fois sous forme glycosylée et sous forme non-glycosylée, indiquant une glycosylation hétérogène et aléatoire. L'étude par fragmentation MS/MS du peptide I-O issu de la double digestion Endo H/GluC a confirmé que le résidu Asn 37 portait une chaîne glycosylée. Les fragments comportant les extrémités C-terminales contiennent plusieurs sites potentiels de N-glycosylation, ce qui n'a pas permis de discriminer précisément le(les) site(s) porteur(s) de la glycosylation. Ces polypeptides présentent également une glycosylation ou non. La taille de ces fragments n'a pas permis leur analyse par MS/MS PSD (hors de la gamme de masse des ions sélectionnables).

Ces observations sont en accord avec l'analyse de II-1 traitée par l'endoH qui a révélé que II-1 porte majoritairement une seule glycosylation qui semble distribuée de manière aléatoire sur différents sites bien que majoritairement sur Asn 37.

La protéine II-1 purifiée présente donc une diversité importante à la fois au niveau de la longueur de la chaîne polypeptidique mais également au niveau de la position des N-glycanes qu'elle porte et de leur longueur. L'importante hétérogénéité de la chaîne peptidique résulte de phénomènes de protéolyse auxquels les PRP sont particulièrement sensibles (Receveur-Brechot *et al.*, 2006).

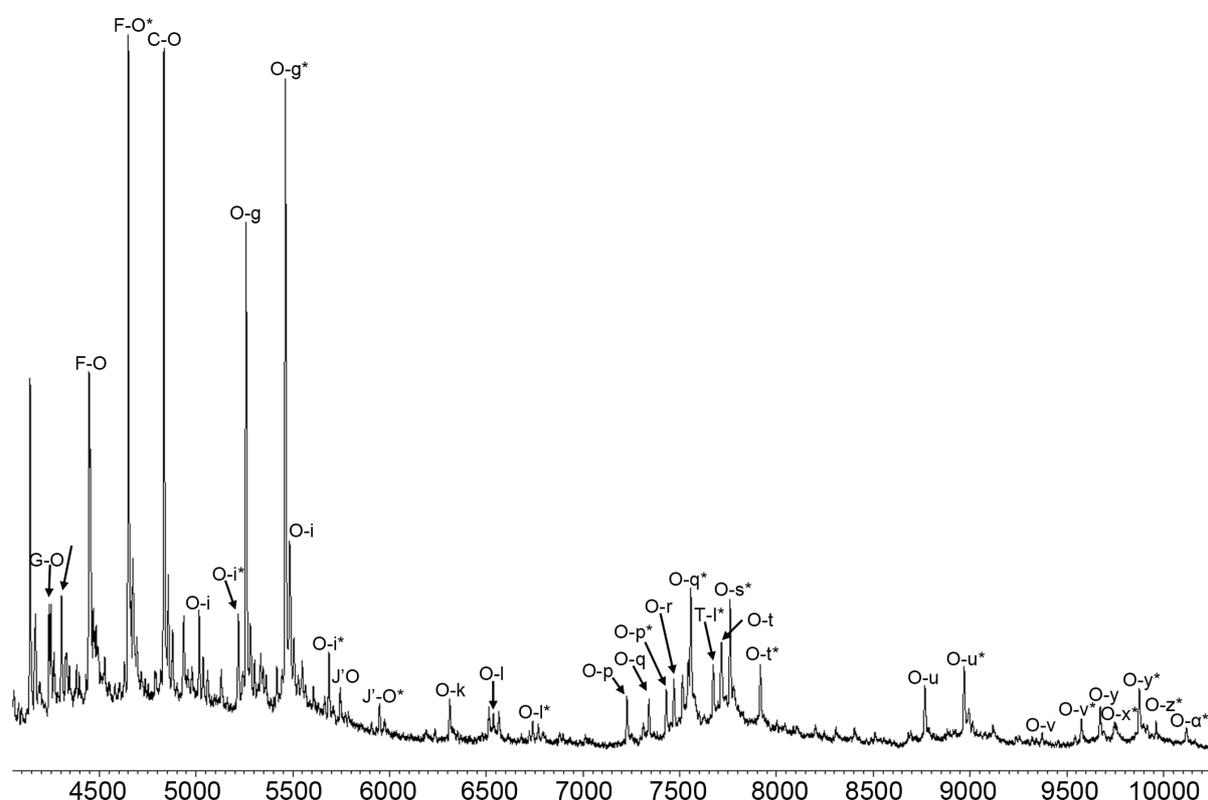


Figure 62. Spectre MALDI-MS en mode linéaire de la protéine II-1 déglycosylée par l'enzyme endoH et clivée par l'enzyme GluC. Les espèces marquées * présentent une seule glycosylation et celle marquées ** en possèdent deux (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).

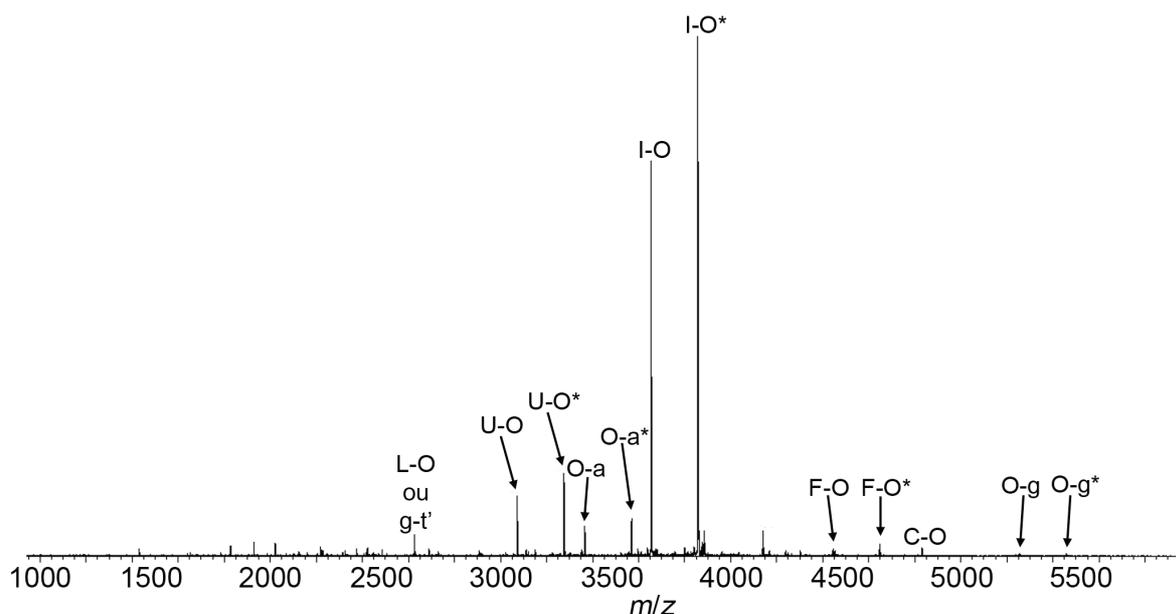


Figure 63. Spectre MALDI-MS en mode réflectron de la protéine II-1 déglycosylée par l'enzyme endoH et clivée par l'enzyme GluC. Les espèces marquées * présentent une seule glycosylation et celle marqués ** en possèdent deux (les séquences qui correspondent aux codes des peptides sont donnés dans la figure 55).

I.3. Caractérisation de la protéine IB5

Les lots de protéine IB5, *bPRP* non glycosylée, ont été caractérisés par électrophorèse, séquençage N-terminal d'Edman et ESI-MS.

La figure 49 illustre l'analyse SDS-PAGE des profils protéiques obtenus au cours de des différentes étapes de purification du protocole n°1. Outre le comportement métachromique de cette protéine lors de la coloration au Bleu de Coomassie (coloration rose au lieu de la coloration bleu typique des protéines (Beeley *et al.*, 1991b)), le caractère particulier des PID a encore une fois été révélé par la migration « atypique » d'IB5 en SDS-PAGE en conditions dénaturantes. En effet, la MM calculée théorique pour IB5 est de 6,9 kDa alors que sa masse apparente sur gel est d'environ 14 kDa (figure 50).

Le séquençage N-terminal de la protéine IB5 purifiée a révélé la présence de plusieurs chaînes polypeptidiques différant de quelques acides aminés. Parmi celles-ci, la séquence attendue d'IB5 (figure 42) est majoritaire.

La caractérisation d'IB5 par MS a été réalisée à l'aide d'un spectromètre de masse de type ESI-TOF muni d'un réflectron. La résolution du spectre a permis d'attribuer les états de charge de chacun des pics et de déduire la masse des ions correspondants. A partir des

masses obtenues et de la séquence théorique d'IB5, les séquences des polypeptides présents ont été déterminées. Les données relatives à IB5 purifiée selon le protocole n°1, décrit dans la partie *Matériels et Méthodes*, sont détaillées dans la publication « Characterization, stoichiometry and stability of salivary protein-tannin complexes by ESI-MS and ESI-MS/MS » (Publication n°1, chapitre II, pp. 144 à 167). La présence de trois chaînes peptidiques qui diffèrent de quelques acides aminés a été mise en évidence. La forme majoritaire correspond à la séquence attendue de la protéine et les deux autres formes comportent respectivement 3 et 6 acides aminés en moins du côté N-terminal (figure 2, publication n°1). Ces résultats sont en accord avec ceux obtenus par séquençage N-terminal d'Edman.

Les fractions d'IB5, purifiées selon le protocole n°2, ont été caractérisées par MS suivant la même méthodologie (figure 64.a).

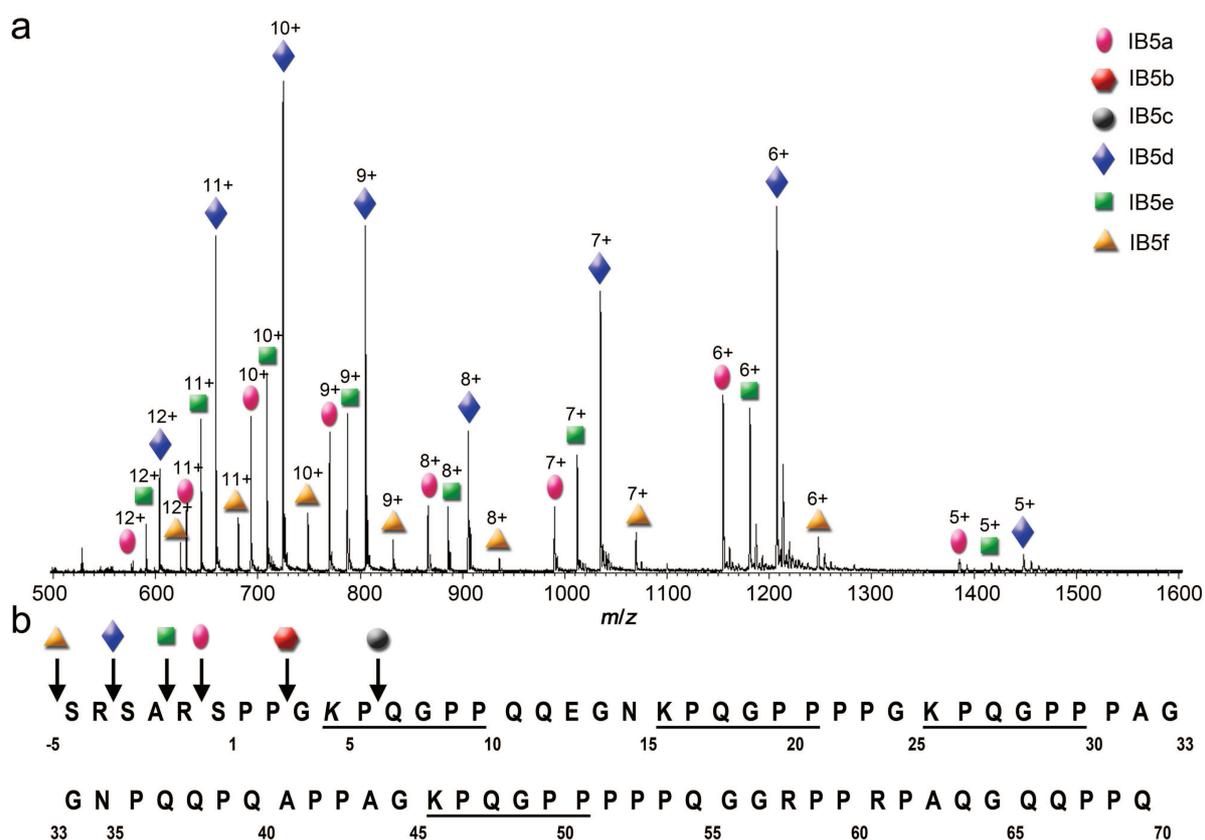


Figure 64. a. spectre de masse d'une fraction d'IB5 purifiée selon le protocole n°2 ; b. séquence des chaînes polypeptidiques identifiées (les séquences répétées sont soulignées).

A partir des masses obtenues et de la séquence théorique d'IB5, les séquences des polypeptides présents ont été déterminées et sont présentées sur la figure 64.b. Alors que le surnageant de fermenteur est le même, la protéine IB5 issue de ce mode de purification n°2

est composée de chaînes polypeptidiques dont les séquences terminales sont légèrement différentes de celles des protéines purifiées selon la première méthode. Parmi ces différentes formes, la chaîne polypeptidique attendue d'IB5 (IB5a) est présente mais elle n'est plus la forme majoritaire. Les trois autres chaînes polypeptidiques ont respectivement un, trois et cinq acides aminés en plus au niveau de l'extrémité N-terminale. Cette différence de quelques acides aminés pourrait être le résultat d'une protéolyse au niveau des extrémités N-terminales au cours de la purification suivant la procédure n°1, dont la durée totale est largement supérieure à celle de la procédure n°2.

La majorité des travaux présentés dans la suite de ce mémoire a été réalisée avec le lot de protéine IB5 purifié selon la première méthode.

Chapitre II. Etude des interactions IB5-EgCG

II.1. Introduction

Afin d'étudier les interactions PRP-tanin, nous avons développé une approche par spectrométrie de masse (MS). Avec l'apparition de sources d'ionisation douces de type ESI, cette technique est devenue un outil puissant pour l'étude des complexes non-covalents impliquant des protéines (Robinson, 2002). Sa précision de mesure permet une caractérisation fine des partenaires et des objets supramoléculaires résultant de leur interaction. L'analyse par MS détermine entre autres la composition et les stœchiométries des complexes formés en solution. Certains spectromètres de masse sont également capables d'effectuer des expériences dites de MS/MS au cours desquelles un ion est isolé avant d'être dissocié. Ceci permet à la fois de confirmer l'identification de cet ion et d'étudier sa stabilité. En effet, des courbes de dissociation peuvent être établies en augmentant progressivement l'énergie interne ajoutée aux ions.

II.2. Etudes des interactions IB5-EgCG par ESI-MS et ESI-MS/MS

CID

Les études des interactions PRP-tanin par MS ont concerné tout d'abord les interactions s'établissant entre la protéine IB5 et un tanin modèle, l'épigallocatechine gallate (EgCG, figure 43.a). Ce monomère a été choisi en raison de son « affinité » reconnue pour les protéines. En effet, la présence d'un acide gallique comme substituant sur le noyau flavanol semble favoriser les interactions protéine-tanin insolubles et solubles (Cheynier *et al.*, 1997 ; Sarni-Manchado et Cheynier, 1999 ; de Freitas et Mateus, 2001 ; Charlton *et al.*, 2002b ; Sarni-Manchado et Cheynier, 2002 ; Maury *et al.*, 2003 ; Pascal *et al.*, 2007). Les complexes formés entre IB5 et l'EgCG ont tout d'abord été détectés et caractérisés par ESI-TOF MS puis leur stabilité a été évaluée par des expériences de CID à l'aide d'un spectromètre de masse en tandem de type ESI-Q-TOF.

Les résultats obtenus ont fait l'objet d'une première publication dans la revue internationale à comité de lecture *Analytical and Bioanalytical Chemistry* (publication n°1).

Publication n°1 : Characterization, Stoichiometry and Stability of Salivary Protein-tannin Complexes by ESI-MS and ESI-MS/MS

Francis Canon†, Franck Paté†, Emmanuelle Meudec†, Thérèse Marlin†, Véronique Cheynier†, Alexandre Giuliani‡, § and Pascale Sarni-Manchado†*

† *Unité Mixte de Recherche 1083 SPO, Polyphenol Interaction Research Group, Institut National de la Recherche Agronomique (INRA), Bât 28, 2 place Viala F-34060 Montpellier France, sarni@supagro.inra.fr, Phone: 33 4 99 61 22 98, Fax: 33 4 99 61 26 83.*

‡ *DISCO beamline, Synchrotron Soleil, l'Orme des Merisiers, F-91192 Gif sur Yvette France.*

§ *Cepia, Institut National de la Recherche Agronomique (INRA), BP 71627, F-44316 Nantes Cedex 3 France.*

*Reprint request to Dr. Pascale Sarni-Manchado, INRA, Unité Mixte de Recherche 1083 SPO, Polyphenol Interaction Research Group, INRA, Bât 28, 2 place Viala F-34060 Montpellier France, sarni@supagro.inra.fr, Phone: 33 4 99 61 22 98, Fax: 33 4 99 61 26 83.

Keywords polyphenol, interaction, Proline Rich Protein, saliva, astringency, IUP

Abbreviations CID, Collision-Induced Dissociation; EgCG, epigallocatechin gallate; ESI, electrospray ionization; IUP, Intrinsically Unstructured Protein; MS, mass spectrometry; MS/MS, tandem mass spectrometry; PRP, Proline Rich Protein; Q-TOF, quadrupole/time-of-flight

ABSTRACT

Numerous protein-polyphenol interactions occur in biological and food domains particularly involving proline rich proteins, representative of the Intrinsically Unstructured Protein group (IUP). Noncovalent protein–ligand complexes are readily detected by electrospray ionization mass spectrometry (ESI-MS), which also gives access to ligand binding stoichiometry. Surprisingly, the study of interactions between polyphenolic molecules and proteins is still an area where ESI-MS has poorly benefited whereas it has been extensively applied to the detection of noncovalent complexes. Electrospray ionization-mass spectrometry has been applied to the detection and the characterization of the complexes formed between tannins and a human salivary proline rich protein (PRP), namely IB5. The study of the complex stability was achieved by low-energy Collision-Induced-Dissociation (CID) measurements, which are commonly implemented using triple quadrupole, hybrid quadrupole time-of-flight or ion trap instruments. Complexes composed of IB5 bound to a model polyphenol

II. Etude des interactions IB5-EgCG

EgCG have been detected by ESI-MS and further analyzed by MS/MS. Mild ESI interface conditions allowed us to observe intact noncovalent PRP-tannin complexes with stoichiometries ranging from 1:1 to 1:5. Thus, ESI-MS shows its efficiency for 1) the study of PRP-tannin interactions, 2) the determination of stoichiometry and 3) the study of complex stability. We were able to establish unambiguously both their stoichiometries and their overall subunit architecture via tandem mass spectrometry and solution disruption experiments. Our results prove that IB5•EgCG complexes are maintained intact in the gas phase.

INTRODUCTION

Demonstrations of proteins having little or no ordered structure under physiological conditions have increased exponentially over the past 15 years. Particular terms have been proposed to describe these proteins: Intrinsically disordered proteins (IDPs), natively unfolded or intrinsically unstructured proteins (IUPs) [1-3]. Proline-rich proteins (PRPs), which are IUPs, are particularly abundant in human saliva. They may constitute 70% of total secreted proteins in parotid saliva [4]. The salivary proteins play various roles (lubrication, digestion, taste, protection ...) as saliva is the first fluid which interacts with food constituents. Herbivorous and omnivorous mammals have developed the ability to synthesize salivary basic proline rich proteins whose only known role is to bind tannins [5].

Plants synthesize a diverse group of phenolic compounds (*syn.* polyphenols) [6]. Among them, proanthocyanidins (*syn.* condensed tannins) are ubiquitous in plants. They are major components of foods (*e.g.* fruit, cocoa), beverages (*e.g.* wine, tea, beer, cider,...), animal feeds (*e.g.* pasture legumes) and herbal remedies. These molecules have attracted considerable interest because of their major contribution to organoleptic and biological properties. One of their prominent characteristics is their ability to interact with proteins, which can be either detrimental or beneficial. Haze developments are observed in beverages [7-9], reducing their quality. Inactivation of digestive enzymes and/or tanning of diet proteins impede their digestibility and assimilation, and thus lowers nutritional value of some feeds (*e.g.* legumes) [10]. On the other hand, polyphenol-protein interactions are also involved in the beneficial antioxidant activity of polyphenols as enzyme inhibitors that limit superoxide and uric acid formation [11]. Recently, their participation in the regulation of gene expression, due to their ability to interact with receptors and nuclear factors, has been demonstrated [11].

According to Red Queen hypothesis² [12], animals developed shunning mechanisms along with the ability to perceive astringency. Astringency is a tactile sensation [13-16] resulting from interactions of tannins and showing some specificity [17] for which salivary PRPs are often evoked [18]. However, its physico-chemical bases are still poorly understood. Extensively studied, using a variety of techniques and experimental models in relation with astringency and anti-nutritional effects, protein-tannin interactions are reported to generate insoluble and soluble complexes [19-22]. Earlier works have measured protein and/or tannins remaining in solution after precipitation. Particular proteins have been proposed to

² Van Valen's Red Queen hypothesis is a model of coevolution driven by competitive interactions between species. It contrasts with the stationary or 'lost world' model, in which evolution is driven primarily by environmental change.

study protein-polyphenol interaction [23,24]. However, these measurements may not reflect astringency as salivary protein-tannin interactions show some specificity [17]. Moreover, they do not take into account the formation of soluble complexes. The first information on the structure of soluble complexes has been provided by spectrometric methods such as Nuclear Magnetic Resonance (NMR) [25-28] sometimes associated with molecular modeling [20,29], circular dichroism (CD) [20,21], and mass spectrometry [20,30-32]. However, only peptides have been considered in these studies because of the unavailability of model natively unstructured PRPs.

Mass spectrometric procedures have greatly evolved and became methods of choice for characterization of biological macromolecules as well as their interactions. Electrospray ionisation (ESI) has been demonstrated to be efficient to probe weak associations of molecules in solution such as receptor-ligand and protein-nucleic acid interactions [33,34]. ESI-MS offers speed and sensitivity in monitoring components of mixtures. Determination of dissociation or relative binding affinities of noncovalent complexes is now well established [35-37]. The stabilities of noncovalent complexes have been assessed by their resistance to dissociation in the mass spectrometer using Collision-Induced Dissociation (CID) or thermal denaturation experiments [38-40]. Recently, Xie and Loo wrote that "The direct detection of protein-ligand noncovalent complexes using ESI mass spectrometry can be challenging, depending on the types of compounds tested and the types of physical interactions necessary to maintain a stable complex in both solution and in the gas phase.". This is certainly one of the reasons why few mass spectrometry publications concern protein-polyphenol complexes. Although these experiments have to be carefully interpreted with regards to solution-phase studies, they can provide information that is difficult to determine otherwise: the contribution of intermolecular interactions to ligand binding, free of any solvent contribution. MS and MS/MS allow probing experimentally intermolecular interactions [41].

We report here the first study of noncovalent complexes of the human salivary protein IB5 with tannins using ESI-MS and MS/MS. IB5 is an intrinsically unstructured proteins of human saliva, whose interactions with tannins may be involved in astringency.

MATERIALS AND METHODS

Samples

Epigallocatechin gallate (EgCG) and reserpine (Fig. 1a and 1b) were purchased from Sigma (Sigma Chemical Co., Poole, Dorset, U.K.).

The human salivary proline rich protein, IB5, was produced by use of the yeast *Pischia pastoris* as a host organism and purified as previously described [42].

EgCG and protein stock solutions were prepared in the following medium: water/ethanol, 88:12 (v/v) acidified to pH 3.1 with acetic acid. They were immediately frozen (-20°C) to prevent EgCG oxidation and proteolysis. Before use, they were diluted at room temperature to the desired concentrations.

EgCG-Protein interactions

Formation of complexes was studied at pH 3.1 that corresponds to the average pH value of the mouth a red wine. For IB5-EgCG interactions, the protein concentration was kept constant at 5µM in the water/ethanol medium described above. EgCG and protein solutions were mixed extemporaneously at room temperature (regulated at 24°C) so as to obtain a protein/polyphenol molar ratio of 1:10.

IB5 was mixed with EgCG (final concentration 5µM and 50 µM respectively) and allowed to equilibrate at room temperature for at least 5 min before ESI-MS analysis. All samples were prepared in triplicate and analyzed by ESI-MS.

Electrospray ionization mass spectrometry

MS spectra were obtained on a AccuTOF (JEOL, Tokyo, Japan) mass spectrometer equipped with an ESI source and a time-of-flight (TOF) mass analyzer, providing a resolving power of 6000 at m/z 609. Each solution was injected into the mass spectrometer by means of a syringe infusion pump with a flow rate of 20 µL.min⁻¹. The source voltage was set at 2350 V in positive ion mode, the orifice 1 voltage at 35 V, the orifice 2 voltage at 5 V, the ring voltage at 10 V and the capillary temperature at 150 °C. Spectra were acquired in positive ion mode with a m/z ranging from m/z 300 to 1800. Each mass spectrum was acquired with an integration time of 1 s. The signal from 28 scans was accumulated in the multi-channel acquisition mode and averaged.

Tandem mass spectrometry

Tandem mass spectrometry experiments were performed on a hybrid quadrupole time-of-flight (Q-TOF) Qstar Pulsar i mass spectrometer (Applied Biosystems, Forster City, CA), providing a resolving power of 10000 at m/z 1080.8. The Q-TOF was operated with the

Turbospray (electrospray ion source). The interface was fitted with the ion cooler guide to preserve the noncovalent complexes upon transfer into the gas phase. The IonCooler Guide is fitted around the initial ion entrance region of Q0 and increases the local pressure to greater than 30 mTorr, which greatly improves detection of weakly-bound complexes [43]. The sample was injected into the mass spectrometer by means of a syringe infusion pump with a flow rate of 7 $\mu\text{L}\cdot\text{min}^{-1}$. The source voltage was set to 5800 V in positive ion mode, the declustering potential at 47 V, the focusing potential at 209 V, the declustering potential 2 at 17 V, the ion source gas 1 set at 22, the ion source gas 2 at 0, the curtain gas at 10 and the capillary was not heated up. Nitrogen from a generator, containing less than 6 ppm oxygen was used both as nebulizing and desolvation gas. During MS/MS experiments, nitrogen was used as collision gas in the collision cell (Linac) with CAD gas setting of 5. The collision-set energy voltage V_c was increased from 0 to 50 V. The kinetic energy of the precursor ion in the laboratory frame of reference is $E_{Lab} = z \times e \times V_c$, where z is the number of charges on the ion and e is the charge of an electron [44]. Data were acquired by the TOF analyzer at 1 acquisition/s with m/z ranging from 250 to 1800.

Analysis of ESI MS and MS/MS experiments and calculations

All experiments were repeated at least three times. Ion pattern and intensities were compared. ESI-MS and collision MS/MS data were analysed with Analyst QS [45] and MagTran [46] softwares.

A procedure adapted from Jørgensen et al. [47] and Wan et al. [48] was used to establish the dissociation curves. The 50% dissociation energy is defined as the energy giving a 50% loss of precursor complex ions. Based on the data of Akashi et al. [49], one of the multiprotonated molecules of the IB5•EgCG complex (M), $[\text{M}+6\text{H}]^{6+}$ was selected and submitted to CID.

RESULTS AND DISCUSSION

ESI analysis of the protein solution

Fig. 2a shows a representative positive electrospray ionization spectrum of purified IB5 (5 μM in $\text{H}_2\text{O}/\text{EtOH}$, 88:12 v/v pH 3.1). The theoretical average molecular mass (MM) of IB5 is 6923.74 Da (Table 1). The positive ESI mass spectra of IB5 showed a series of ion peaks from m/z 1385 to 630, with charge states ranging from +5 to +11 which was confirmed by TOF-MS analysis (data not shown). Peaks have been assigned to protonated molecules.

The IB5 spectrum analysis and deconvolution resulted in three molecular mass values of 6923.70, 6642.63, and 6360.39 Da. It reveals the presence of three sequences of IB5, namely IB5a, IB5b and IB5c, which correspond respectively to the whole protein and to N-term three and six amino acid deleted proteins (Fig. 2b), with IB5a being the main form (>65%). Each form of IB5 exhibits the same charge state distribution (CSD). Typical protein purification techniques (and particularly for IUPs [1]) do not allow the separation of proteins with such small differences in length, the overall sequence remaining exactly the same. Edman sequencing experiments confirmed that IB5b and c are not the result of fragmentation during the ionization but co-exist in the purified fraction with IB5a. They may result from proteolytic activity of *Pichia pastoris* during protein expression or remaining during the protein purification procedure. It is worth noticing that peaks corresponding to multimers (i.e. aggregates) of IB5 were not observed, consistent with the absence of contribution from non-specific association under the solution and instrumental conditions selected.

Under acidic conditions, the basic residues Arg, Lys, His and the NH₂ terminus are often the only sites of protonation, but in the gas phase [50,51], Pro and Gln behave as basic residues that can be protonated. The high charge states observed for IB5 can thus be related to the ability of Pro residues to be easily protonated. Investigations on the effect of pH and organic solvent on the CSD of IB5 performed from pH 2.6 to 5.7 in the experimental medium and at a pH value of 1.3 in acetonitrile/water/TFA 80:20:0.1% (v/v) showed no difference in the MS spectra. The natively unfolded nature of IB5 explains these data as evidenced earlier by CD experiments [21]. The highest charged ions may represent extensively unfolded solution conformers whereas the low charge states are assigned to more “compact” series [52].

IB5 fragmentation experiments via CID MS/MS performed on the Q-TOF instrument demonstrated that the highest charged ions were less stable than the lowest charged ones as reported for other proteins and according to the theory of proton mobility [53]. Whereas most investigations on proline rich sequences under CID have concerned peptides [54,55], we examined the fragmentation of a whole PRP, IB5, 70 aa long. The ESI MS/MS spectrum of the IB5⁷⁺ ion is shown in Fig. 3a Fig. 3b presents cleavages that occur on the IB5 sequence. The CID spectra of the parent ion are primarily composed of singly, doubly, triply and some quadruply charged fragments, and the C-terminal y series dominates. y₂₉ and y₃₂ ions, corresponding to cleavages on the N-terminal side of proline residues, are the most abundant product ions. Y sequence ions are mainly observed in the center part of the protein backbone. b₃₈, b₃₇, and b₃₅ ions are present with their respective corresponding c ions. Moreover, many internal fragment ions are detected. Their interpretation remains difficult because of repeated sequences and proline abundance in IB5. This protein presents a low complexity primary sequence, with only 9 different residues, and within the 70 amino acids,

29 are P, 13 are G and 13 are K, with repeats of KPQGPP(P). P, G and K residues are interesting from a mass spectrometry point of view [56]. Indeed, the presence of proline in a peptide induces the so-called 'proline effect' which is characterized by the presence of a labile amide bond on the N-terminal side of proline and a stable amide bond on its C-terminal side which is associated with the abundance of γ sequence ions [55]. Fragmentation of proline containing peptides, particularly when poly-Pro stretches exist, produces unusual fragment ions in high-energy experiments, some of which are suggested to be produced by charge-remote cleavage [54]. Complete sequence information cannot be obtained because fragment ions arising by cleavages at not adjacent to proline amide bonds have very low abundances or are missing [57].

ESI-MS analysis of IB5•EgCG complexes

Concentrations of condensed tannin in red wine are approximately $1\text{-}2\text{ g}\cdot\text{L}^{-1}$ corresponding to a mean molar concentration of $0.5\text{-}1\text{ mM}$. In biological media as well as in processing media, polyphenols are often present together with proteins. For example, fining treatment, consisting in the addition of a protein fining agent, is applied to precipitate tannins from plant-derived beverages, so as to decrease haze and/or astringency. Fining agents are gelatin fractions submitted to different hydrolysis leading to different protein sizes. In the case of a red wine, addition from 8 to 20 g of protein for one hL of wine is of general use [58]. Thereby, the concentration of a fining protein showing a 10 kDa apparent MM at $10\text{ g}\cdot\text{hL}^{-1}$ can be estimated around $20\text{ }\mu\text{M}$. The presence of food in the mouth initiates both mechanical and chemical stimuli via neural reflexes that result in an increased secretion of saliva with taste being the main stimulant for this secretion. Thus, the confrontation of polyphenols to salivary proteins is the first step of ingestion of plant products. The concentration of proteins in saliva is approximately $1\text{ mg}\cdot\text{mL}^{-1}$ circa $83\text{ }\mu\text{M}$. Polyphenols are present in large excess compared to proteins both in wine and in the mouth after sipping of wine, making the protein/polyphenol ratio always in the benefit of polyphenols. On this basis, to investigate soluble protein-polyphenol complexes we use a protein/polyphenol ratio of 1:10 that does give rise neither to the formation of precipitate nor cloudiness.

Premixed interaction medium with a 1:10 IB5/EgCG molar ratio ($5\mu\text{M}$ of IB5 with $50\mu\text{M}$ of EgCG in $\text{H}_2\text{O}/\text{EtOH}$ 88:12 v/v pH 3.1) was introduced to ESI mass spectrometer. The successful analysis of noncovalent assemblies by MS requires the preservation of the native state of the complex in solution. It is crucial to use aqueous solutions containing only volatile buffers [59]. The MS analysis has been performed in water/ethanol allowing satisfactory levels of desolvation [59]. Analysis of the ESI mass spectrum (Fig. 4a) revealed three distinct

entities, IB5, EgCG and noncovalent soluble complexes in agreement with earlier Isothermal Titration Microcalorimetry (ITC) and CD experiments [21]. It is noteworthy that no indication about their stoichiometries have been obtained by other approaches. The CSD of these ions ranged from +6 to +10 like those of individual IB5 proteins. IB5a complexes remain predominant as expected from the protein sample analysis. Fig. 4b presents a blow-up on the region of the spectrum corresponding to the +7 ions. This is the main charge state for both IB5 and IB5•EgCG complexes. It can be noted that numerous stoichiometries ranging from 1:1 to 1:5 are observed. ESI mass spectra performed on a mixture of IB5 and reserpine, which presents a chemical structure with aromatic rings and a molecular mass close to that of EgCG did not give rise to complex peaks (data not shown). If the interaction is not driven by polyphenols, ESI MS should have generated complexes between IB5 and reserpine. Use of reserpine allows us to confirm that the formation of complexes is related to tannins.

From previous solution studies, proline rich peptide-tannin interaction involves stacking of the planar proline cycle with the phenolic ring, stabilized by hydrogen bonding interactions with the peptide bonds next to the proline and with other amino acids like glycine [20,26,60]. Involvement of both enthalpy-driven (attributed to hydrogen bonding) and entropy-driven (associated to hydrophobic effect and conformational changes) phenomena in interactions of flavan 3-ols with poly-L-proline was confirmed by ITC [61]. The former is prevalent in the case of monomers and the latter in that of polymers. Survival of small protein-ligand complexes is largely dictated by the nature of the interaction. In the gas phase, electrostatic and hydrogen bonding interactions are enhanced whereas hydrophobic interactions are partly or completely lost [62,63]. Full scan simple mass spectra with soft ionization conditions obtained in our study probe the composition of the solution and relative abundance of the species. We show that IB5•EgCG complexes are preserved and may involve more than one tannin. The preservation of supramolecular complexes formed by IB5 and EgGC, model of biological interest with regards to astringency, in the gas phase demonstrates the strength and the contribution of hydrogen bondings [62], in agreement with previous studies for the medium considered here [20]. As previously stated, the main function described for basic salivary PRPs, like IB5, is to be a scavenger which binds tannins [22,64-66]. Scavenger IUPs function by molecular recognition and often can bind several partners at the same time [67]. Observation of multiple stoichiometries ranging from 1:1 to 1:5 is consistent with IB5 alleged function. Basic PRPs have to bind more than one molecule of tannins to counteract their effects, especially when tannin concentration is rather high.

ESI-MS/MS analysis of IB5•EgCG complexes

Collision-induced dissociation is a process whereby precursor ions are activated upon collision with a non-reactive gas, such as argon or nitrogen in the present case. Each collision increases the overall vibrational energy of the molecule until there is enough internal energy accumulated within the molecule to overcome the bond energy, thus causing dissociation [68]. Different cares need to be considered. Proteins with the lowest charge states are reported to have a conformation close to that in solution [69]. Therefore ESI conditions must be controlled, but also the charge state of the complex ion selected. ESI MS/MS was performed on IB5-EgCG mixtures. Low-energy CID investigated the relative gas phase stability of the noncovalent IB5•EgCG complexes with 1:1 and 1:2 stoichiometries in the +6 charge state. The same experiments performed on +7 ions gave similar results but IB5 fragmentation was observed. During CID experiments, complex ions are accelerated with voltage offsets up to 50 V range into the gas-filled collision cell where their translational energy is partly converted into internal energy in a multi-collision process. As mentioned earlier, the kinetic energy of the precursor ion in the laboratory frame of reference is obtained by $E_{Lab} = z e V_c$, where z is the number of charges on the ion and e is the charge of an electron [44]. Therefore, the collision energy for the +6 precursor was increased from 0 eV to 300 eV. Collision-Induced total dissociation of both complexes (1:1 and 1:2) gave rise to the ions corresponding to $[IB5]^{6+}$ and a neutral EgCG (Fig. 5). It remains with the overall charge as expected, whereas EgCG is released as neutral species. Moreover, collision energy thresholds of total dissociation of $[IB5 \cdot EgCG]$ and $[IB5 \cdot 2 EgCG]$ complex ions are 180 eV and 210 eV respectively, showing the greater stability of the latter. Spectra obtained at CID lower collision energies allowed detection of both the remaining parent complex and $IB5^{6+}$ ions (Fig. 5a). In the case of 1:2 complexes, ions corresponding to the 1:1 complex are also detected. Precursor and product ions exhibit the same charge state (Fig. 5b). Thus, dissociation pathways for $[IB5 \cdot 2 EgCG]^{6+}$ result from neutral loss of EgCG :



Dissociation of the precursor complex ions leads to a decrease in stoichiometry down to the protein alone. Increasing energy on the parent ion releases the first tannin before the second one. Both tannins do not dissociate together from IB5. Nonetheless, this observation alone does not allow us to conclude whether tannins are bound on two different sites over IB5 or not.

Dissociation energies for $[IB5 \cdot xEgCG]$ ions ($x=1$ or 2) were determined to give a measurement of the relative EgCG binding energies in these complexes. Relative abundance of the precursor ion was calculated as $I_p / (I_p + \sum I_{frag})$, where I_p is the peak

intensity of the precursor ion, and $\sum I_{\text{frag}}$ is the sum of the peak intensities corresponding to all fragments [70]. This ratio is used to monitor the internal energy added to the complex ions to dissociate. The maximum energy that ions can acquire in the collision cell is the center-of-mass collision energy (E_{CoM}) of each collision summed over all. Laboratory frame collision energies (E_{Lab}) are converted to center-of-mass energies by using the equation

$$E_{\text{CoM}} = E_{\text{Lab}} \frac{m_g}{m_g + m_i} \quad (1)$$

where m_g and m_i are the mass of the target gas and of the complex ion respectively.

The number of collision (N_{col}) is determined by three parameters, the gas number density (n), the length of the collision cell (l) and the collision cross section (σ) of the selected ion :

$$N_{\text{col}} = n \sigma l \quad (2)$$

Finally, the internal energy E_{Int} can be calculated from equation

$$E_{\text{Int}} = E_{\text{CoM}} N_{\text{col}} \quad (3)$$

As a first order approximation, both complexes are expected to present very similar collision cross section because of the mere modification of their MM and their similar charge state [71]. Equation (2) shows that ions with similar cross section experience identical number of collisions, thus from equation (3) their E_{Int} can be directly compared. From equations (1) and (3), direct access to E_{Int} can be obtained via measurement of E_{CoM} for each species present in the spectrum. Moreover, previous works on proteins by Douglas group provided evidence of the efficiency of energy transfer in a quadrupole, showing that almost 90% of the center-of-mass-energy is converted to internal energy [72]. Thus E_{Int} is approximately given by the center of mass energy E_{CoM} [72,73]. For $[\text{IB5}\cdot\text{EgCG}]^{6+}$ and $[\text{IB5}\cdot 2 \text{ EgCG}]^{6+}$ complex ions, dissociation curves as a function of collision energy were plotted, and the half-wave collision energy (E_{50}) was defined. E_{50} in the laboratory frame is the collision energy required to disrupt 50 % of the parent complex ions. From the dissociation curves, E_{50} values have been established at 120 eV and 127.5 eV respectively (Table 2). From these measurements, E_{50} , in the center of mass (referred to as E_{CoM}) for $[\text{IB5}\cdot\text{EgCG}]^{6+}$ and $[\text{IB5}\cdot 2 \text{ EgCG}]^{6+}$ ions are obtained (Table 2)). The magnitude of that energy is related to the stability of the complex. Hence, we find that both $[\text{IB5}\cdot\text{EgCG}]^{6+}$ and $[\text{IB5}\cdot 2 \text{ EgCG}]^{6+}$ ions (Table 2) have the same E_{CoM} . This implies that the same energy is required to either dissociate one EgCG from the IB5•2 EgCG complex or to dissociate the 1:1 complex. This feature pointed out that EgCGs are involved within the complex with the same strength, indicating that they are bound on equivalent sites on IB5. This data supports the presence of at least two potential equivalent interaction sites on IB5, as schematically designed on figure 6. Moreover, the intrinsic

disorder of IB5 makes room for quite a large interaction surface, thus increasing the number of tannins that IB5 can bind [3].

CONCLUSIONS

Our experiments have demonstrated the ability of ESI-MS and ESI MS/MS to detect and characterize noncovalent protein-polyphenol complexes with a charge state distribution similar to that of the protein alone. Direct access to the complex stoichiometry has been reached through ESI-MS analysis, where a maximum of 5 EgCG for 1 protein has been observed. For the first time, this approach affords an opportunity to estimate the ability of polyphenols to form with a protein different stoichiometry soluble complexes and to evaluate their stability. Our results indicate that IB5 possesses several interaction sites to bind tannins. In the light of both observed stoichiometries and IB5 sequence, the repeated motif “KPQGPP” could be the main site for interaction. This study confirms that the stability of protein-tannin complexes increases with the number of bound tannins and presumably with the number of repeated amino-acid sequences [74]. Moreover, EgCG is likely to interact simultaneously through several sites of IB5. Another interesting consequence is that only PRP-tannin complexes made with full length PRP sequences resist degradation under conditions such as encountered in digestion [75]. This study is part of a general theme centered on the relevance of ESI-MS to probe protein-tannin solution-phase properties. Indeed, it clearly demonstrates that ESI-MS is well suited to detect associated polyphenols and will be used in further work to evaluate the ability of structurally distinguishable tannins to interact with peptides or proteins keeping in mind their astringency.

The authors thank H el ene Boze and Claire Bouchut (UMR 1083 SPO) for their help in protein production and purification and MS experiments.

We acknowledge synchrotron SOLEIL and thank all staff for assistance in using beamline DISCO.

AG thanks Applied Biosystems (Les Ulis, France) for the loan of the ion cooler guide.

This work is supported by grant 07-BLAN-0279 from the French Agence Nationale de la Recherche (A.N.R.).

REFERENCES

- [1] Tompa P (2002) *Trends Biochem Sci* 27:527-533
- [2] Dyson HJ, Wright PE (2005) *Nat Rev Mol Cell Biol* 6:197-208
- [3] Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z (2002) *Biochemistry* 41:6573-6582
- [4] Bennick A (1982) *Mol Cell Biochem* 45:83-99
- [5] Robbins CT, Hagerman AE, Austin PJ, Arthur CM, Hanley TA (1991) *Journal of Mammalogy* 72:480-486
- [6] Haslam E, Lilley TH, Cai Y, Martin R, Magnolato D (1989) *Planta Med* 55:v
- [7] McMurrough I, Madigan D, Kelly RJ, Smyth MR (1996) *J Am Soc Brew Chem* 54:141-148
- [8] McMurrough I, Hennigan GP, Loughrey MJ (1983) *J Inst Brew* 89:15-23
- [9] Waters EJ, Peng Z, Pocock KF, Williams PJ (1995) *Aust J Grape Wine Res* 1:86-93
- [10] Butler LG (1992) In: Hemingway RW, and Laks PE (eds.) *Antinutritional effects of condensed and hydrolyzable tannins*. Plenum Press, New York
- [11] Dangles O (2006) *Agro Food Ind Hi-Tec* 17:64-67
- [12] van Valen L (1973) *Evolutionary Theory* 1:1-30
- [13] Breslin PA, Gilmore M, Beauchamp GK, Green BG (1993) *Chem Senses* 18:405-417
- [14] Green BG (1993) *Acta Psychol* 84:119-125
- [15] Clifford MN (1997) In: Tomas-Barberan F, and Robins R (eds.) *Astringency*. Clarendon Press, Oxford
- [16] Noble A (1998) In: Waterhouse AL, and Eberler S (eds.) *Why do wines taste bitter and feel astringent ?* American Chemical Society, Washington, DC
- [17] Hagerman AE (1989) In: Hemingway RW, and Karchesy JJ (eds.) *Chemistry of tannin-protein complexation*. Plenum Press, New York, London
- [18] Bate-Smith EC (1954) *Food* 23:124-135
- [19] Haslam E (1996) *J Nat Prod* 59:205-215
- [20] Simon C, Barathieu K, Laguerre M, Schmitter JM, Fouquet E, Pianet I, Dufourc EJ (2003) *Biochemistry* 42:10385-10395
- [21] Pascal C, Poncet-Legrand C, Imberty A, Gautier C, Sarni-Manchado P, Cheynier V, Vernhet A (2007) *J Agric Food Chem* 55:4895-4901
- [22] Sarni-Manchado P, Canals-Bosch J, Mazerolles G, Cheynier V (2008) *J Agric Food Chem* 56:9563-9569
- [23] Porter LJ, Woodruffe J (1984) *Phytochemistry* 23:1255-1256
- [24] Okuda T, Mori K, Hatano T (1985) *Chem Pharm Bull* 33:1424-1433
- [25] Baxter NJ, Lilley TH, Haslam E, Williamson MP (1997) *Biochemistry* 36:5566-5577
- [26] Hatano T, Hemingway RW (1996) *J Chem Soc Chem Comm*:2537-2538
- [27] Luck G, Liao H, Murray NJ, Grimmer HR, Warminski EE, Williamson MP, Lilley TH, Haslam E (1994) *Phytochemistry* 37:357-371
- [28] Murray NJ, Williamson MP, Lilley TH, Haslam E (1994) *Eur J Biochem* 219:923-935
- [29] Vergé S, Richard T, Moreau S, Nurich A, Merillon J-M, Vercauteren J, Monti J-P (2002) *Biochim Biophys Acta* 1571:89-101
- [30] Sarni-Manchado P, Cheynier V (2002) *J Mass Spectrom* 37:609-616
- [31] Vergé S, Richard T, Moreau S, Richelme-David S, Vercauteren J, Promé J-C, Monti J-P (2002) *Tetrahedron Lett* 43:2363-2366
- [32] Chen Y, Hagerman AE (2004) *J Agric Food Chem* 52:4008-4011
- [33] Loo JA, Sannes-Lowery KA (1997) *Mass Spectrom Rev* 16:1-23
- [34] Pramanik BN, Bartner PL, Mirza UA, Liu YH, Ganguly AK (1998) *J Mass Spectrom* 33:911-920

- [35] Jorgensen TJD, Roepstorff P, Heck AJR (1998) *Anal Chem* 70:4427-4432
- [36] Kapur A, Beck JL, Brown SE, Dixon NE, Sheil MM (2002) *Protein Sci* 11:147-157
- [37] Bligh SWA, Haley T, Lowe PN (2003) *J Mol Recognit* 16:139-148
- [38] Gupta R, Kapur A, Beck JL, Sheil MM (2001) *Rapid Commun Mass Spectrom* 15:2472-2480
- [39] de Brouwer APM, Versluis C, Westerman J, Roelofsen B, Heck AJR, Wirtz KWA (2002) *Biochemistry* 41:8013-8018
- [40] Benesch JLP, Sobott F, Robinson CV (2003) *Anal Chem* 75:2208-2214
- [41] Rosu F, De Pauw E, Gabelica V (2008) *Biochimie* 90:1074-1087
- [42] Pascal C, Bigey F, Ratomahenina R, Boze H, Moulin G, Sarni-Manchado P (2006) *Protein Expr Purif* 47:524-532
- [43] Yin S, Xie Y, Loo JA (2008) *J Am Soc Mass Spectrom* 19:1199-1208
- [44] Haller I, Mirza UA, Chait BT (1996) *J Am Soc Mass Spectrom* 7:677-681
- [45] Robinson CV (2001) *J Am Soc Mass Spectrom* 12:126-126
- [46] Zhang J, Kashket S (1998) *Caries Res* 32:233-238
- [47] Jørgensen TJD, Delforge D, Remacle J, Bojesen G, Roepstorff P (1999) *Int J Mass Spectrom Ion Processes* 188:63-85
- [48] Wan KX, Gross ML, Shibue T (2000) *J Am Soc Mass Spectrom* 11:450-457
- [49] Akashi S, Osawa R, Nishimura Y (2005) *J Am Soc Mass Spectrom* 16:116-125
- [50] Harrison AG (1997) *Mass Spectrom Rev* 16:201-217
- [51] Bleiholder C, Suhai S, Paizs B (2006) *J Am Soc Mass Spectrom* 17:1275-1281
- [52] Yi SL, Boys BL, Brickenden A, Konermann L, Choy WY (2007) *Biochemistry* 46:13120-13130
- [53] Engel BJ, Pan P, Reid GE, Wells JM, McLuckey SA (2002) *Int J Mass Spectrom Ion Processes* 219:171-187
- [54] Breci LA, Tabb DL, III JRY, Wysocki VH (2003) *Anal Chem* 75:1963-1971
- [55] Vaisar T, Urban J (1996) *J Mass Spectrom* 31:1185-1187
- [56] Zhang X, Jai-nhuknan J, Cassady CJ (1997) *Int J Mass Spectrom Ion Processes* 171:135-145
- [57] Leymarie N, Berg EA, McComb ME, O'Connor PB, Grogan J, Oppenheim FG, Costello CE (2002) *Anal Chem* 74:4124-4132
- [58] Flanzky C (1998) *Oenologie - Fondements scientifiques et technologiques*. Lavoisier, Paris
- [59] Sharon M, Robinson CV (2007) *Annu Rev Biochem* 76:167-193
- [60] Hagerman AE, Rice ME, Richard NT (1998) *J Agric Food Chem* 46:2590-2595
- [61] Poncet-Legrand C, Gautier C, Cheynier V, Imberty A (2007) *J Agric Food Chem* 55:9235-9240
- [62] Robinson CV, Chung EW, Kragelund BB, Knudsen J, Aplin RT, Poulsen FM, Dobson CM (1996) *J Am Chem Soc* 118:8646-8653
- [63] Sobott FM, MG; Hernández, H; Robinson, CV. (2005) *Phil Trans R Soc A* 363:379-391
- [64] Mehansho H, Hagerman A, Clements S, Butler LG, Rogler JC, Carlson DM (1983) *Proc Natl Acad Sci USA* 80:3948-3952
- [65] Carlson DM (1993) *Crit Rev Oral Biol Med* 4:495-502
- [66] Sarni-Manchado P, Cheynier V, Moutounet M (1999) *J Agric Food Chem* 47:42-47
- [67] Tompa P (2003) *J Mol Struct (Theochem)*:361-371
- [68] Khalsa-Moyers G, McDonald WH (2006) *Brief Funct Genomic Proteomic* 5:98-111
- [69] Ruotolo BT, Robinson CV (2006) *Current Opinion in Chemical Biology* 10:402-408
- [70] Ham BM, Cole RB (2005) *Anal Chem* 77:4148-4159

- [71] Chen Y-LC, J. M.; Collings, B. A.; Konermann, L.; Douglas, D. J. (1998) *Rapid Commun Mass Spectrom* 12:1003-1010
- [72] Chen Y-L, Collings BA, Douglas DJ (1997) *J Am Soc Mass Spectrom* 8:681-687
- [73] Wells JM, McLuckey SA (2005) In: *Collision Induced Dissociation (CID) of Peptides and Proteins*. Academic Press
- [74] Charlton AJ, Baxter NJ, Lilley TH, Haslam E, McDonald CJ, Williamson MP (1996) *FEBS Lett* 382:289-292
- [75] Lu Y, Bennick A (1998) *Arch Oral Biol* 43:717-728

Table 1 Theoretical and measured mass (Da) of IB5 purified from heterologeous *Pichia pastoris* expressed PRPs

	IB5a	IB5b	IB5c
Theoretical average mass	6923.74	6642.43	6360.08
Measured average mass	6923.70	6642.63	6360.39

Table 2 Collision energy corresponding to 50% abundance of the precursor ions (E_{50}) in (E_{Lab}) in the laboratory frame and in the center of mass (E_{CoM}) (see text)

E_{50} (eV)	IB5•EgCG (1:1)	IB5•2EgCG (1:2)
E_{Lab}	120	127.5
E_{CoM}	0.45	0.45

Fig. 1 Structures of **a** epigallocatechin-3-O-gallate (EgCG) and **b** reserpine

Fig. 2 **a** Positive-ion ESI-mass spectrum of salivary PRP IB5 purified from PRPs produced by PRB4S transformant *Pichia pastoris* yeast. **b** Deduced primary sequences of the purified IB5x

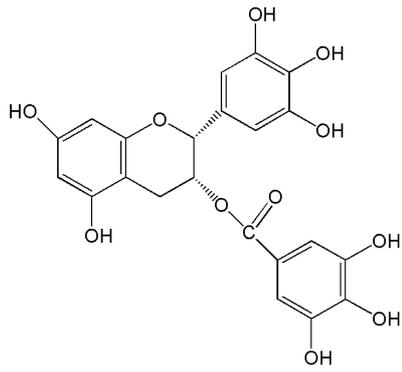
Fig. 3 **a** MS/MS spectrum of product ions obtained by CID of the 7+ charged ion of IB5, collision energy 280 eV. **b** Primary CID fragmentation reactions of IB5

Fig. 4 **a** Positive-ion ESI-mass spectrum obtained from mixture of IB5 and EgCG (5 μ M:50 μ M). **b** Expanded mass region of the 7+ complex ions

Fig. 5 **a** Plots of dissociation of [IB5 • EgCG]⁶⁺ ion and appearance of the [IB5]⁶⁺ ion. **b** Plots of dissociation of [IB5 • 2 EgCG]⁶⁺ ion, appearance and dissociation of [IB5 • EgCG]⁶⁺ ion and appearance of the [IB5]⁶⁺ ion. The y axis represents relative abundance of ions; Insets are CID MS/MS spectrum of the selected [IB5 • 2 EgCG]⁶⁺ complex ion (m/z 1307) for collision energy 120 eV and collision energy 240 eV

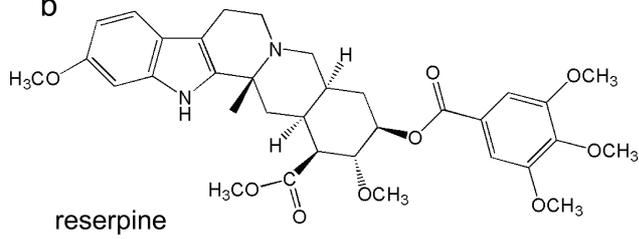
Fig. 6 Model scheme of interaction between IB5 and EgCG

a

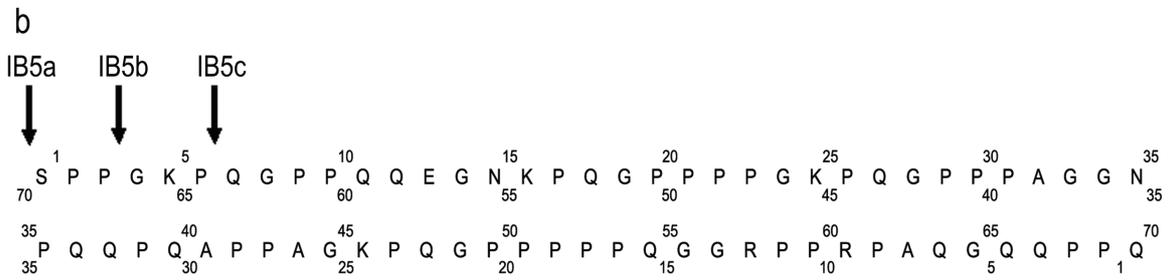
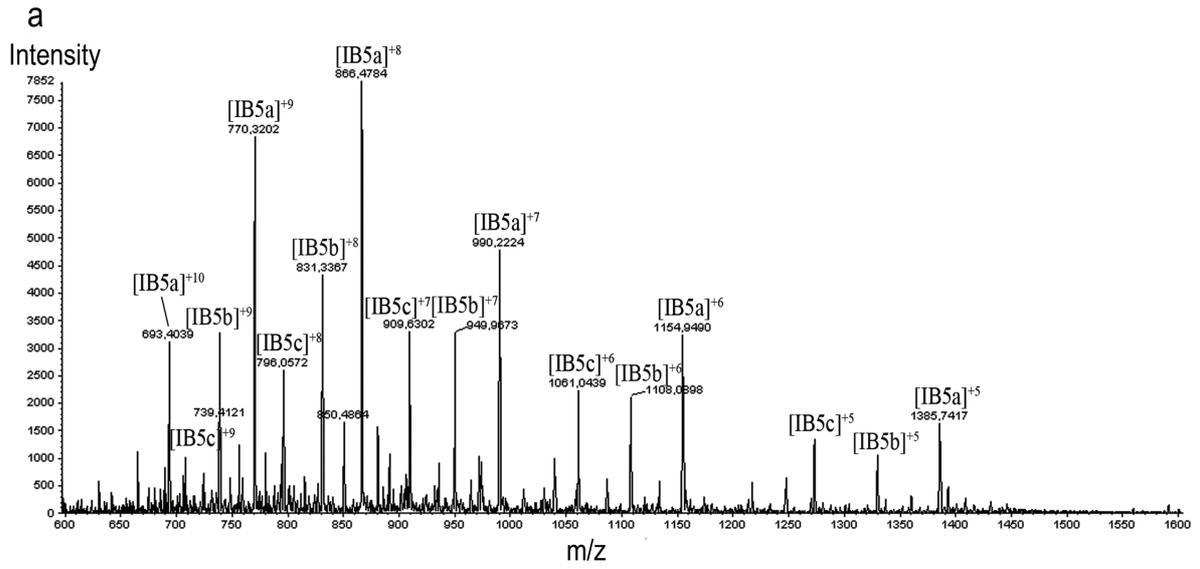


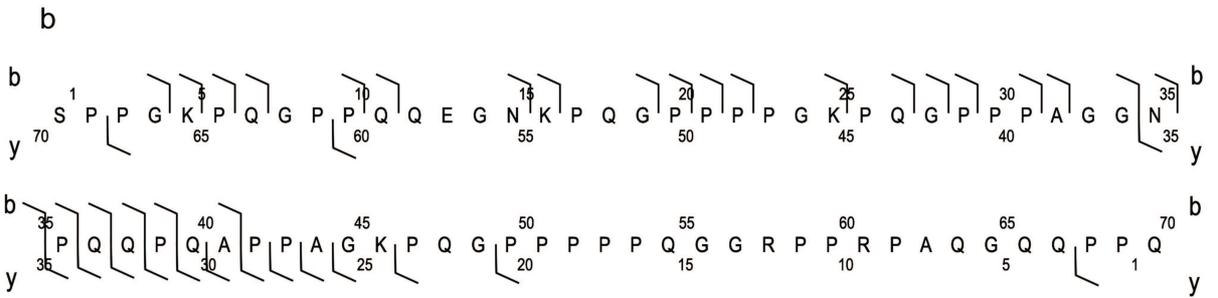
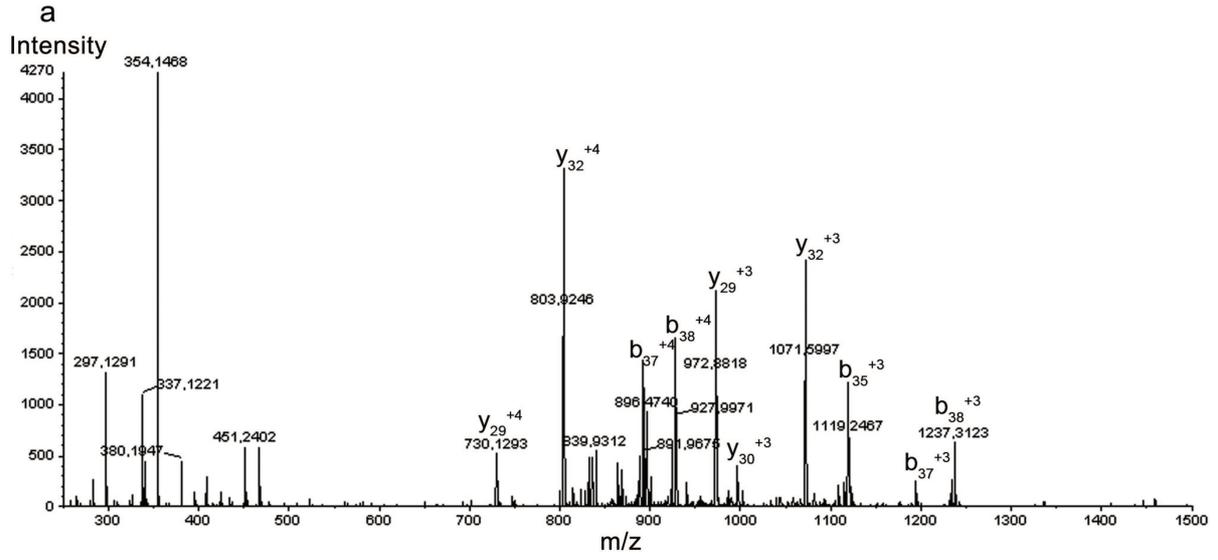
(-)-epigallocatechin-3-O-gallate (EgCG)

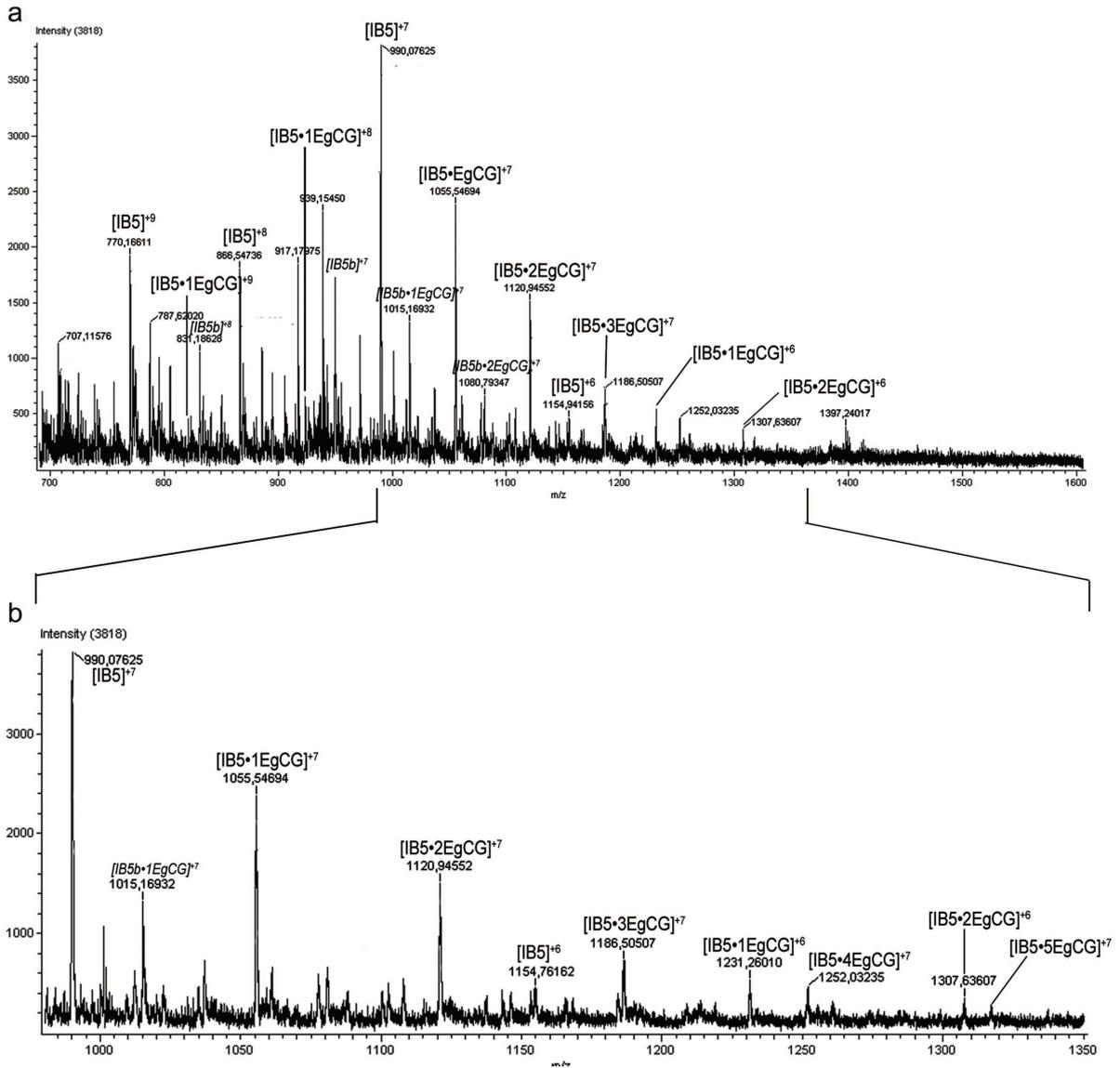
b

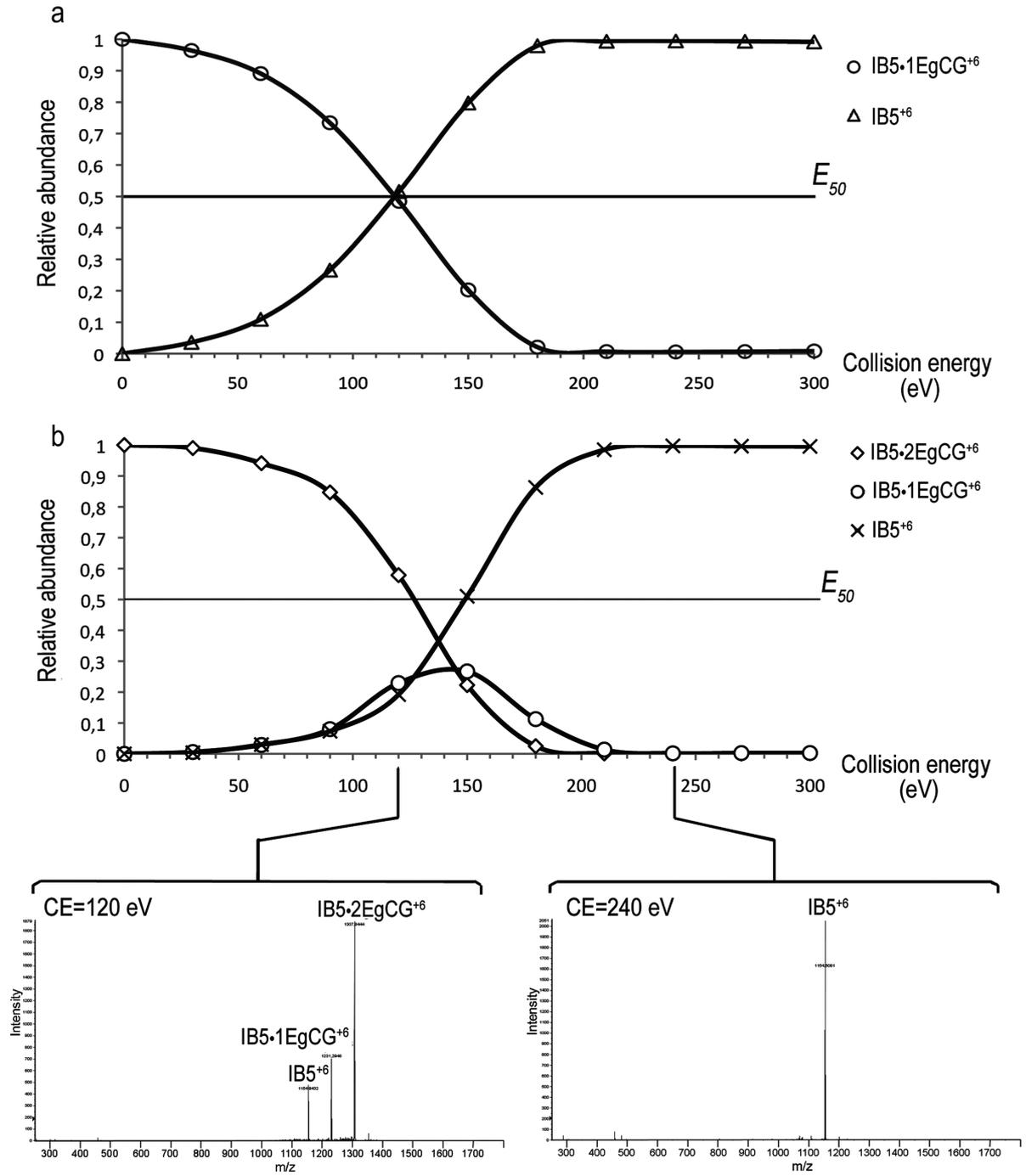


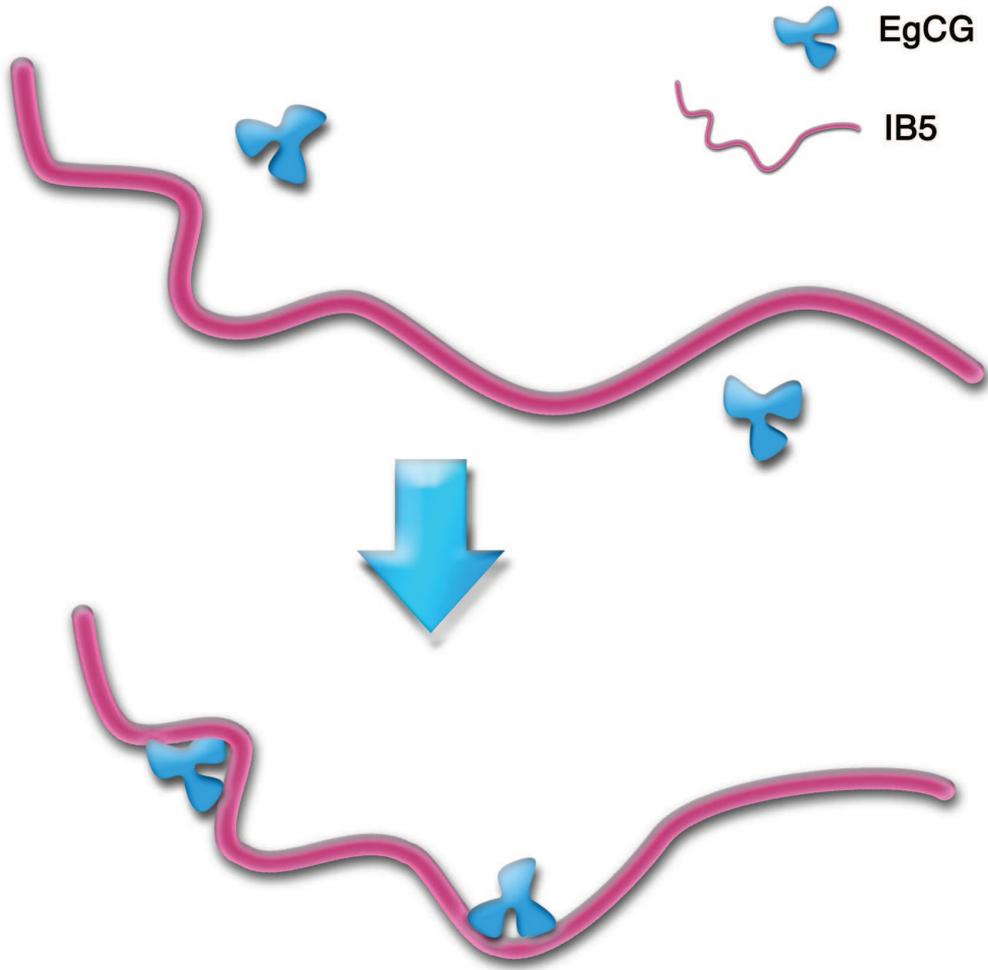
reserpine











La présence de complexes IB5•EgCG en solution a été démontrée pour un mélange IB5:EgCG de rapport molaire 1/10 (5 µM : 50 µM). Plusieurs stœchiométries allant de 1•1 à 1•5 ont été détectées.

Plusieurs questions sont soulevées par ces premiers résultats:

- des complexes ayant des stœchiométries supérieures peuvent-ils se former en solution?
- Peut-on déterminer une constante d'affinité correspondant à la formation de complexes IB5•EgCG ?

Pour tenter de répondre à ces deux interrogations, une expérience de titration a été réalisée. Différents mélanges IB5:EgCG ont été préparés en maintenant la concentration en IB5 égale à 5 µM et en faisant varier la concentration en EgCG entre 1 µM et 150 µM. Les rapports molaires qui ont été testés avec les concentrations en protéine et en tanin correspondantes sont récapitulés dans le tableau 10.

Tableau 10. Concentrations en IB5 et en EgCG testées durant les expériences de titration.

Rapport IB5/EgCG	5/1	2/1	1/1	1/2	1/5	1/10	1/15	1/20	1/30
[IB5] (µM)	5	5	5	5	5	5	5	5	5
[EgCG] (µM)	1	2,5	5	10	25	50	75	100	150

Chaque mélange a été étudié par ESI-MS à l'aide d'un Q-TOF (QStar, Applied Biosystem). Les spectres ESI-MS ont été acquis en mode positif.

L'ensemble des spectres présente des pics correspondant aux entités initiales du mélange (IB5 libre, EgCG libre) et aux complexes IB5•EgCG formés. Ces derniers présentent différentes stœchiométries qui augmentent avec la concentration en EgCG. La figure 65 illustre l'évolution des proportions de protéine libre et de protéine complexée en fonction de la concentration en EgCG. Cette figure montre une diminution de la proportion de protéine libre concomitante à une augmentation de celle de protéine complexée en présence de quantité croissante d'EgCG. La portion du spectre MS correspondant aux ions des complexes $[IB5 \cdot nEgCG]^{+7}$ de ce mélange au rapport molaire 1/30 est présentée sur la figure 66. L'état de charge étudié est celui qui est majoritaire. Des stœchiométries de complexes allant de 1•1 à 1•9 sont observées dans ces conditions d'interaction. Les complexes $[IB5 \cdot 2EgCG]^{7+}$ et $[IB5 \cdot 3EgCG]^{7+}$ prédominent.

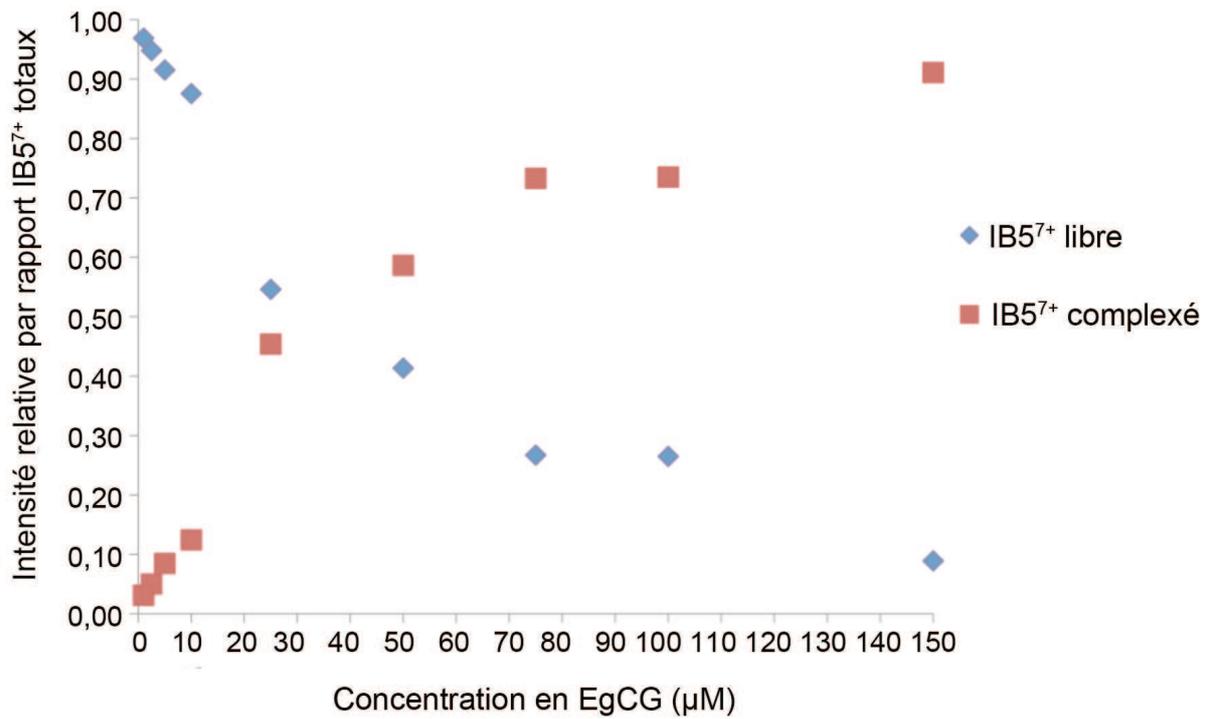


Figure 65. Evolution des proportions de protéine libre et de protéine complexée en fonction de la concentration en EgCG présent dans le mélange (ions dont état de charge est +7).

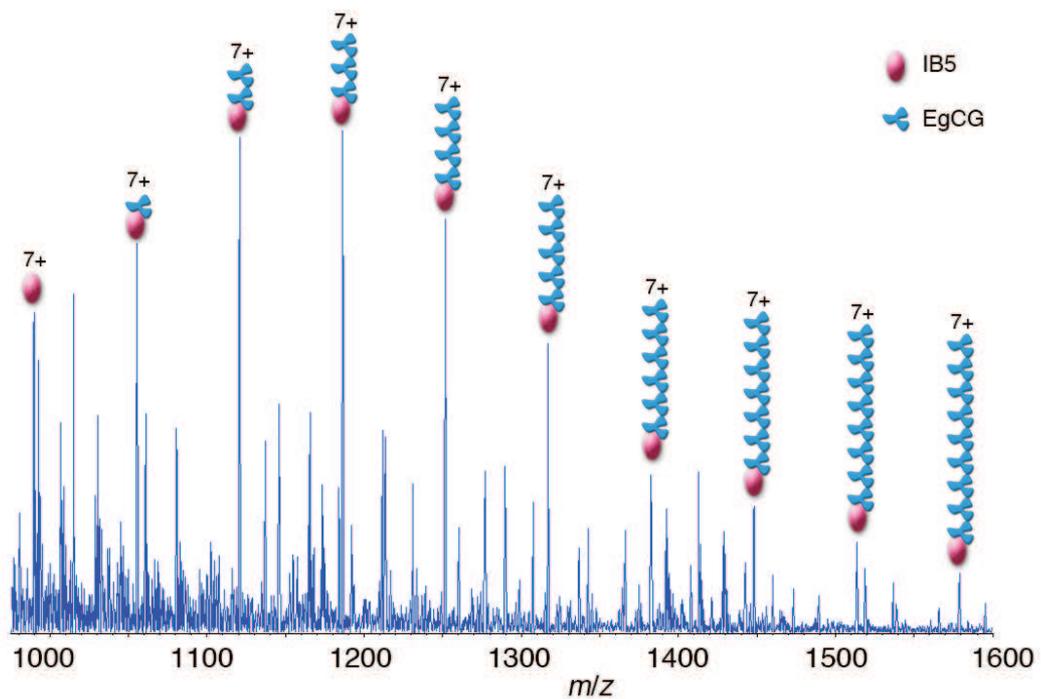


Figure 66. Spectre MS du mélange IB5:EgCG au rapport molaire 1/30 (5 μM /150 μM). Les pics correspondant à $[\text{IB5}]^{7+}$ et aux complexes $[\text{IB5}\cdot n\text{EgCG}]^{7+}$ sont légendés.

La présence de différentes stoechiométries au sein d'une même solution d'interaction associée à l'augmentation progressive de leur valeur avec la concentration en EgCG ne nous a pas permis d'atteindre un état d'équilibre. Ainsi, le calcul de la (des) constante(s) d'affinité n'a pas été possible.

II.3. Discussion

La MS supramoléculaire a connu un essor important depuis ses débuts dans les années 1990 : le nombre d'études publiées est maintenant devenu considérable (Ganem *et al.*, 1991 ; Ganem, 1993 ; Loo et Sannes-Lowery, 1997 ; Loo, 2000 ; Hernandez et Robinson, 2001 ; Cole, 2003 ; Ganem et Henion, 2003 ; Heck et van den Heuvel, 2004 ; Van den Heuvel et Heck, 2004).

Contrairement à d'autres techniques qui donnent une vue globale et moyenne d'un échantillon, la MS présente l'avantage unique de détecter individuellement toutes les espèces présentes en solution et permet notamment de démontrer l'existence de complexes de stoechiométries différentes dans le cas de systèmes à interactions multiples. Cette approche permet directement de mettre en évidence des effets de coopérativité négatifs ou positifs (Hu et Loo, 1995 ; Chazin et Veenstra, 1999 ; Rogniaux *et al.*, 2001). Ainsi, cette technique a confirmé la formation de complexes protéine•tanin solubles à travers l'étude du modèle d'interaction développé entre la protéine IB5 et le tanin EgCG. Elle a également mis en évidence l'existence de plusieurs stoechiométries au niveau des complexes formés. La singularité de l'interaction entre les tanins et les PRP a été démontrée par l'utilisation d'une molécule témoin, la réserpine (figure 1.b, publication n°1, page 162). Cette dernière ne forme pas de complexes avec IB5 malgré une structure proche de celle de l'EgCG.

Dans le cas d'interactions hôte-ligand dans lesquelles la protéine hôte présente un site de liaison structuré adapté à la forme du ligand (concept "clé-serrure" introduit par Emil Fisher (Fischer, 1894)), la validation des complexes supramoléculaires et de leurs stoechiométries consiste à enregistrer les spectres de masse de la solution hôte-ligand à un pH « physiologique » et à un pH acide (figure 67) (Sanglier *et al.*, 2005). La dénaturation du site de liaison à pH acide empêche l'interaction d'avoir lieu. Ainsi, la présence de complexes est observée sur les spectres de masse des solutions ayant un pH neutre alors que ces complexes ne sont pas observés dans le cas des solutions à pH acide.

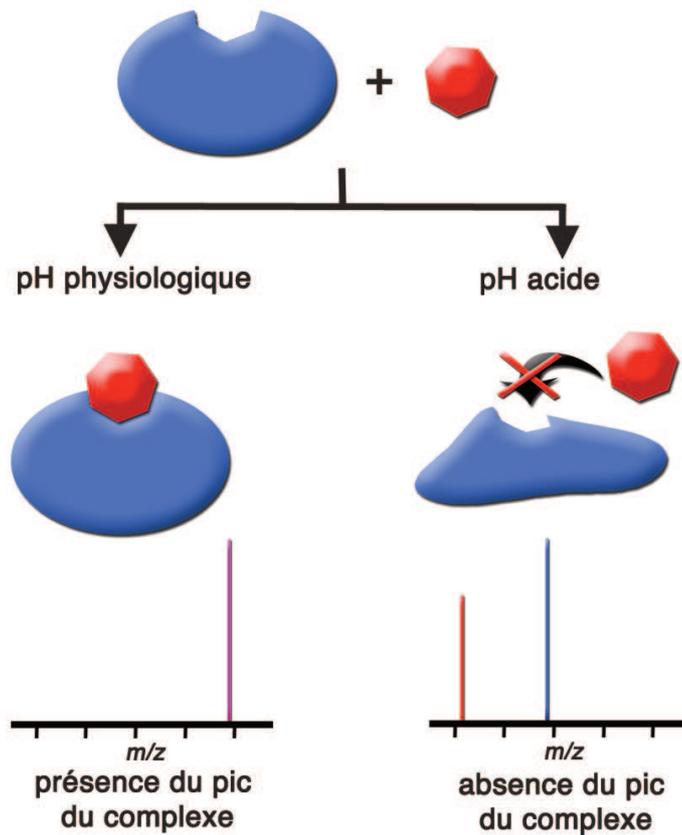


Figure 67. Principe du test de validation de l'observation des complexes non-covalents par spectrométrie de masse. La protéine est dénaturée à pH acide et son site de liaison n'est plus adapté à la forme du ligand.

Nous avons employé ce test, dans le but de montrer que la MS est capable de confirmer le caractère peu structuré des PRP et de montrer que le pH ne joue pas de rôle majeur dans l'établissement des interactions entre IB5 et l'EgCG. Les interactions IB5-EgCG ont été réalisées dans un milieu eau-éthanol acidifié pH 3,2, dans un milieu eau-éthanol pH 5,6 et dans un milieu eau-éthanol/acétate d'ammonium 20 mM pH 6,8. Les spectres acquis à partir de ces solutions présentent des ions correspondant aux mêmes complexes (figure 68). Le pH n'affectant pas les interactions PRP-tanin dans les milieux testés (pH 3,2, 5,5 et 6,8), le choix a été fait de maintenir la valeur de pH de notre milieu d'interaction à celle d'un milieu de type vin ou jus de fruit. De plus, la valeur du pH choisie de 3,2 est proche de celle du mélange du chyme alimentaire et des sucs gastriques dans l'estomac, au sein duquel agit une partie des enzymes digestives susceptibles d'être complexées par les tanins.

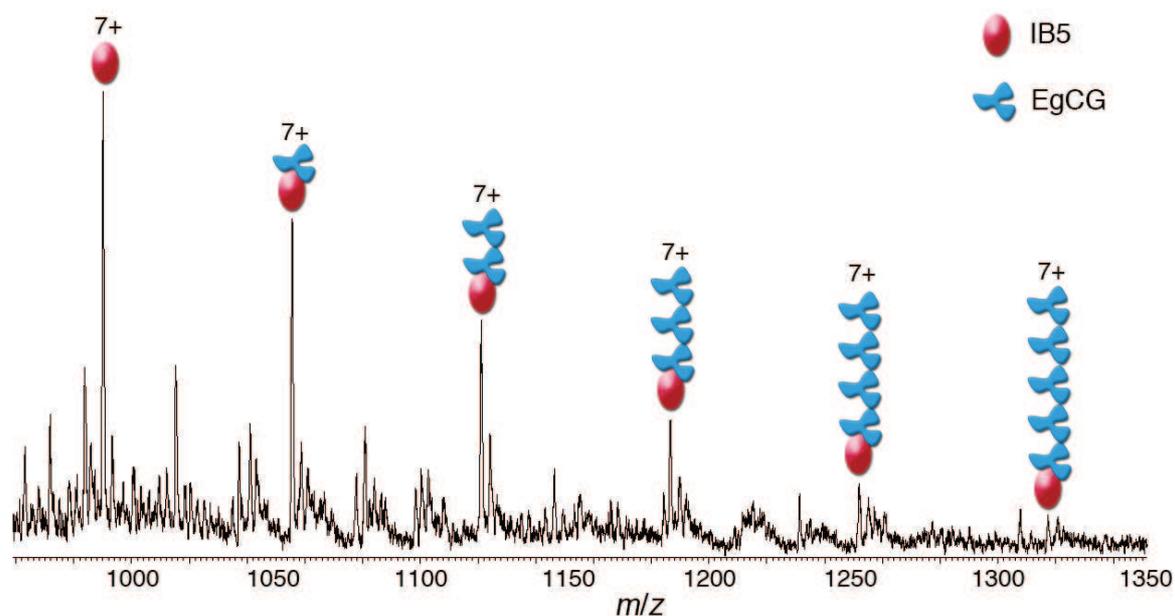


Figure 68. Région d'un spectre de masse d'une solution IB5-EgCG (5 μ M-50 μ M) eau-éthanol 88/12 (v/v) non acidifiée pH 5,7.

II.4. Conclusions et perspectives

La caractérisation de la protéine IB5 par spectrométrie de masse a montré que cette PID présente de hauts états de charge allant de 5+ à 11+ en mode d'ionisation électrospray positif, à des pH acides et neutres. Les états de charge observés, supérieurs au nombre d'acides aminés susceptibles de se charger positivement (acides aminés basiques : Arg, Lys et His), sont en accord avec les résultats montrant que la protéine IB5 présente une conformation étendue en solution (Boze *et al.*, 2010).

Cette conformation ouverte offre une grande surface potentielle d'interaction aux tanins. De hautes stœchiométries de complexes peuvent ainsi être atteintes. En effet, les stœchiométries observées vont de 1•1 à 1•5 si la concentration en EgCG est inférieure à 50 μ M et atteignent 1•9 pour une concentration de 150 μ M en EgCG. Le caractère « multi-récepteur » des PRP (pouvoir lier plusieurs tanins en même temps) est particulièrement intéressant pour piéger les tanins. La capacité de IB5 à lier plusieurs tanins simultanément étant démontrée, IB5 peut être classée dans la catégorie fonctionnelle des collecteurs décrite par Tompa (Tompa, 2003b ; Tompa, 2003a).

D'après Tompa (Tompa, 2003a), les séquences répétées des PRP constituent des unités fonctionnelles au niveau desquelles viennent interagir les tanins. D'autres observations l'ont conduit à proposer un schéma d'évolution des PID par expansion de leur

séquence répétée. Trois voies d'évolution seraient possibles (figure 69). Selon la voie de type I, les séquences répétées ne se diversifient pas et restent toutes fonctionnellement équivalentes. Dans la voie de type II, les régions répétées se différencient, soit à cause de leur position dans la chaîne, soit à cause de mutations d'acides aminés dans la séquence répétée. Cette différenciation conduit chaque unité fonctionnelle à assurer un rôle différent ou à lier un ligand différent. Dans l'évolution de type III, l'augmentation du nombre de séquences répétées crée une nouvelle fonction.

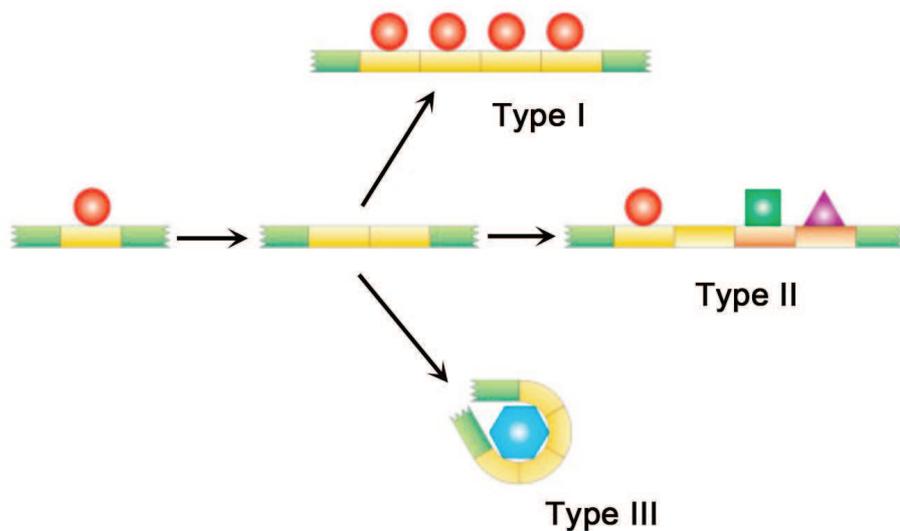


Figure 69. Voies évolutives suivies par expansion des séquences répétées chez les PID, tiré de (Tompa, 2003a).

Les PRP auraient évolué selon la voie de type I. Elles auraient ainsi acquis plusieurs unités fonctionnelles adaptées à la liaison des tanins. En accord avec ces théories, l'étude de la stabilité des complexes de stœchiométries 1•1 et 1•2 que nous avons réalisée semble indiquer que les tanins se fixent sur des régions différentes de la protéine.

Cette première partie a concerné l'étude d'un seul tanin modèle, l'EgCG. Cependant, les tanins présentent une diversité structurale remarquable. La suite de ce travail a porté sur les interactions d'IB5 avec quatre autres tanins modèles. En effet, il est important de s'assurer que cette interaction n'est pas spécifique à un seul tanin et d'envisager l'influence de leur structure.

Chapitre III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

III.1. Introduction

La diversité structurale des tanins est rapportée comme intervenant dans leur capacité de précipitation des protéines et leur pouvoir astringent. Afin de vérifier que l'influence de la structure des flavanols sur l'interaction initiale avec les protéines débute dès la formation des complexes solubles, nous avons étudié les interactions entre IB5 et cinq tanins modèles. L'approche MS validée lors de l'étude des complexes IB5•EgCG a été appliquée à l'étude des interactions s'établissant au sein de ces solutions modèles. Ces molécules sont choisies car elles permettent d'appréhender les principales caractéristiques structurales des tanins que sont :

- le degré de polymérisation,
- l'estérification de l'hydroxyle en position C3 par l'acide gallique (cf. figure 43.a),
- le degré d'hydroxylation du noyau B (cf. figure 43.a).

III.2 Etude des interactions IB5-tanin par ESI-MS et ESI-MS/MS CID

Les résultats obtenus ont fait l'objet d'une publication dans la revue internationale à comité de lecture *Analytical and Bioanalytical Chemistry* (publication n°2).

Publication n°2 : Ability of a Salivary Intrinsically Unstructured Protein to Bind Different Tannin Targets Revealed By Mass Spectrometry

Francis Canon†, Alexandre Giuliani‡§, Franck Paté† and Pascale Sarni-Manchado†*

†INRA, UMR1083 Sciences Pour l'Oenologie, 2, place Viala, F-34060 Montpellier, France

sarni@supagro.inra.fr, Phone: 33 4 99 61 22 98, Fax: 33 4 99 61 26 83.

‡ *DISCO beamline, Synchrotron Soleil, l'Orme des Merisiers, 91192 Gif sur Yvette France.*

§ *CEPIA, INRA, BP 71627, F-44316 Nantes Cedex 3 France.*

*Reprint request to Dr. Pascale Sarni-Manchado, INRA, UMR1083 SPO, INRA, Bât 28, 2 place Viala F-34060 Montpellier France, *sarni@supagro.inra.fr, Phone: 33 4 99 61 22 98, Fax: 33 4 99 61 26 83.*

Keywords: polyphenol, noncovalent interaction, Proline Rich Protein, saliva, astringency, Intrinsically Unstructured Protein.

Abbreviations:

- CID, Collision-Induced Dissociation;
- EgC, epigallocatechin;
- ECG, epicatechin gallate;
- EgCG, epigallocatechin gallate;
- B2 3'OG, B2 3'O-gallate;
- ESI, electrospray ionization;
- IUP, Intrinsically Unstructured Protein;
- MS, mass spectrometry;
- MS/MS, tandem mass spectrometry;
- PRP, Proline Rich Protein;
- b*PRP, basic Proline Rich Protein;
- a*PRP, acidic Proline Rich Protein;
- Q-TOF, quadrupole/time-of-flight;
- T, tannin;
- UGT, ungalloylated tannin;

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

GT, galloylated tannin;

M, monomer;

GM, galloylated monomer;

D, dimer;

GD, galloylated dimer;

DP, degree of polymerisation;

ABSTRACT

Astringency is thought to result from the interaction between salivary proline-rich proteins (PRP) that belong to the Intrinsically Unstructured Protein group (IUP), and tannins, which are phenolic compounds. IUPs have the ability to bind several and/or different targets. In the same time, tannins have different chemical features reported to contribute to the sensation of astringency. The ability of both electrospray ionization mass spectrometry (ESI-MS) and tandem mass spectrometry (MS/MS) to investigate the noncovalent interaction occurring between a human salivary PRP, IB5, and a model tannin, epigallocatechin 3-O-gallate (EgCG), has been reported. Herein, we extend this method to study the effect of tannin chemical features on their interaction with IB5. We used five model tannins, epigallocatechin (EgC), epicatechin 3-O-gallate (ECG), epigallocatechin 3-O-gallate (EgCG), procyanidin dimer B2 and B2 3'-O-gallate, which cover the main tannin chemical features: presence of a gallate moiety (galloylation), the degree of polymerization, and the degree of B ring hydroxylation. We show the ability of IB5 to bind these tannins. We report differences in stoichiometries and in stability of the IB5•tannin complexes. These results demonstrate the main role of hydroxyl groups in these interactions and show the involvement of hydrogen bonds. Finally, these results are in line with sensory analysis, pointing out that the chain length and the level of galloylation are the main factors affecting astringency perception [1].

INTRODUCTION

Over the last decade, the sequence-to-structure-to-function paradigm based on the view that proteins need to have a well-defined three-dimensional structure for their function, has been re-assessed. Numerous proteins do have a well-defined function that requires intrinsic disorder, and the term of Intrinsically Unstructured Protein (IUP) has been proposed to describe this group [2]. IUPs appear to be rather common in living organisms, especially in higher eukaryotes: primary sequence analyses indicate that about 25% of full-length mammal proteins belongs to this class [3]. The intrinsic lack of structure can confer functional advantages such as the ability to bind several and/or different ligands [2,4,5]. This ability often leads to mixtures of complexes with heterogeneous composition. While techniques are emerging for the characterization of flexible proteins [6,7], the study of complexes of IUP is highly challenging and requires new approaches. Electrospray ionisation (ESI) has demonstrated its efficiency to probe weak associations of molecules in solution such as receptor-ligand and protein-nucleic acid interactions [8,9]. ESI-MS offers speed and sensitivity in monitoring components of mixtures and allows the determination of supramolecular edifices stoichiometry. Determination of relative binding affinities of noncovalent complexes is now well established [10-12]. Tandem mass spectrometry (MS/MS) provides access to non-covalent interaction strengths, and qualitative information on the complex structure [13,14]. Herein, the ability of a model IUP, namely IB5, to form supramolecular edifices with different ligands is explored through MS and MS/MS approaches.

The IUP IB5 is a human salivary protein belonging to the class of proline-rich proteins (PRPs). For numerous mammals, PRPs are the most prevalent group of proteins in the saliva; they may constitute 70% of all proteins in human parotid saliva [15]. The salivary proteins play various roles (lubrication, digestion, taste, protection ...) as saliva is the first fluid which interacts with food constituents. Salivary PRPs are classically divided in two groups, acidic PRPs (*a*PRPs) and basic PRPs (*b*PRPs) [16]. *a*PRPs act primarily in maintaining oral homeostasis, whereas the only known function of *b*PRPs is to bind and to scavenge tannins [17-19]. Moreover, in rats and mice, synthesis of salivary PRPs is stimulated by a tannin rich diet [20,21]. PRPs affinity for tannins is estimated to be 5-80 and 1000 times higher than that of bovine serum albumine and lysozyme, respectively [22-24]. PRP involvement in adaptation to tannin rich diet either as astringency mediator or/and as scavenger, is related to their ability to establish non-covalent interactions with tannins [16,25].

Tannins are phenolic compounds ubiquitous in plant and plant-based foods and beverages. From a biological point of view, tannins act in plant defence mechanisms as molecules that protect against herbivores [26]. This role has been attributed to their ability to interact with proteins, to precipitate them and to inhibit gastrointestinal enzymes, thereby reducing the digestibility of plant proteins [27]. Proanthocyanidins are the main tannins in plant based-food (*e.g.* fruit, cocoa) and beverages (*e.g.* wine, tea, beer, cider,...). These molecules have attracted considerable interest because of their quantitative importance and major contribution to organoleptic and biological properties. The chemical characteristics of proanthocyanidins are believed to affect their organoleptic properties, in particular their degree of polymerisation, their level of B ring hydroxylation and their level of galloylation [1]. Numerous studies on the impact of these chemical characteristics over the protein•tannin interaction, were realized on poly L-proline [28] or peptides [29]. Since the affinity for tannins differs among proteins [23], PRPs [22] and with the length of PRP [30], it is essential to examine the effect of tannin chemical features on tannin interaction with a full *b*PRP.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

Our previous work on the interaction between IB5, a 70 a.a. long human basic salivary proline-rich protein (*bPRP*), and EgCG, a model tannin, revealed the presence in solution of IB5•EgCG supramolecular edifices with several stoichiometries [31]. The gas phase stability of complex ions corresponding to two stoichiometries was investigated, and demonstrated the presence of several interaction sites on IB5. This approach affords the opportunity to estimate the ability of polyphenols to form different stoichiometry soluble complexes with a protein and to evaluate their stability. Since PRP-tannin interactions are thought to be involved in astringency [32], it is important to understand how the tannin chemical features have an effect on the stoichiometry, and on the strength of these interactions. To address these questions, we investigated the interactions between IB5, and five model tannins (T): three are monomers (M) epigallocatechin (EgC), epicatechin gallate (ECG), epigallocatechin gallate (EgCG), and two are dimers (D): B2 and B2 3'-O-gallate (B2 3'OG), (Figure 1a). They cover the main chemical characteristics of tannins: number of units (DP), hydroxylation pattern of the B ring and presence of a galloylated moiety on position 3.

This paper describes the use of the ESI-MS CID approach developed previously, which has proven its efficiency to probe protein-tannin solution-state properties, to gain deeper understanding on the relation between tannin structure and astringency perception.

MATERIALS AND METHODS

Samples

Epigallocatechin (EgC), epicatechin gallate (ECG), epigallocatechin gallate (EgCG) and reserpine (Figure 1a and 1b) were purchased from Sigma (Sigma Chemical Co., Poole, Dorset, U.K.). B2 and B2 3'-O-gallate were purified as previously described [33].

The human salivary proline rich protein, IB5, was produced by use of the yeast *Pischia pastoris* as a host organism and purified as previously described [34].

Tannin and protein stock solutions were prepared in the following medium: water/ethanol, 88:12 (v/v) acidified to pH 3.2 with acetic acid. They were immediately frozen (-20°C) to prevent protein proteolysis and tannin oxidation. Prior to use, they were diluted at room temperature to the desired concentrations.

Protein-Tannin interactions

For IB5-tannin interactions, the protein concentration was kept constant at 5µM in the water/ethanol medium described above. Tannins and protein solutions were mixed extemporaneously at room temperature (regulated at 24°C) to obtain a protein/polyphenol molar ratio of 1:10.

The control solution, addressing the issue of nonspecific aggregates that could form during electrospray ionization, was prepared by mixing IB5 and reserpine at the same 1:10 final ratio as the IB5-T mixtures.

Electrospray ionization mass spectrometry and tandem mass spectrometry

Mass spectrometry and tandem mass spectrometry experiments were performed on a hybrid quadrupole time-of-flight (Q-TOF) Qstar Pulsar i mass spectrometer (Applied Biosystems, Forster City, CA), providing a resolving power of 10000 at m/z 1080.8. The interface was fitted with the ion cooler guide to preserve the noncovalent complexes upon transfer into the gas phase. The Ion cooler guide is mounted after the skimmer around Q0 and increases the local pressure to greater than 30 mTorr, which greatly improves preservation of weakly-bound complexes [35]. The Q-TOF was operated with the Turbospray (electrospray ion source). The sample was infused in the mass spectrometer by means of a syringe infusion pump at $7 \mu\text{L}\cdot\text{min}^{-1}$ flow rate. The source voltage was set to 5800 V in positive ion mode, the declustering potential at 47 V, the focusing potential at 209 V, the declustering potential 2 at 17 V, the ion source gas 1 set at 22, the ion source gas 2 at 0, the curtain gas at 10 and the capillary was not heated up. Nitrogen from a generator, containing less than 6 ppm oxygen was used both as nebulizing and desolvation gas. During MS/MS experiments, nitrogen was used as a collision gas in the collision cell (Linac) with CAD gas setting of 5. The collision-set energy voltage V_c was increased from 0 to 50 V. The kinetic energy of the precursor ion in the laboratory frame of reference is $E_{Lab} = z \times e \times V_c$, where z is the number of charges on the ion and e is the charge of an electron [36].

Analysis of ESI MS and MS/MS experiments and calculations

ESI-MS and collision MS/MS data were analysed with Analyst QS [37] and MagTran [38] softwares.

A procedure adapted from Jørgensen et al. [39] and Wan et al. [40] was used to establish the dissociation curves. The 50% dissociation energy is defined as the energy giving a 50% loss of precursor complex ions. Based on the data of Akashi et al. [41], one of the multiprotonated molecules of the IB5•1tannin complex (M), $[\text{M}+6\text{H}]^{6+}$ was selected and submitted to CID.

RESULTS AND DISCUSSION

ESI-MS analysis of IB5•tannin complexes

The mass spectrum obtained by electrospraying the protein solution displayed a series of peaks corresponding to protonated isoforms of IB5 (a 6923.70, b 6642.63 and c 6360.39 Da) with charge states ranging from +5 to +10 as previously observed [31].

The concentrations chosen for protein and tannin reflect biological conditions. Concentrations of condensed tannins in red wine are approximately $1\text{-}2 \text{ g}\cdot\text{L}^{-1}$, corresponding to a mean molar concentration of 0.5-1mM. The concentration of proteins in saliva is approximately $1 \text{ mg}\cdot\text{mL}^{-1}$ circa 83 μM . Polyphenols are present in large excess compared to proteins both in wine and in the mouth after sipping wine, making the protein/polyphenol ratio always in the benefit of polyphenols. On this basis, to investigate soluble protein-polyphenol complexes we use a protein/polyphenol ratio of 1:10 that gives rise neither to the formation of precipitate nor to cloudiness.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

The interaction medium mimics the mouth condition during red wine consumption. As wine is more abundant than saliva, the buffering ability of wine tartaric acid leads to an acidic pH. The pH value of 3.2 used during these experiment is in the range of wine pH values, which are comprised between 2.8 and 3.8 [42,43]. Moreover, the function of PRPs is to bind and scavenge tannins to protect digestive enzymes, for example in the stomach where pH is approximately 1.5 and PRP•tannin complexes are stable [44]. Thus, the pH used in this study reflects the biological pH of PRP-tannin interaction medium.

Premixed interaction medium with a 1:10 IB5/tannin molar ratio was submitted to ESI mass spectrometry. MS analysis reveals IB5•nT complex ion peaks for each mixture with several stoichiometries (n being the number of T molecules involved in the complexes, Figure 2). Complexes involving galloylated tannins (IB5•nGT) show the highest stoichiometries and also the highest relative intensities. The presence of a galloylated moiety favours the interaction with IB5. Closer analysis of MS spectra shows that IB5•nUGT complexes (UnGalloylated Tannins) present similar stoichiometries, while stoichiometries of IB5•nGM complexes (Galloylated Monomers) are higher than those of IB5•nGD (Galloylated Dimers). This result may be related to the bulkier size of the dimers, which may restrict the number of sites that can be reached simultaneously on IB5. To address the issue of nonspecific aggregates, which could form during electrospray ionization, ESI mass spectra of mixtures of IB5 and reserpine have been recorded. Reserpine presents a chemical structure and a molecular mass close to those of our model tannins (Figure 1). These spectra did not give rise to complex peaks. If the interaction was not driven by polyphenols, ESI MS should have generated complexes between IB5 and reserpine. The absence of IB5•reserpine peak, and the presence of IB5•T peaks, demonstrated that the observed noncovalent complexes were not some ESI artefacts, sustaining the specificity of the protein-tannin interaction. The observation of IB5•T complexes demonstrates the ability of IB5 to bind different tannins. This result is in agreement with the characteristics of IUPs, which have often enough malleability to bind different partners [45,46]. Observation of multiple and high stoichiometries is consistent with IB5 alleged function. The main function described for basic salivary PRPs, like IB5, is to be a scavenger that binds tannins [17-19,47]. The extended structure of PRPs [48] allows them to bind more than one molecule of tannins to counteract their effects, especially when tannin concentration is rather high. These results confirm that PRP belong to the IUP functional class of scavenger [4,49]. Scavenger IUPs work by molecular recognition and, can bind several partners at the same time [2,4]. However, the stoichiometry differences observed among the complexes involving the different tannins reveal an impact of the tannin chemical features on the interaction. To pursue the investigation on this effect, the stability of IB5•1T complexes has been evaluated by MS/MS.

ESI-MS/MS analysis of IB5•1tannin complexes

The relative stability of IB5•1T complexes was investigated by monitoring $[IB5•1T]^{+6}$ ion dissociation by CID MS/MS experiments. Collision-induced dissociation is a multi-collisions process whereby precursor ions are activated upon collision with a non-reactive gas, nitrogen in the present case. Each collision increases the overall vibrational energy of the ion until there is enough internal energy accumulated within the species to overcome the bond dissociation energy, thus causing fragmentation [50]. In the present study, low-energy CID investigated the relative gas phase stability of the noncovalent complexes $[IB5•1T]^{+6}$.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

Ions were injected into the collision cell with increasing energies to dissociate them through collisions with nitrogen. Collision-Induced total dissociation of all IB5•1T complexes gave rise to the ions corresponding to [IB5]⁶⁺ and a neutral tannin.



Dissociation curves for each parent complex ion were derived and correlated the relative abundance of the parent ion to the collision energy (Figure 3).

The maximum energy that ions can acquire in the collision cell is the center-of-mass collision energy (E_{CoM}) of each collision summed over all. Laboratory frame collision energies (E_{Lab}) are converted to center-of-mass energies by using the equation

$$E_{CoM} = E_{Lab} \frac{m_g}{m_g + m_i} \quad (1)$$

where m_g and m_i are the mass of the target gas and of the complex ion respectively.

The number of collision (N_{col}) is determined by three parameters, the gas number density (n), the length of the collision cell (l) and the collision cross section (σ) of the selected ion :

$$N_{col} = n \sigma l \quad (2)$$

Finally, the internal energy E_{Int} can be calculated from equation

$$E_{Int} = E_{CoM} N_{col} \quad (3)$$

Equation (2) shows that ions with similar cross section experience the same number of collisions, on average. Thus from equation (3), direct comparison of E_{Int} can be realised via comparison of E_{CoM} for complexes with a similar collision cross section. Given the slight modification of the molecular mass between [IB5•1T]⁺⁶ complexes and their similar charge state [31,51], we expect complex ions to present similar collision cross sections. Consequently, the centre of mass energy of the complex (E_{CoM}) is likely to fit the scaling of the internal energy transferred to the complex upon dissociation.

The dissociation curves show similar sigmoid shape for each [IB5•1T]⁺⁶, indicating that the interaction of IB5 with each tannin involves a similar mechanism independently of the tannin structure. During a first phase, dissociation curves depict an upper plateau where dissociation is minimal. Ions accumulate internal energy. Then the curves bend downward into a log-linear dissociation region. Finally, the curves bend into a plateau toward the total complex dissociation. E_{50} in the laboratory frame is the collision energy required to disrupt 50 % of the parent complex ions. E_{50} are determined graphically from the dissociation curves (Figure 3, table 1). We used them to calculate the corresponding E_{CoM} (table 1) reflecting complex stability[31,51]. The higher the E_{CoM} , the stronger the interaction. Thus, classification of the complexes as a function of their stability can be drawn as :



It is worth noting that IB5•dimer complexes are more stable than those formed with monomers. Moreover, IB5•GT are more stable than IB5•UGT for a same degree of polymerization. The extra OH group on the B-ring leads to a higher E_{CoM} (EgCG vs ECG). Through analysis of this set of tannins, we have shown the impact of

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

chemical features on the stability of the complexes, and on the ability of IB5 to adapt to ligands, as previously observed for other IUPs [52]. Hence, we are able to classify these chemical characteristics as a function of their impact on the interaction strength with IB5:

$$\text{B ring hydroxylation degree} < \text{galloylation} < \text{DP}$$

Figure 4 shows that a linear relationship exists between the number of tannin hydroxyl groups and the E_{CoM} of IB5•tannin complexes. The more hydroxyl groups on tannins, the more stable the interactions with IB5. This indicates that interactions between IB5 and tannins mainly involve hydrogen bonds in the gas phase as stated earlier [31]. Electrostatic and hydrogen bonding interactions are enhanced in the gas phase whereas hydrophobic interactions are partly or completely lost [53,54]. Therefore, the involvement of hydrophobic effect may not be excluded as it may occur in solution. Our results are in agreement with those of Simon [55], who has shown the importance of hydrogen bonding in the interaction between a Proline Rich peptide (PRp: 14 a.a.) and tannins, with different techniques such as NMR, MS, CD and molecular modelling.

Higher protein affinity has been reported for prodelphinidins (tannins with EgC units) compared to procyanidins (tannins mainly composed with EC and C units) [56]. In agreement with this finding, epigallocatechin gallate exhibited a stonger interaction than epicatechin gallate with immobilized salivary proteins [57]. The same trend was found when studying their complexation with the human salivary PRP IB5 by mass spectrometry but the reverse was found when calculating their binding constants with poly L-proline [58]. This difference can be explained by different affinities for tannins as observed between PRPs from various species [22] and by the random coil nature of *b*PRPs [48], which probably gives them enough malleability to be adaptable for different ligands in contrast with the polyproline helix conformation of poly L-proline.

Moreover, the astringency sensory analysis of different tannins has shown that DP and galloylation are major factors driving astringency perception [1]. This is parallel with our finding here, for which such factors are involved in the interaction strength with salivary proteins.

CONCLUSION

It is often asked why wine astringency sensation and intensity are dependent upon the wine itself and the wine taster. On one hand, wine molecules involved in astringency are tannins that gather a very large family of compounds with numerous and peculiar chemical structures. On the other hand, taster saliva holds a high level of proteins and particularly the so-called *b*PRPs whose only reported function is to bind tannins. Astringency is supposed to be related to their interactions and the stability of the arising complexes may have an impact on this perception. The stability of the complexes formed between tannins and *b*PRPs comes from the strength of the interaction forces. At this point, it makes compulsory to watch for them with regards to tannin chemical features. In this work, we used MS technologies benefiting from MS instrumentation developments to detect and characterize soluble noncovalent complexes involving a model *b*PRP, IB5, and tannins. The results show the effect of the tannin structure and highlight the involvement of hydrogen bounding on PRP-tannin interactions. The conformational rearrangement of PRP may favor the establishment of more hydrogen bonds. These results

demonstrate the ability of PRP to universally bind tannins that have a variety of shapes and sizes and are in agreement with the tannin scavenging function of IB5. Moreover, sensory analyses and MS studies have revealed that DP and galloylation are major factors driving tannin astringency perception [1] and PRP-tannin interaction respectively. This parallel is of particular interest because PRP-tannin interactions are thought to contribute to astringency mechanisms.

Acknowledgements

The authors thank Dr. Véronique Cheynier for helpful scientific discussions, Thérèse Marlin for protein purification, Jean-Paul Mazauric for tannin purification and Emmanuelle Meudec for mass spectrometry assistance. Francis Canon was supported by a grant of French Ministry of Research. This work is supported by grant 07-BLAN-0279 from the French Agence Nationale de la Recherche (A.N.R.). We acknowledge synchrotron SOLEIL and thank all staff for assistance in using beamline DISCO. AG thanks ABSciex (Les Ulis, France) for the loan of the IonCooler Guide.

REFERENCES

- [1] Vidal S, Francis L, Guyot S, Marnet N, Kwiatkowski M, Gawel R, Cheynier V, Waters EJ (2003) *J Sci Food Agric* 83:564-573
- [2] Dyson HJ, Wright PE (2005) *Nat Rev Mol Cell Biol* 6:197-208
- [3] Dunker AK, Silman I, Uversky VN, Sussman JL (2008) *Curr Opin Struct Biol* 18:756-764
- [4] Tompa P (2003) *J Mol Struct (Theochem)*:361-371
- [5] Wright PE, Dyson HJ (1999) *J Mol Biol* 293:321-331
- [6] Bernstein SL, Dupuis NF, Lazo ND, Wytttenbach T, Condrón MM, Bitan G, Teplow DB, Shea J-E, Ruotolo BT, Robinson CV, Bowers MT (2009) *Nature Chemistry* 1:326-331
- [7] Murray MM, Bernstein SL, Nyugen V, Condrón MM, Teplow DB, Bowers MT (2009) *J Am Chem Soc* 131:6316-6317
- [8] Loo JA, Ogorzalek-Loo RR (1997) In: Cole RB (ed) *Electrospray ionization mass spectrometry of peptides and proteins*. John Wiley and Sons, Inc., New York
- [9] Pramanik BN, Bartner PL, Mirza UA, Liu YH, Ganguly AK (1998) *J Mass Spectrom* 33:911-920
- [10] Jorgensen TJD, Roepstorff P, Heck AJR (1998) *Anal Chem* 70:4427-4432
- [11] Kapur A, Beck JL, Brown SE, Dixon NE, Sheil MM (2002) *Protein Sci* 11:147-157
- [12] Bligh SWA, Haley T, Lowe PN (2003) *J Mol Recognit* 16:139-148
- [13] Sobott F, McCammon MG, Robinson CV (2003) *Int J Mass Spectrom Ion Processes* 230:193-200
- [14] Jorgensen TJD, Hvelplund P, Andersen JU, Roepstorff P (2002) *Int J Mass Spectrom Ion Processes* 219:659-670
- [15] Mehansho H, Butler LG, Carlson DM (1987) *Annu Rev Nutr* 7:423-440

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

- [16] Bennick A (2002) *Critical Reviews in Oral Biology & Medicine* 13:184-196
- [17] Carlson DM (1993) *Crit Rev Oral Biol Med* 4:495-502
- [18] Sarni-Manchado P, Canals-Bosch J, Mazerolles G, Cheynier V (2008) *J Agric Food Chem* 56:9563-9569
- [19] Sarni-Manchado P, Cheynier V, Moutounet M (1999) *J Agric Food Chem* 47:42-47
- [20] Mehansho H, Carlson DM (1983) *J Biol Chem* 258:6616-6620
- [21] Mehansho H, Clements S, Sheares BT, Smith S, Carlson DM (1985) *J Biol Chem* 260:4418-4423
- [22] Mole S, Butler LG, Iason G (1990) *Biochemical Systematics and Ecology* 18:287-293
- [23] Asquith TN, Uhlig J, Mehansho H, Putnam L, Carlson DM, Butler L (1987) *J Agric Food Chem* 35:331-334
- [24] Austin PJ, Suchar LA, Robbins CT, Hagerman AE (1989) *J Chem Ecol* 15:1335-1347
- [25] McArthur C, Sanson GD, Beal AM (1995) *J Chem Ecol* 21:663-691
- [26] Dixon R, Xie D, Sharma S (2005) *New Phytologist* 165:9-28
- [27] Zucker WV (1983) *The American Naturalist* 121:335-365
- [28] Poncet-Legrand C, Edelmann A, Putaux J-L, Cartalade D, Sarni-Manchado P, Vernhet A (2006) *Food Hydrocoll* 20:687-697
- [29] Sarni-Manchado P, Cheynier V (2002) *J Mass Spectrom* 37:609-616
- [30] Charlton AJ, Baxter NJ, Lilley TH, Haslam E, McDonald CJ, Williamson MP (1996) *FEBS Lett* 382:289-292
- [31] Canon F, Paté F, Meudec E, Marlin T, Cheynier V, Giuliani A, Sarni-Manchado P (2009) *Anal and Bioanal Chem* 395:2535-2545
- [32] Jobstl E, O'Connell J, Fairclough JPA, Williamson MP (2004) *Biomacromolecules* 5:942-949
- [33] Ricardo da Silva JM, Rigaud J, Cheynier V, Cheminat A, Moutounet M (1991) *Phytochemistry* 30:1259-1264
- [34] Pascal C, Bigey F, Ratomahenina R, Boze H, Moulin G, Sarni-Manchado P (2006) *Protein Expr Purif* 47:524-532
- [35] Yin S, Xie Y, Loo JA (2008) *J Am Soc Mass Spectrom* 19:1199-1208
- [36] Haller I, Mirza UA, Chait BT (1996) *J Am Soc Mass Spectrom* 7:677-681
- [37] Robinson CV (2001) *J Am Soc Mass Spectrom* 12:126-126
- [38] Zhang J, Kashket S (1998) *Caries Res* 32:233-238
- [39] Jørgensen TJD, Delforge D, Remacle J, Bojesen G, Roepstorff P (1999) *Int J Mass Spectrom Ion Processes* 188:63-85
- [40] Wan KX, Gross ML, Shibue T (2000) *J Am Soc Mass Spectrom* 11:450-457
- [41] Akashi S, Osawa R, Nishimura Y (2005) *J Am Soc Mass Spectrom* 16:116-125
- [42] Flanzy C (1998) *Oenologie - Fondements scientifiques et technologiques*. Lavoisier, Paris
- [43] Champagnol F (1986) *Revue Française d'Oenologie* 26:26-57
- [44] Shimada T (2006) *J Chem Ecol* 32:1149-1163
- [45] Fuxreiter M, Simon I, Friedrich P, Tompa P (2004) *J Mol Biol* 338:1015-1026
- [46] Oldfield C, Meng J, Yang J, Yang MQ, Uversky V, Dunker AK (2008) *BMC Genomics* 9:S1

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

- [47] Mehansho H, Hagerman A, Clements S, Butler LG, Rogler JC, Carlson DM (1983) *Proc Natl Acad Sci USA* 80:3948-3952
- [48] Boze H, Marlin T, Durand D, Pérez J, Vernhet A, Canon F, Sarni-Manchado P, Cheynier V, Cabane B (2010) *Biophys J* 99:656-665
- [49] Tompa P (2003) *BioEssays* 25:847-855
- [50] Khalsa-Moyers G, McDonald WH (2006) *Brief Funct Genomic Proteomic* 5:98-111
- [51] Chen Y-LC, J. M.; Collings, B. A.; Konermann, L.; Douglas, D. J. (1998) *Rapid Commun Mass Spectrom* 12:1003-1010
- [52] Wright PE, Dyson HJ (2009) *Curr Opin Struct Biol* 19:31-38
- [53] Robinson CV, Chung EW, Kragelund BB, Knudsen J, Aplin RT, Poulsen FM, Dobson CM (1996) *J Am Chem Soc* 118:8646-8653
- [54] Sobott FM, MG; Hernández, H; Robinson, CV. (2005) *Phil Trans R Soc A* 363:379-391
- [55] Simon C, Barathieu K, Laguerre M, Schmitter JM, Fouquet E, Pianet I, Dufoure EJ (2003) *Biochemistry* 42:10385-10395
- [56] Hagerman AE (1989) In: Hemingway RW, and Karchesy JJ (eds.) *Chemistry of tannin-protein complexation*. Plenum Press, New York, London
- [57] Bacon JR, Rhodes MJC (1998) *J Agric Food Chem* 46:5083-5088
- [58] Poncet-Legrand C, Gautier C, Cheynier V, Imberty A (2007) *J Agric Food Chem* 55:9235-9240

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

Table 1. E_{50} and E_{CoM} of the $[IB5 \cdot IT]^{+6}$ complexes.

	IB5•EgC	IB5•ECG	IB5•EgCG	IB5•B2	IB5•B2,3'OG
E_{50} (eV)	67.5	105	120	150	165
E_{CoM} (eV)	0.26	0.40	0.45	0.55	0.65

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

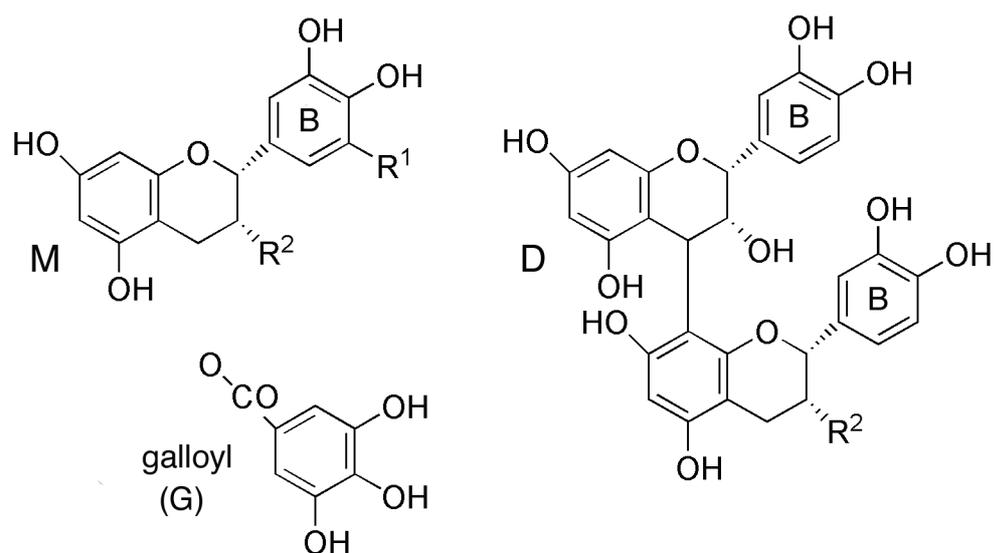
Figure 1. Structures of **a** studied tannins (T) and **b** reserpine

Figure 2. Positive ion mass spectra of IB5:T (1:10) interaction mixtures. Peaks of IB5a and its complexes are labelled.

Figure 3. Dissociation curves of $[\text{IB5}\cdot\text{1T}]^{+6}$ and MS/MS spectra of $[\text{IB5}\cdot\text{1EgCG}]^{+6}$ at collision energies of respectively 90eV and 150eV.

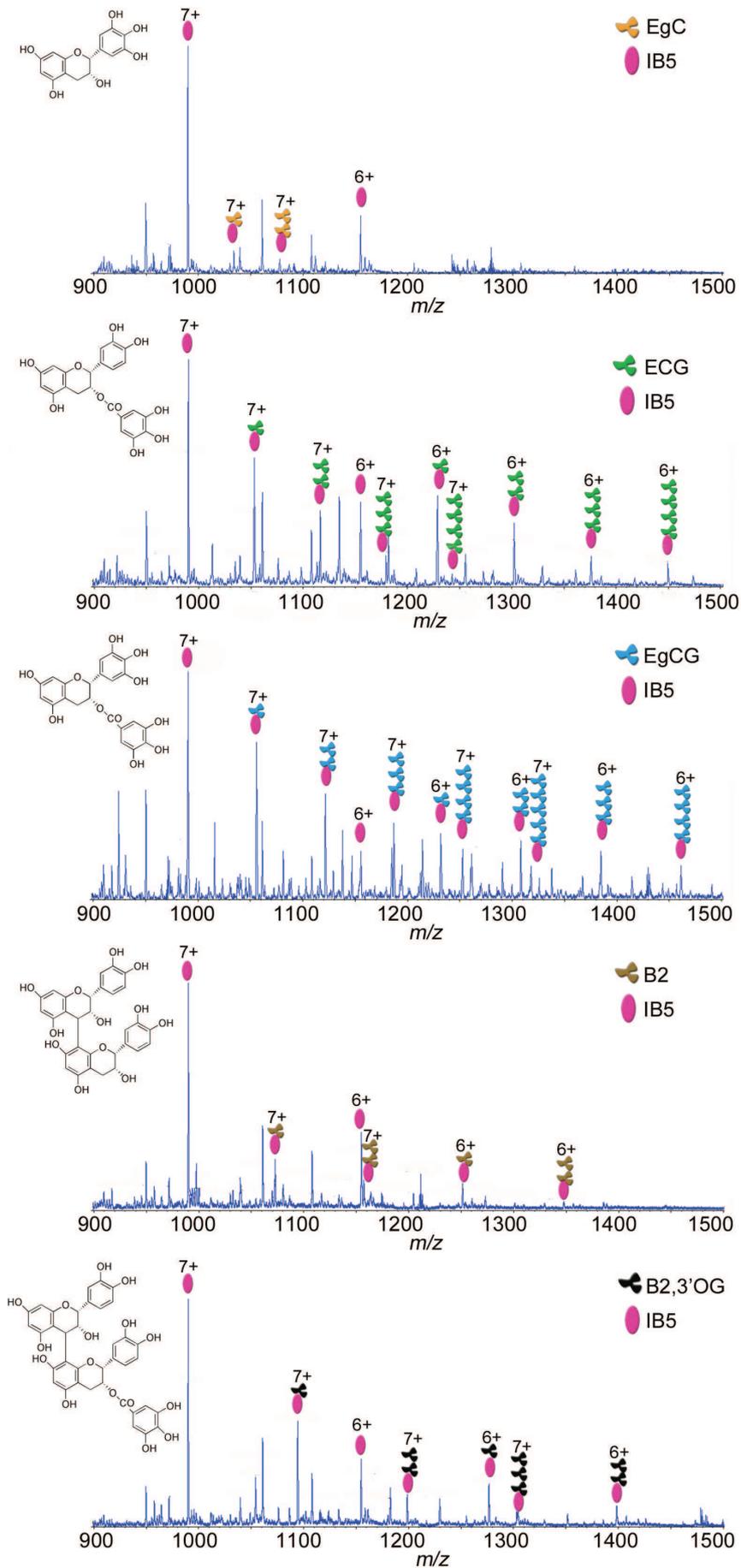
Figure 4. Relation between the E_{CoM} of IB5•tannin complexes and the number of tannin hydroxyl groups.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

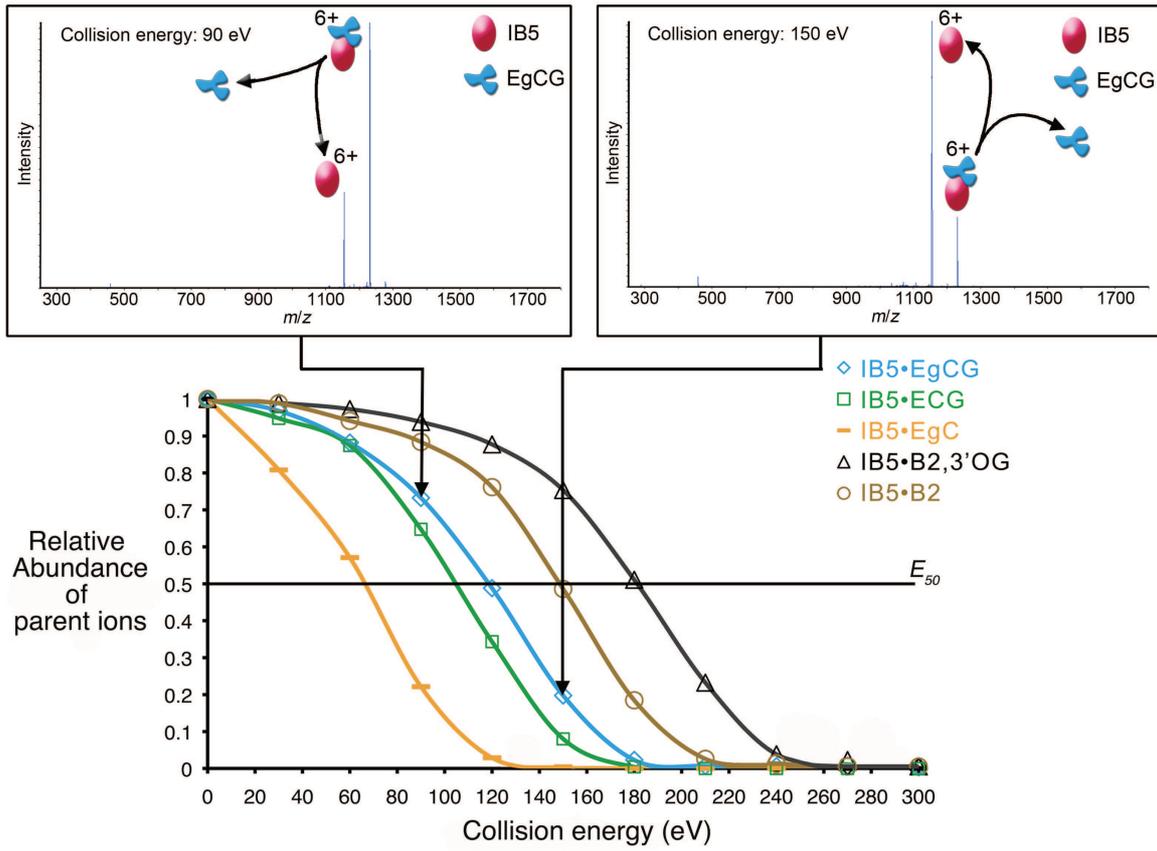


Tannins	skeleton	R ¹	R ²
Epigallocatechin	M	OH	OH (UG)
Epicatechin gallate	M	H	galloyl (G)
Epigallocatechin gallate	M	OH	galloyl (G)
B2	D		OH (UG)
B2,3'OG	D		galloyl (G)

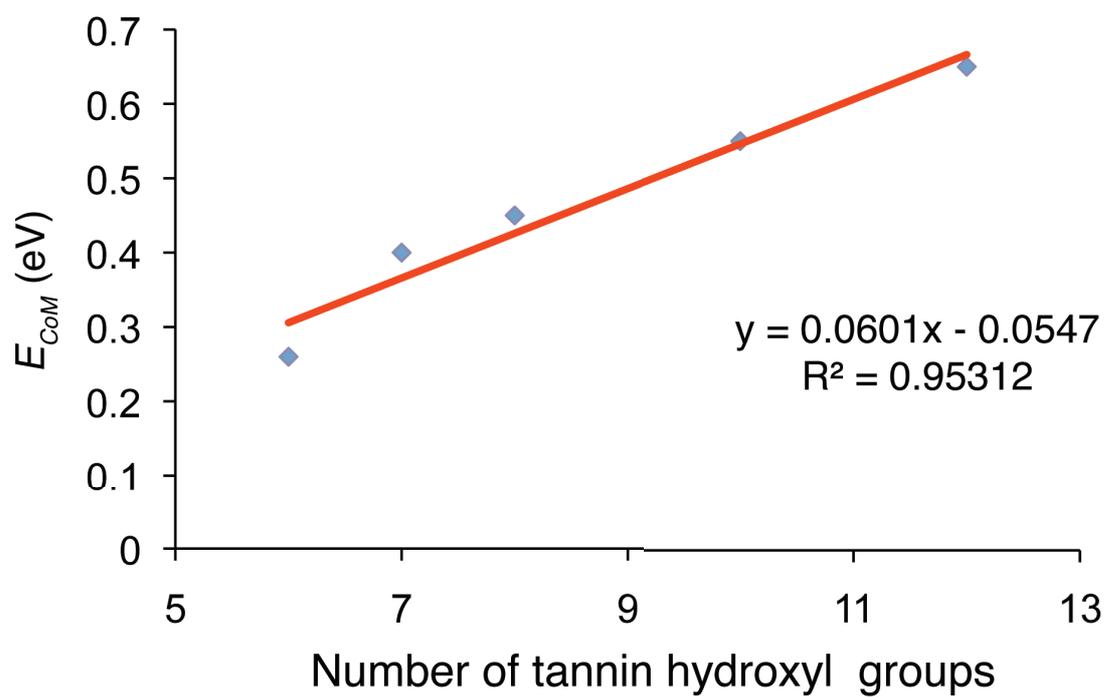
III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin



III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin



3



Au cours de cette étude, nous avons à la fois montré des différences de stœchiométries entre complexes qui impliquent divers tanins modèles et des différences de stabilité des complexes de stœchiométrie 1•1. Ainsi, IB5 est capable de lier différents tanins selon des « affinités » liées à leur nature. Cette étude a également permis de confirmer l'implication de liaisons hydrogène dans la stabilité des complexes.

III.3. Etude des interactions IB5-tanin par compétition par ESI-MS

Les approches de compétition récemment développées par spectrométrie de masse ont mis en évidence des différences d'affinité dans le cadre des systèmes biologiques hôte-ligand (Jorgensen *et al.*, 1998 ; Kempen et Brodbelt, 2000). Des expériences de compétition entre tanins vis-à-vis d'IB5 ont donc été réalisées. Cette démarche offre également la possibilité d'évaluer la stabilité des complexes hétéro-tanins par MS/MS et de la comparer avec celles précédemment obtenues.

Au cours de ces expériences, les tanins modèles ont été mis en compétition deux à deux vis-à-vis de la protéine IB5. La solution de compétiteurs qui contient les deux tanins modèles est mélangée à la solution de protéine, ce qui permet de procéder à des compétitions directes en solution (Sanglier *et al.*, 2005). Les concentrations utilisées sont de 25 μ M pour chaque tanin et de 5 μ M pour IB5.

Les spectres MS sont acquis à l'aide d'un spectromètre de masse de type Q-TOF (QStar, Applied Biosystem).

Les régions des spectres correspondant aux états de charge majoritaires d'IB5 et de ses complexes sont présentées dans les figures 70 à 79. Les pics *m/z* non légendés correspondent aux isoformes minoritaires d'IB5 et de leurs complexes avec les tanins.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

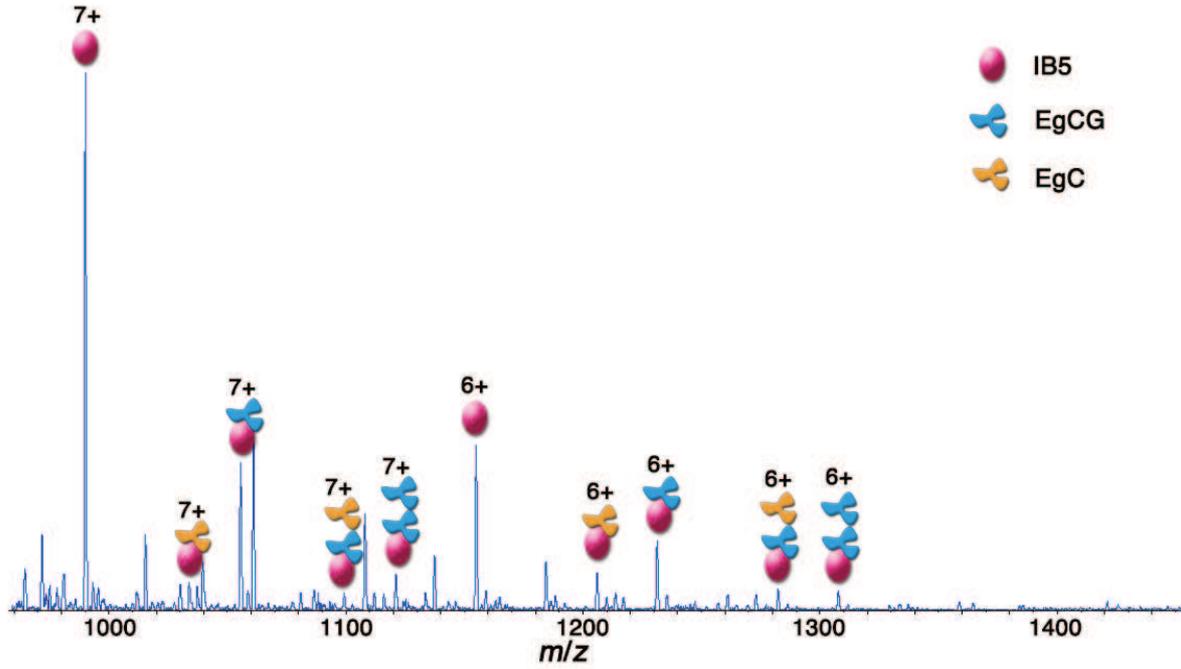


Figure 70. Spectre MS de la solution d'interaction IB5 5µM - EgCG 25µM - EgC 25µM.

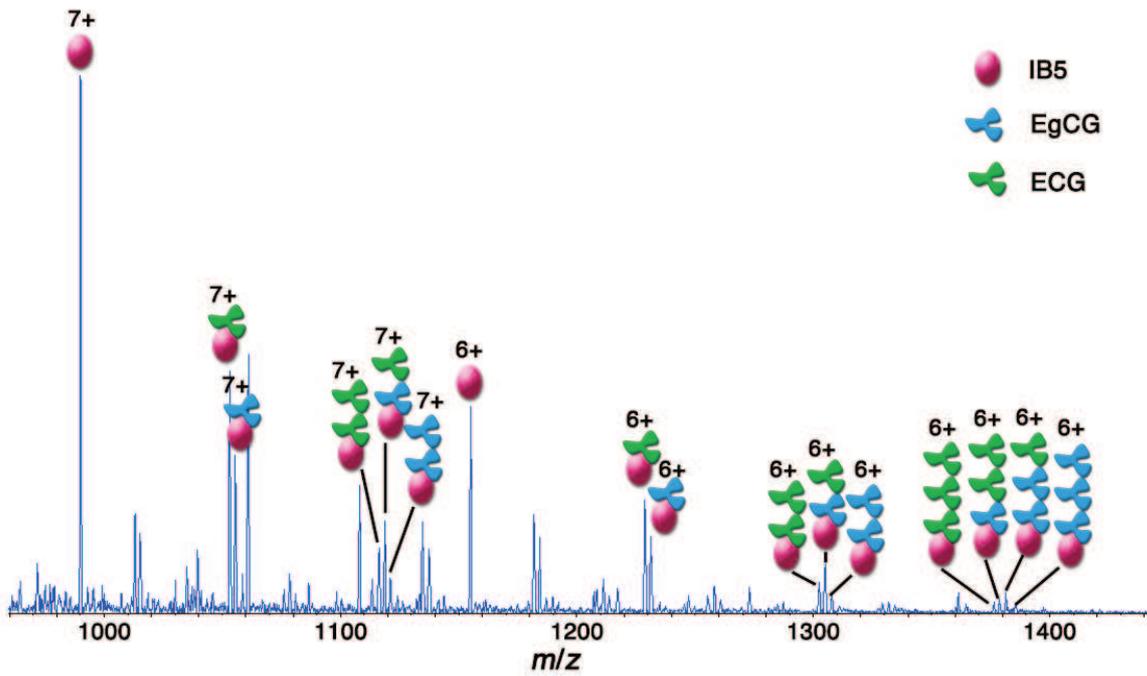


Figure 71. Spectre MS de la solution d'interaction IB5 5 µM - EgCG 25 µM - ECG 25 µM.

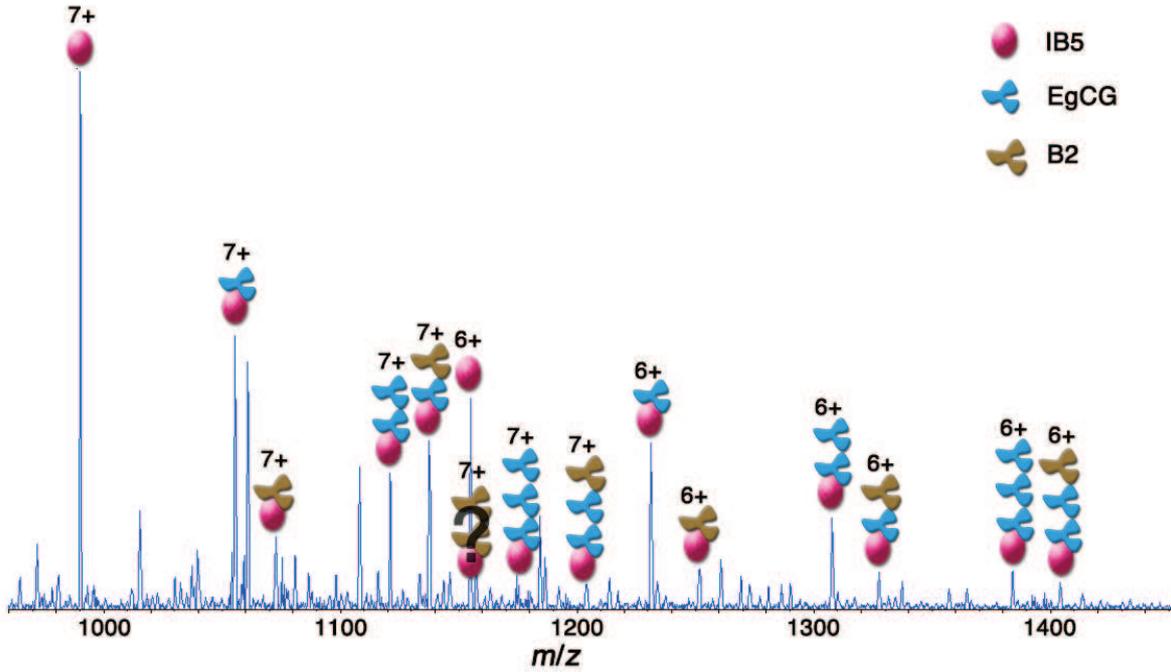


Figure 72. Spectre MS de la solution d'interaction IB5 5 μ M - EgCG 25 μ M – B2 25 μ M. Le point d'interrogation indique que la présence du pic de m/z correspondant à $[\text{IB5}\cdot 2\text{B2}]^{7+}$ ne peut être déterminée car ce pic a le même rapport m/z que $[\text{IB5}]^{6+}$.

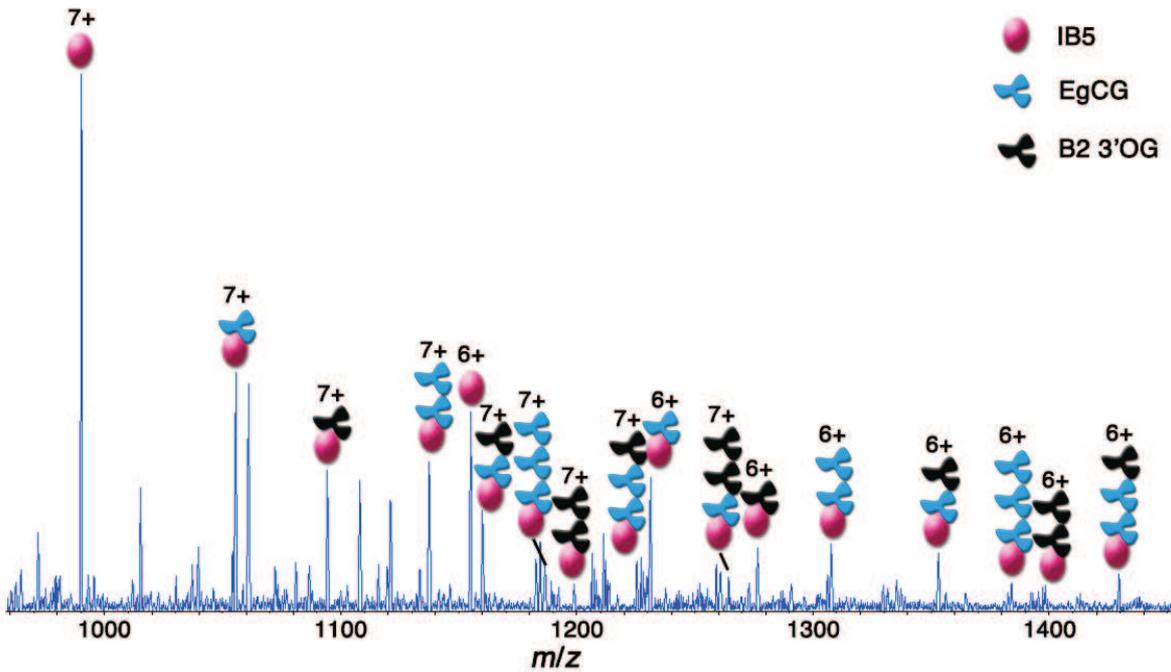


Figure 73. Spectre MS de la solution d'interaction IB5 5 μ M - EgCG 25 μ M – B2 3'OG 25 μ M.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

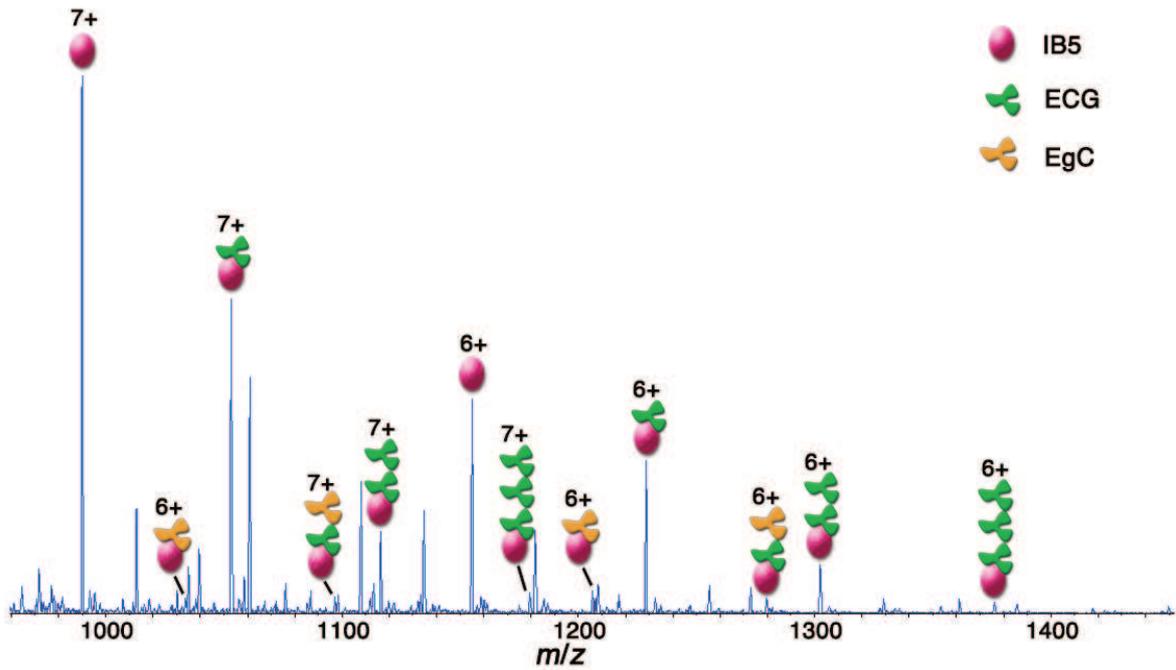


Figure 74. Spectre MS de la solution d'interaction IB5 5 μ M – EgG 25 μ M – ECG 25 μ M.

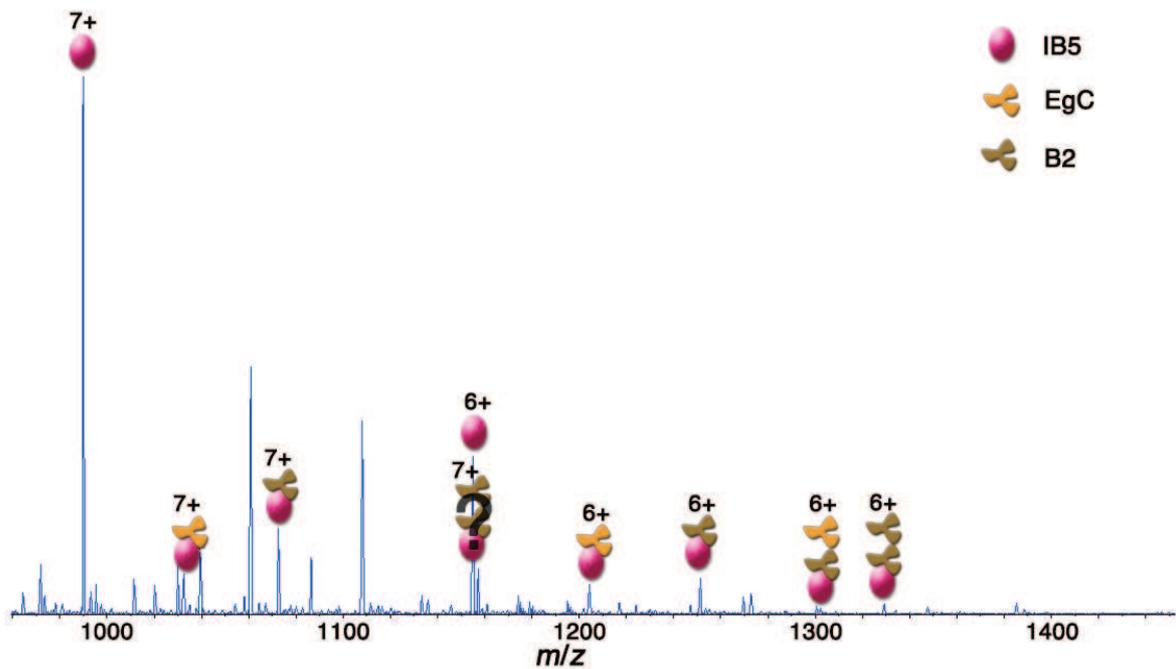


Figure 75. Spectre MS de la solution d'interaction IB5 5 μ M – EgC 25 μ M – B2 25 μ M. Le point d'interrogation indique que la présence du pic de m/z correspondant à $[\text{IB5}\cdot 2\text{B2}]^{7+}$ ne peut être déterminée car ce pic a le même rapport m/z que $[\text{IB5}]^{6+}$.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

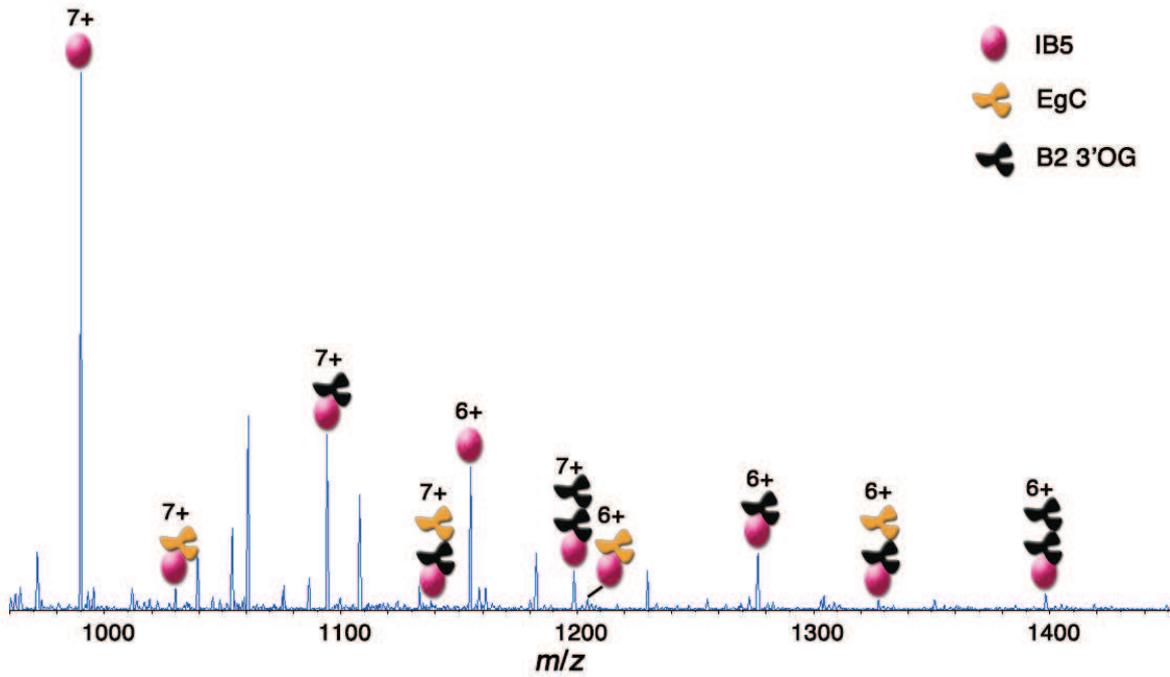


Figure 76. Spectre MS de la solution d'interaction IB5 5 μ M – EgC 25 μ M – B2 3'OG 25 μ M.

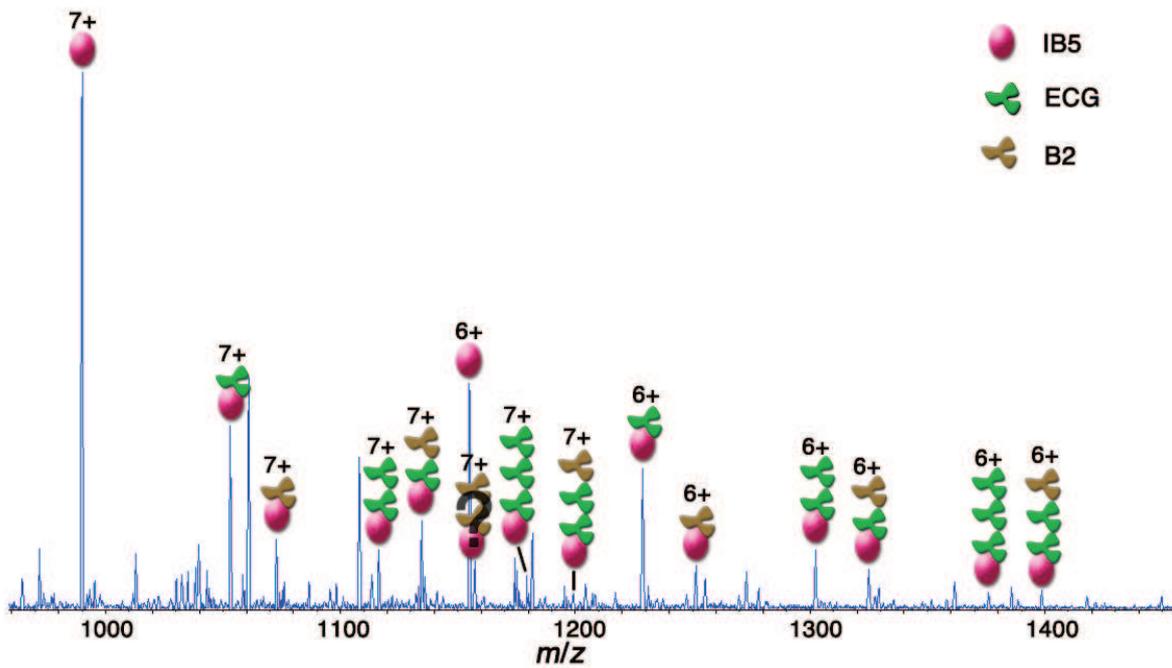


Figure 77. Spectre MS de la solution d'interaction IB5 5 μ M – ECG 25 μ M – B2 25 μ M.
Le point d'interrogation indique que la présence du pic de m/z correspondant à $[IB5 \cdot 2B2]^{7+}$ ne peut être déterminée car ce pic a le même rapport m/z que $[IB5]^{6+}$.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

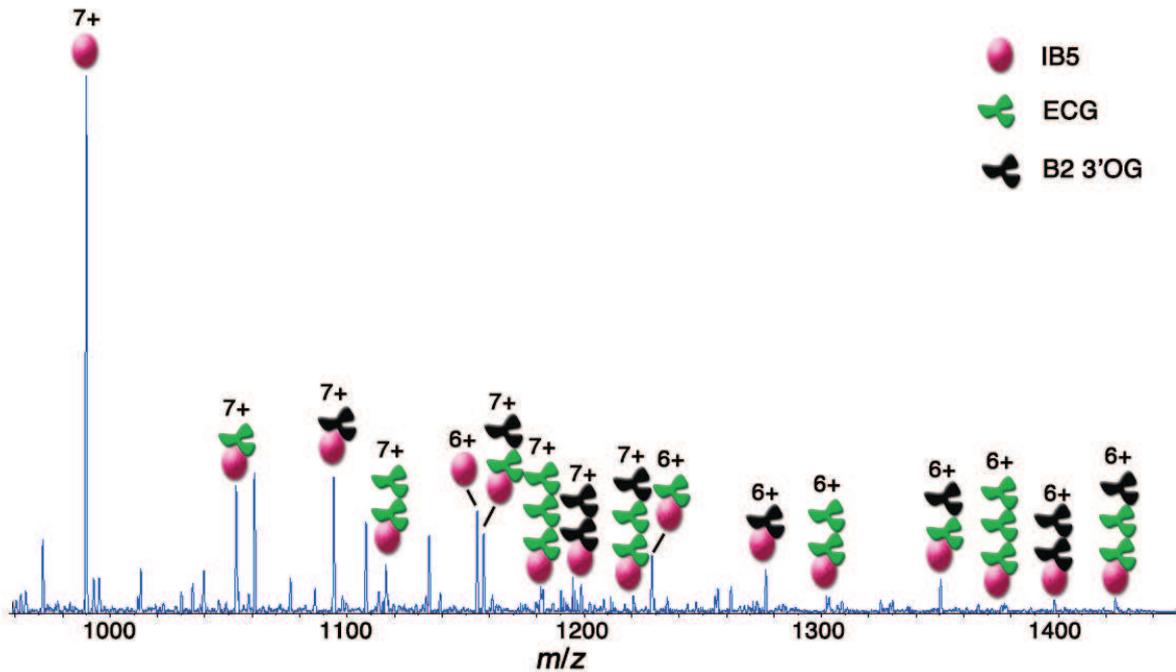


Figure 78. Spectre MS de la solution d'interaction IB5 5 μ M – ECG 25 μ M – B2 3'OG 25 μ M.

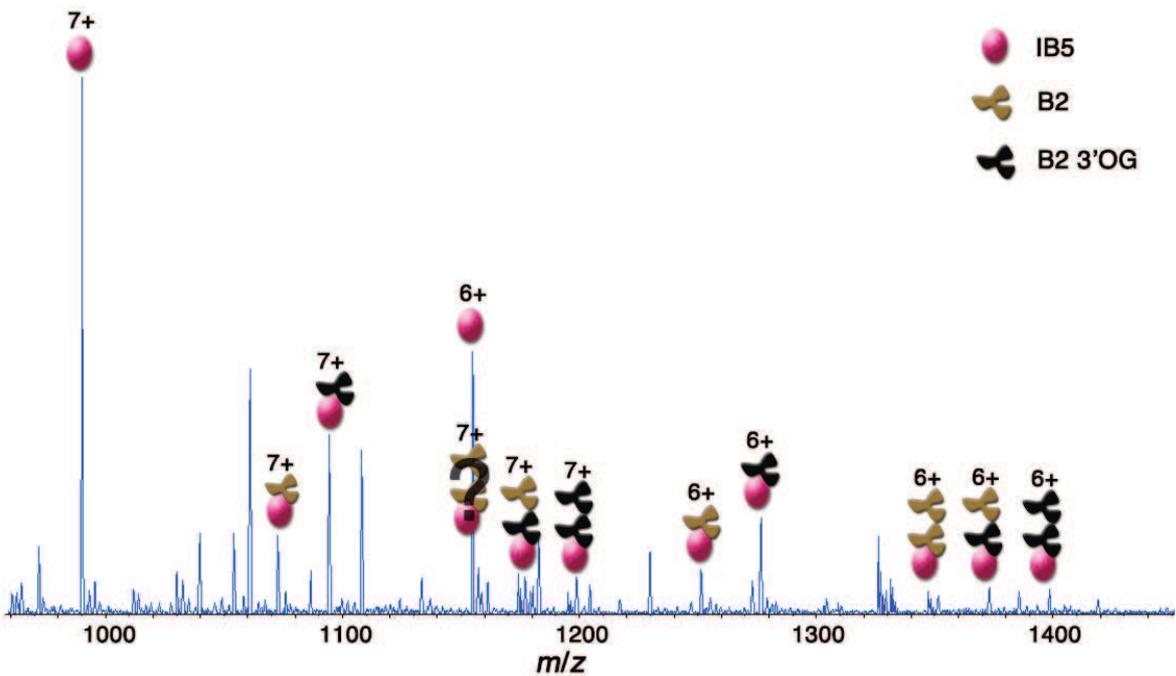


Figure 79. Spectre MS de la solution d'interaction IB5 5 μ M – B2 3'OG 25 μ M – B2 25 μ M. Le point d'interrogation indique que la présence du pic de m/z correspondant à $[IB5 \cdot 2B2]^{7+}$ ne peut être déterminée car ce pic a le même rapport m/z que $[IB5]^{6+}$.

Les spectres de masse des différents mélanges étudiés présentent les signaux des ions correspondant à la protéine libre, aux tanins libres (non visibles sur les régions de

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

spectres présentées ici) et aux complexes IB5•tanin. Pour tous les mélanges, les ions présents correspondent à deux catégories de complexes :

- les "homo-complexes" impliquant un seul type de tanin ;
- les "hétéro-complexes" impliquant des tanins différents.

Ces résultats prouvent que différents tanins se lient simultanément sur IB5 pouvant former des "hétéro-complexes" de différentes stœchiométries (figures 70 à 79).

L'analyse des spectres révèle que les ions des "homo-complexes" impliquant les tanins galloylés sont plus abondants que ceux impliquant les tanins non galloylés. Les plus hautes stœchiométries sont observées pour les complexes comprenant des monomères galloylés. Ces données confirment que la galloylation est un groupement fonctionnel des tanins important dans l'établissement des interactions PRP•tanin. Ces résultats sont en accord avec ceux obtenus lors des expériences précédentes, la présence d'un acide gallique semblant stabiliser l'interaction IB5-tanin.

Selon le couple de compétiteurs, la composition et les stœchiométries maximales observées des "hétéro-complexes" et des "homo-complexes" sont différentes (tableau 11).

Tableau 11. Récapitulatif des stœchiométries maximales observées pour chacune des expériences de compétition.

T1 / T2	EgC / ECG	EgC / EgCG	EgC / B2	EgC / B2 3'OG	ECG / EgCG
Homo-complexes T1	1	1	1	1	3
Homo-complexes T2	3	2	2	2	3
Hétéro-complexes	2	2	2	2	3

T1 / T2	ECG / B2	ECG / B2 3'OG	EgCG / B2	EgCG / B2 3'OG	B2 / B2 3'OG
Homo-complexes T1	3	3	3	3	2
Homo-complexes T2	1	2	1	2	2
Hétéro-complexes	3	3	3	3	2

Lors de la mise en compétition des deux monomères galloylés, EgCG et ECG, les "hétéro-complexes" atteignent les mêmes stœchiométries que les "homo-complexes". Lors

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

de la compétition impliquant les deux dimères B2 et B2 3'OG, la même constatation est faite. Dans le cas d'une solution contenant un dimère et un monomère galloylé, les stœchiométries maximales des "hétéro-complexes" sont supérieures à celles des "homo-complexes" de dimères et égales à celles des "homo-complexes" de monomères. L'encombrement stérique supérieur des dimères par rapport aux monomères peut limiter leur accès à certains sites d'interaction par rapport aux monomères, ce qui peut expliquer les stœchiométries moins importantes atteintes par les complexes impliquant uniquement des dimères. Les expériences de compétition permettent d'établir les affinités relatives de différents ligands pour une même protéine. Dans les études menées par ESI-MS, la mesure de « l'affinité » en solution est une information obtenue, pour un couple de compétiteur donné, par comparaison de l'intensité des espèces libres et complexées sur le spectre de masse (Jorgensen *et al.*, 1998 ; Daniel *et al.*, 2003). Cette approche est basée sur l'approximation que les intensités déduites du spectre de masse sont proportionnelles aux concentrations en solution.

Pour cela, plusieurs hypothèses doivent être vérifiées :

- les complexes ne doivent pas se dissocier en phase gazeuse ;
- les espèces étudiées (protéines complexées et protéines libres) doivent avoir le même rendement d'ionisation ;
- la transmission et la détection des protéines complexées et des protéines libres dans le spectromètre de masse doivent être identiques.

Ces conditions sont souvent remplies dans le cas de systèmes protéine-ligand (Li *et al.*, 1993 ; Zhang *et al.*, 2003). En effet, on considère généralement que la fixation d'un ligand de petite taille n'influence pas le rendement d'ionisation de la protéine et que la différence de masse engendrée n'est pas suffisante pour induire une transmission et une détection différentes. Par ailleurs, les conditions d'analyses sont optimisées pour éviter la dissociation des complexes en phase gazeuse. Il est cependant admis que les interactions où l'effet hydrophobe en solution est prédominant peuvent se révéler difficiles à préserver en phase gazeuse (Li *et al.*, 1993). Ainsi, dans le cas des systèmes IB5/tanins étudiés, ces hypothèses se trouvent vérifiées, permettant de classer par ordre d'affinité relative les tanins à partir des expériences de compétition. Le classement des cinq tanins modèles est établi en fonction de leur affinité pour IB5 :



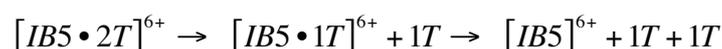
Ce classement confirme la plus forte affinité des tanins galloylés pour IB5. L'augmentation du degré de polymérisation accroît l'affinité des tanins non galloylés pour IB5, mais limite l'effet de la galloylation (B2 3'OG par rapport à ECG). Enfin, le degré

d'hydroxylation du noyau B (comparaison ECG/EgCG) ne semble pas modifier significativement l'affinité des tanins. Ce résultat apparaît pour partie en contradiction avec celui obtenu précédemment en MS/MS CID (cf. page 183). Ces différences sont discutées plus loin dans ce chapitre.

III.4. Etude des complexes de stœchiométrie 1•2 par ESI-MS/MS CID

Afin d'étudier l'influence de la liaison d'un tanin supplémentaire sur la stabilité des édifices supramoléculaires formés, tous les complexes de stœchiométrie 1•2 ont été analysés par MS/MS. La stabilité des "homo" et "hétéro-complexes" de stœchiométrie 1•2 (IB5•2T) a été étudiée par MS/MS CID. Pour cette étude, les ions d'état de charge 6+ ont été préférés. La sélection de cet état de charge, correspondant à la protonation des seuls acides aminés basiques présents sur IB5, limite les problèmes de fragmentation de la chaîne peptidique. En effet, pour les états de charge supérieurs, la présence de protons supplémentaires extrêmement mobiles augmente cette fragmentation, ce qui rend l'interprétation des spectres plus laborieuse.

Quel que soit le complexe 1•2 considéré, l'augmentation de l'énergie de collision conduit à la dissociation des complexes 1•2 en deux étapes :



La disparition progressive du pic de l'ion parent $[IB5 \cdot 2T]^{6+}$ est accompagnée par l'apparition d'un pic $[IB5 \cdot 1T]^{6+}$, qui disparaît à son tour engendrant un pic $[IB5]^{6+}$. Les tanins sont libérés sous forme neutre (figure 80). Pour chacun des ions $[IB5 \cdot 2T]^{6+}$, les courbes de dissociation des tanins liés à IB5 sont établies en fonction de l'énergie de collision appliquée. Elles présentent une allure similaire, mais les valeurs d'énergie de collision nécessaire pour la dissociation sont différentes selon les tanins impliqués dans le complexe. L'apparition de protéine libre débute avec la dissociation du deuxième tanin. Les courbes de dissociation du deuxième tanin, d'une part, et de libération de la protéine, d'autre part, se croisent lorsque 50% de la protéine est libre et 50% du deuxième tanin est dissocié. Ces observations sont valables pour tous les complexes $[IB5 \cdot 2T]^{6+}$.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

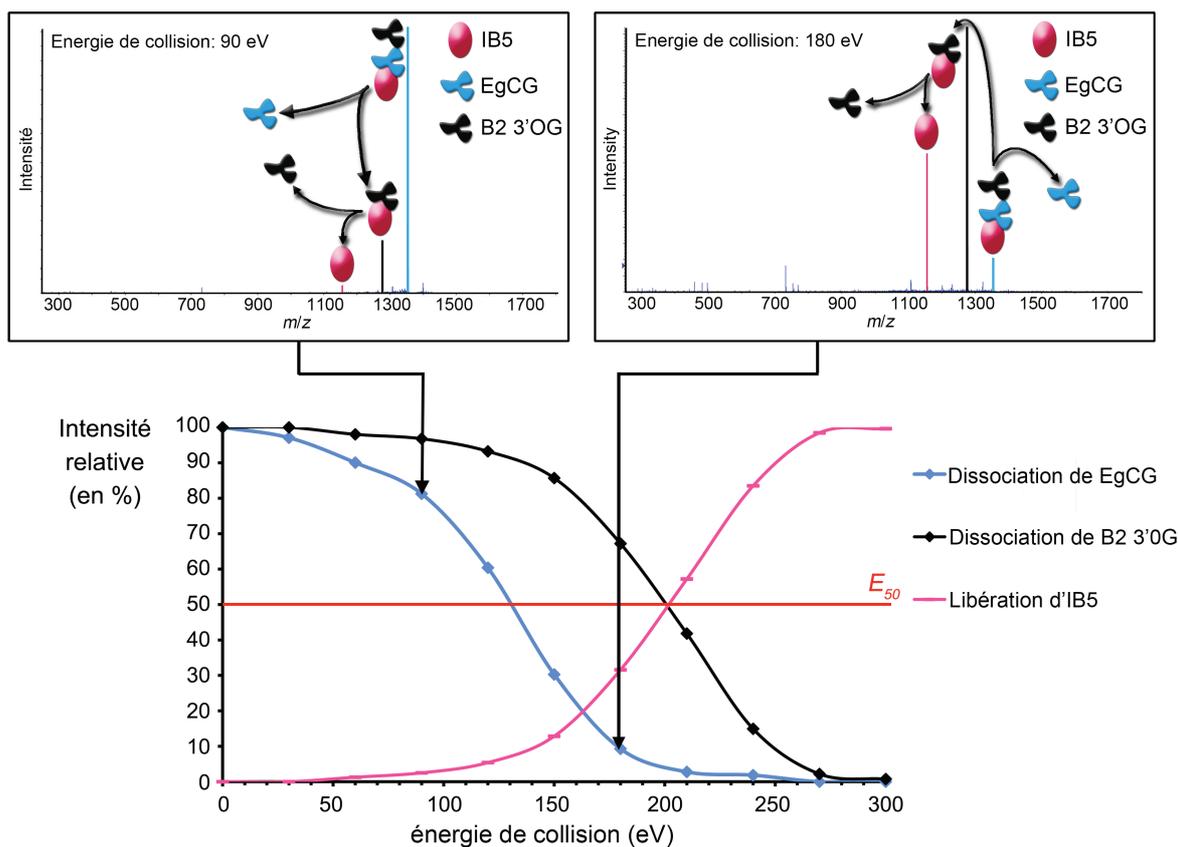


Figure 80. Courbes de dissociation des tanins du complexe $[IB5 \cdot EgCG \cdot B2 \ 3'OG]^{6+}$ et de libération d'IB5.

L'ensemble des courbes de dissociation des complexes $[IB5 \cdot 2T]^{6+}$ est présenté en annexe 1. Ces expériences ont été répétées 2 à 3 fois, les barres d'erreurs n'ont pas été indiquées à cause de leur faible amplitude. Dans cette partie, nous ne présenterons que les courbes de dissociation des ions des complexes impliquant l'EgC ou le B2 3'OG de charge 6+ (figures 81 et 82). A partir de ces dernières, les énergies de collision dans le système du laboratoire entraînant 50% de dissociation de l'ion parent (E_{50}) sont déterminées graphiquement. Elles sont ensuite converties en énergie au centre-des-masses (E_{CoM}) (cf. publications 1 et 2). Les sections efficaces des complexes composés d'IB5 et de un ou deux tanins sont très proches et donc équivalentes. La comparaison des énergies internes transférées aux complexes aux cours des collisions peut ainsi être assimilée à celle des E_{CoM} correspondantes (cf. publications 1 et 2).

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

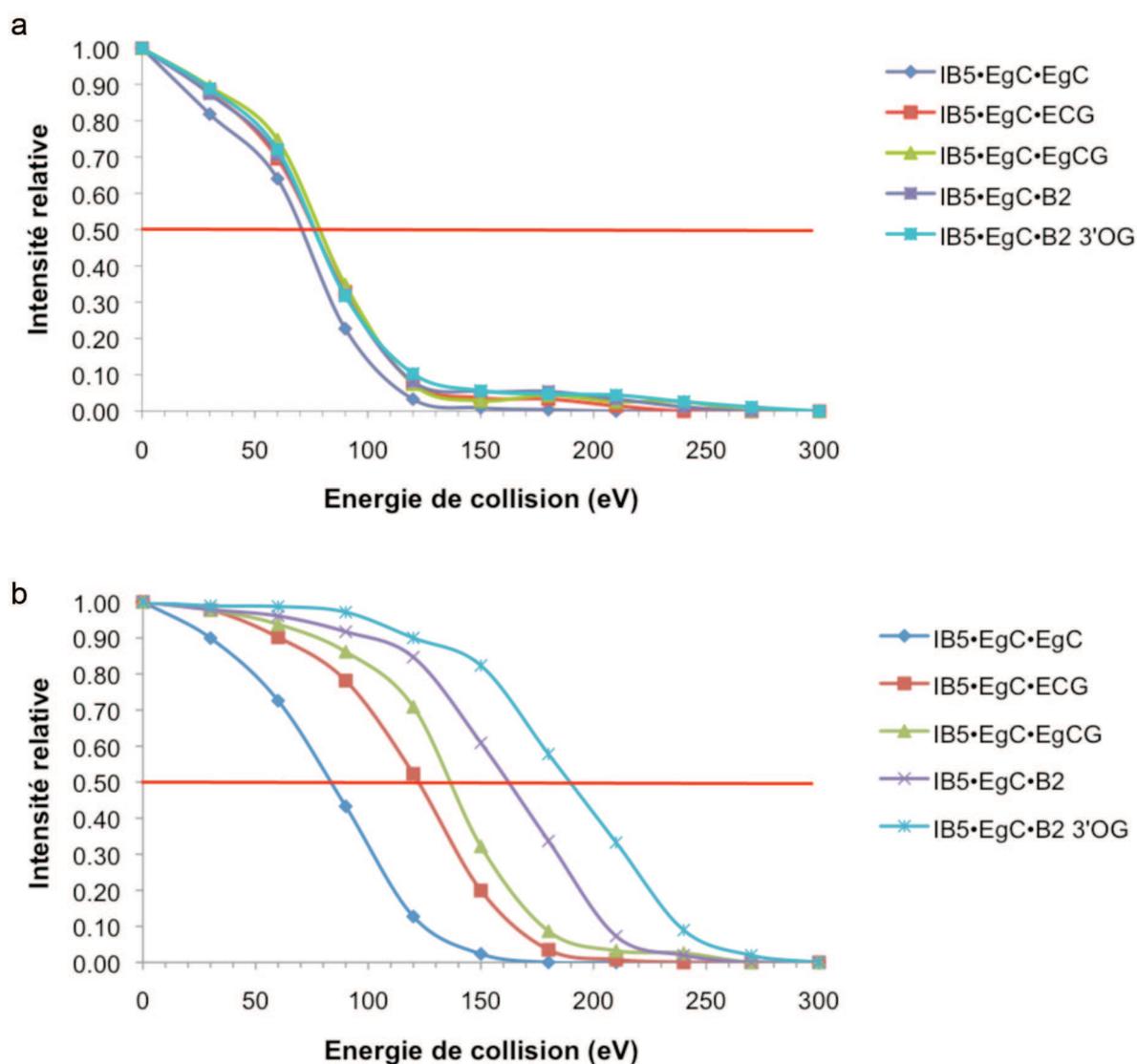


Figure 81. a. courbes de dissociation du premier tanin des complexes $[IB5 \cdot EgC \cdot T]^{6+}$ avec $T = EgC, ECG, EgCG, B2$ ou $B2\ 3'OG$; b. courbes de dissociation du second tanin des complexes $[IB5 \cdot EgC \cdot T]^{6+}$.

Tableau 12. E_{CoM} de demi dissociation des tanins des complexes $[IB5 \cdot EgC \cdot T]^{6+}$

	IB5•EgC•EgC	IB5•EgC•ECG	IB5•EgC•EgCG	IB5•EgC•B2	IB5•EgC•B2 3'OG
E_{CoM} 1 ^{er} tanin (eV)	0,26	0,27	0,28	0,27	0,27
E_{CoM} 2 nd tanin (eV)	0,30	0,44	0,49	0,59	0,68

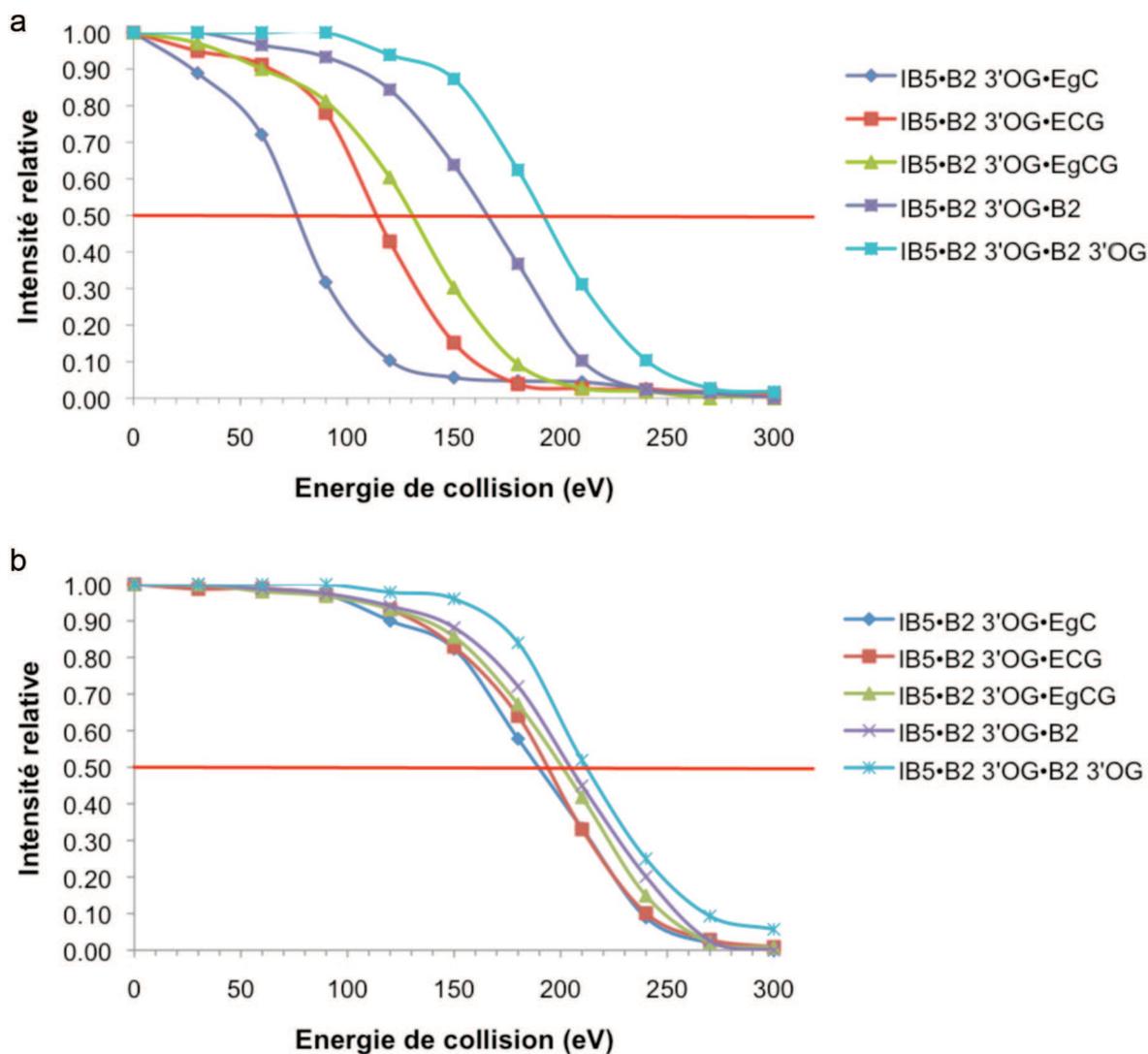


Figure 82. a. courbes de dissociation du premier tanin des complexes $[IB5•B2 3'OG•T]^{6+}$ avec $T= EgC, ECG, EgCG, B2$ ou $B2 3'OG$; b. courbes de dissociation du second tanin des complexes $[IB5•B2 3'OG•T]^{6+}$.

Tableau 13. E_{CoM} de demi dissociation des tanins des complexes $[IB5•B2 3'OG•T]^{6+}$.

	IB5•B2 3'OG •EgC	IB5•B2 3'OG •ECG	IB5•B2 3'OG •EgCG	IB5•B2 3'OG •B2	IB5•B2 3'OG •B2 3'OG
E_{CoM} 1 ^{er} tanin (eV)	0,26	0,39	0,45	0,55	0,64
E_{CoM} 2 nd tanin (eV)	0,68	0,67	0,69	0,69	0,71

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

L'étude des complexes 1•1 a permis d'établir le classement des tanins en fonction de la « stabilité de leur liaison » à IB5 : EgC < ECG < EgCG < B2 < B2 3'OG (les E_{CoM} de demi dissociation des complexes $[IB5 \cdot 1T]^{6+}$ sont rappelées dans le tableau 14).

L'analyse de l'ensemble des courbes de dissociation des tanins des "homo" et "hétéro-complexes" de stoechiométrie 1•2 révèle que le tanin dissocié en premier est toujours celui qui est classé d'après l'étude des complexes 1•1 comme nécessitant la plus faible énergie pour être dissocié d'IB5. Par exemple, les courbes obtenues pour les complexes $[IB5 \cdot EgC \cdot T]^{6+}$ montrent que l'EgC est toujours le tanin qui est dissocié en premier (figure 81).

La comparaison des valeurs des E_{CoM} de demi dissociation de l'EgC des "hétéro-complexes" $[IB5 \cdot EgC \cdot T]^{6+}$ (tableau 12, 1^{ère} ligne) montre que des énergies internes équivalentes sont transférées aux différents ions pour dissocier l'EgC, quel que soit l'autre tanin présent dans le complexe. Elles sont équivalentes à celle requise pour dissocier l'EgC lorsqu'il est seul lié à IB5. L' E_{CoM} de demi dissociation du deuxième tanin est la même que celle obtenue lorsque ce tanin est seul lié à IB5 (tableau 13, 2^{ème} ligne et tableau 14, 1^{ère} ligne), excepté pour l'EgC, où elle est légèrement supérieure.

Les courbes de dissociation des tanins des complexes $[IB5 \cdot B2 \text{ 3'OG} \cdot T]^{6+}$ montrent également que le tanin lié avec la force de liaison la plus faible est toujours décroché en premier (figure 82 et tableau 13). La comparaison des E_{CoM} de demi dissociation du 1^{er} tanin (tableau 13, 1^{ère} ligne) montre que des énergies équivalentes sont nécessaires pour dissocier les différents tanins, qu'ils soient seuls ou impliqués dans un complexe de stoechiométrie 1•2 (tableau 14 et tableau 13, 1^{ère} ligne). De plus, l' E_{CoM} de demi dissociation du B2 3'OG est toujours la même, qu'il y ait un deuxième tanin lié ou pas, excepté quand 2 molécules de B2 3'OG constituent le complexe avec IB5 (comparaison tableau 14 et tableau 13, 2^{ème} ligne). Dans ce cas, un transfert d'énergie interne légèrement plus important est requis pour la dissociation de la deuxième molécule de B2 3'OG.

Tableau 14. Rappel des E_{CoM} de demi dissociation des complexes $[IB5 \cdot 1T]^{6+}$.

	IB5•EgC	IB5•ECG	IB5•EgCG	IB5•B2	IB5•B2 3'OG
E_{CoM} Ions parents (eV)	0,26	0,40	0,45	0,55	0,65

Ces résultats suggèrent que les tanins des complexes de stoechiométries 1•2 sont liés au niveau de « sites » différents sur la protéine et que ces « sites » sont équivalents du point

de vue énergétique. Cela indique que les complexes de stœchiométrie 1•2 ne résultent pas du « stacking » d'un deuxième tanin sur le premier (auto-association) mais de leur interaction sur deux « sites » distincts de la protéine. L'énergie de liaison d'une première entité polyphénolique n'est pas modifiée par la fixation d'un deuxième ligand, de même que l'énergie de liaison de celui-ci est équivalente à celle qu'il a lorsqu'il est seul en interaction avec IB5, ainsi ces « sites » sont indépendants.

Dans le cas des interactions entre tanins et IB5, PID salivaire, la MS a rendu possible la démonstration de la formation des complexes ayant de hautes stœchiométries et des "hétéro-complexes". Elle a également conduit à distinguer la fixation préférentielle d'une molécule par rapport aux autres.

III.5. Conclusion et Perspectives

Les expériences menées ont permis d'étudier l'influence des principales caractéristiques structurales des tanins (galloylation, degré d'hydroxylation du noyau B et degré de polymérisation) sur leur interaction avec la *bPRP*, IB5.

Trois types d'expériences ont été réalisées et ont permis d'acquérir des informations complémentaires (tableau 15).

Tableau 15. Expériences réalisées pour étudier l'impact de la structure des tanins sur leur interaction avec IB5.

Type d'expérience	Technique MS utilisée	Informations obtenues
Etude des interactions de IB5 avec un seul tanin	ESI-MS	Stœchiométries des complexes, affinité relative des différents tanins
Expériences de compétition entre 2 tanins	ESI-MS	Stœchiométries des complexes, affinité relative des différents tanins
Etude de la stabilité des complexes IB5•tanin de stœchiométries 1•1 et 1•2	MS/MS CID	Stabilité des complexes en phase gazeuse

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

Ces expériences ont permis de déterminer:

- une échelle d'affinité relative des différents tanins pour IB5:



- une échelle de stabilité des complexes impliquant les tanins en phase gazeuse :



Ces deux échelles n'aboutissent pas au même classement des tanins étudiés. Pour expliquer ce résultat, il est important de mieux définir ce que les différents types d'expériences mesurent. La formation d'un complexe entre un ligand et une protéine dépend de la différence d'énergie libre de Gibbs (ΔG) entre l'état libre et l'état complexé. Cette différence met en jeu des contributions entropiques et enthalpiques :

$$\Delta G = \Delta H - T\Delta S$$

où T est la température, ΔH la variation d'enthalpie et ΔS la variation d'entropie. Les contributions d'origine entropique peuvent être principalement attribuées aux modifications de conformation de la protéine ainsi qu'aux effets hydrophobes (Chandler, 2005). Les contributions enthalpiques proviennent majoritairement des interactions électrostatiques dont font partie les liaisons hydrogène.

Les expériences sur la stabilité des complexes (MS/MS CID) étant réalisées en phase gazeuse, une partie importante des effets hydrophobes est perdue. Par contre, les contributions d'origine électrostatique sont renforcées. Ayant une constante diélectrique de 80, l'eau diminue la force des liaisons électrostatiques d'un facteur 80 par rapport à leur force en phase gazeuse (Berg *et al.*, 2002). Ainsi, les expériences de type MS/MS CID apportent des informations sur les contributions enthalpiques (renforcées) et entropiques de l'interaction conservées en phase gazeuse (liaisons électrostatiques + contraintes de conformation).

Lors des expériences de compétition, la comparaison de l'intensité des signaux et des stœchiométries des complexes donne une image de l'affinité des tanins pour IB5 (Daniel *et al.*, 2002 ; Gabelica *et al.*, 2003). En effet, les intensités des signaux sur les spectres obtenus sont supposées être représentatives des concentrations des espèces en solution, à la condition que soient respectées les hypothèses énoncées dans la partie III.3. Ainsi, l'échelle obtenue tend à comparer le ΔG de l'interaction, donc l'ensemble des contributions (effets hydrophobes, interactions électrostatiques dont liaisons hydrogène, contraintes conformationnelles de la protéine).

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

Dans la littérature, les interactions PRP-tanin sont généralement décrites comme résultant à la fois d'interactions de type effets hydrophobes et liaisons hydrogène (Spencer *et al.*, 1988 ; Haslam, 1996). Les données obtenues par MS/MS CID ont montré une corrélation entre le nombre d'hydroxyles portés par les tanins et leur stabilité en phase gazeuse. Ce résultat confirme l'intervention de liaisons hydrogène dans les interactions PRP•tanins et est en accord avec les résultats obtenus par Simon *et al.* (Simon *et al.*, 2003), Plet et Schmitter (Plet et Schmitter, 2009) et par Hagerman et Butler (Hagerman et Butler, 1981), pour qui les interactions PRP-tanin résultent principalement de l'établissement de liaisons hydrogène entre les groupements hydroxyles des tanins et les fonctions carbonyles des protéines. Cependant les différences obtenues entre les échelles de stabilité en phase gazeuse des complexes et d'affinité des tanins pour IB5 en solution, invitent à ne pas négliger la part des contributions des effets hydrophobes dans ces interactions.

La principale différence observée concerne les dimères : ils forment des complexes avec IB5 plus stables en phase gazeuse que ceux des monomères (cf. tableau 15), alors qu'ils montrent une affinité pour IB5 moins importante que les monomères galloylés. Ce résultat pourrait provenir soit du fait que ce qui se passe en phase gazeuse n'est pas représentatif de la phase aqueuse, ou soit du fait que la constante d'association des dimères avec IB5 (k_a) est plus faible que celle des monomères. Une constante d'association plus faible pour les dimères pourrait avoir pour origine une pénalité entropique plus importante de l'interaction. En effet, il est possible que la structure des dimères impose des restrictions de conformation plus importantes sur la protéine par rapport aux monomères. Ainsi la contrainte de conformation lors de l'interaction IB5-dimère pourrait être plus importante et pénalisante du point de vue entropique que celle de l'interaction IB5-monomère. Ces réarrangements conformationnels des PRP lors de leurs interactions avec les tanins ont été évoqués par différents auteurs (Charlton *et al.*, 1996 ; Pascal *et al.*, 2007). Ainsi, l'association IB5•dimère serait plus stable que celle IB5•monomère mais avec une pénalité entropique plus importante, ce qui se traduirait au final par une affinité plus faible.

Le couplage de la spectrométrie de masse et de la mobilité ionique est capable de mettre en évidence de telles modifications conformationnelles. Dans la suite de notre travail, nous avons donc abordé les questions relatives aux éventuels changements conformationnels par la mise en œuvre d'expériences utilisant ce couplage. De plus, la structure d'IB5 semblant jouer un rôle important dans ces interactions, nous avons entrepris de comparer les interactions IB5-tanin avec celles d'une protéine modèle avec les tanins.

III. Etude de l'effet de la structure des tanins sur l'interaction IB5-tanin

Chapitre IV. Etude structurale de l'interaction IB5-tanin

IV.1. Introduction

La capacité des tanins d'inhiber l'activité des nombreuses enzymes tendrait à prouver que leur interaction non-covalente avec celles-ci ne présente pas de spécificité. Toutefois, les PRP salivaires sécrétées par les mammifères possèdent une affinité supérieure aux autres protéines pour les tanins (Hagerman et Butler, 1981 ; Hagerman et Klucher, 1986 ; Bacon et Rhodes, 1998). Cette affinité pourrait provenir de la présence de séquences particulières répétées au sein de leur séquence primaire. Plus précisément, les tanins interagiraient de façon préférentielle au niveau des clusters de proline présents au sein de ces zones répétées (Murray et Williamson, 1994 ; Murray *et al.*, 1994). Cependant, il a été mis en évidence que les PRP entières ont une affinité pour les tanins supérieure à celle de peptides composés d'une de ces séquences répétées (Charlton *et al.*, 1996). Ainsi, la présence seule de ces répétitions ne peut expliquer la forte affinité des PRP pour les tanins. Certains auteurs ont évoqué la possibilité de réarrangements conformationnels des PRP survenant au cours de l'interaction (Charlton *et al.*, 1996), la protéine venant s'enrouler autour du tanin pour augmenter la surface d'interaction. Cette hypothèse de réarrangement conformationnel a été étayée pour la première fois sur une PRP totale par des expériences de dichroïsme circulaire (Pascal *et al.*, 2007). Les données indiquent l'acquisition d'un plus haut degré de structuration de la protéine IB5 lorsqu'elle est en interaction avec l'EgCG. De plus, les résultats que nous avons obtenu en MS/MS sur les complexes IB5•tanin suggèrent l'intervention de nombreuses liaisons hydrogène. Les tanins sont considérés comme des ligands possédant plusieurs zones d'interaction représentées par leurs différents noyaux phénoliques et les substituants qu'ils comportent.

Dans ce contexte, nous avons poursuivi notre travail avec pour objectif d'évaluer les trois hypothèses suivantes :

- les PRP ont plus d'affinité que les protéines structurées pour les tanins, entraînant la formation d'un plus grand nombre de complexes protéine•tanin, et permettant un « piégeage » plus efficace des tanins;
- les PRP possèdent des sites d'interaction privilégiés pour les tanins ;
- des modifications de conformation des PRP se produisent lors de leur interaction avec les tanins.

La MS est apparue comme une approche adaptée à notre objectif. En effet, la comparaison de l'affinité d'un ligand envers deux protéines différentes peut être réalisée par la mise en oeuvre de différentes expériences de MS (Daniel *et al.*, 2002). Les développements récents relatifs aux appareils de MS font que des informations structurales sur les édifices supramoléculaires sont maintenant accessibles par MS (Sharon et Robinson, 2007).

Plus particulièrement, l'utilisation de procédés de fragmentation tel que la dissociation par capture d'électrons (ECD, *electron capture dissociation*) permet la localisation des sites d'interaction sur les protéines. Xie et ses collaborateurs ont notamment localisé le site d'interaction de la spermine sur l' α -synuclein, une autre PID, en accord avec les données de RMN (Xie *et al.*, 2006). L'ECD semble donc capable de mettre en évidence la présence d'un ou de plusieurs sites d'interaction privilégiés pour les tanins sur la chaîne peptidique des PRP.

Le couplage de la spectrométrie de mobilité ionique (IMS) avec la spectrométrie de masse (MS) est capable de fournir des informations sur la structure des protéines et sur celle de leurs complexes. Ce couplage est particulièrement avantageux dans les cas où coexistent en solution différents types de complexes (Ruotolo *et al.*, 2005 ; Ruotolo *et al.*, 2007). En effet, à l'inverse d'autres techniques qui ne donnent qu'une image de la moyenne des structures adoptées par les différents objets, l'IMS-MS permet de discriminer la structure individuelle de chacun des objets. Ainsi, cette démarche apparaît particulièrement bien adaptée à l'étude des interactions IB5-tanin. La structure des complexes IB5•tanin pourra être comparée entre les différentes stœchiométries et d'éventuels changements de conformation pourront être mis en évidence pour chacune d'entre elles.

IV.2. Structure des protéines et formation de complexes : IB5-EgCG Vs. BPTI-EgCG

Dans le but de vérifier que les PRP présentent une affinité supérieure aux protéines structurées pour les tanins, nous avons comparé l'interaction IB5-EgCG à l'interaction entre l'EgCG et une protéine structurée.

Une protéine structurée stable aux variations de pH et avec une longueur de chaîne peptidique proche de celle d'IB5 a été retenue. Cette protéine est l'inhibiteur pancréatique bovin de la trypsine (BPTI, Bovine Pancreatic Trypsin Inhibitor). Avec une masse moléculaire de 6511 Dalton correspondant à 58 acides aminés et la présence de trois ponts disulfures,

sa structure s'avère particulièrement stable même à des valeurs de pH acide (Wlodawer *et al.*, 1984 ; Amir et Haas, 1988 ; Dadlez et Kim, 1995).

Cette étude a été réalisée à l'aide d'un spectromètre de masse de type TOF avec réflectron couplé à une source d'ionisation de type électrospray. Cet appareil a été choisi à la fois pour sa gamme dynamique et pour sa résolution qui permet de résoudre les massifs isotopiques d'ions présentant jusqu'à une dizaine de charges. Dans un premier temps, les spectres de la protéine BPTI seule en solution ont été enregistrés (figure 83).

Les spectres de masse obtenus sont composés de trois pics correspondant à BPTI sous trois états de charge: 5+, 6+ et 7+.

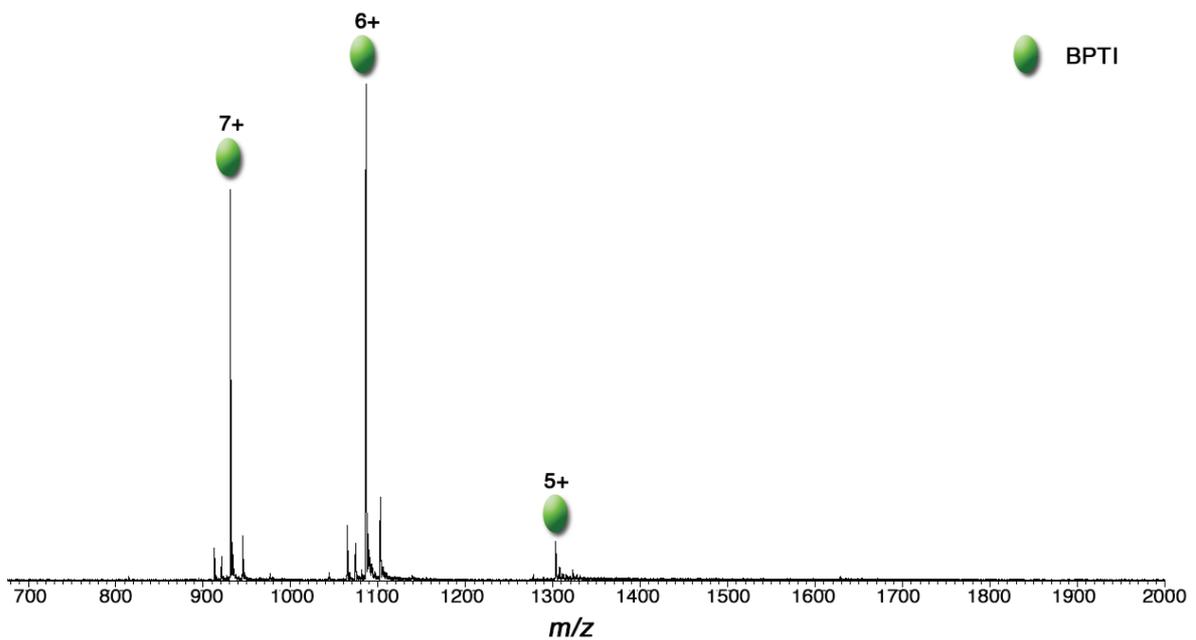


Figure 83. Spectre de masse ESI-TOF d'une solution de BPTI à 5 μ M dans un solvant eau-éthanol 88/12 acidifié à pH 3,2 avec de l'acide acétique.

La comparaison des intensités des signaux sur les spectres individuels de BPTI et d'IB5 montre une intensité nettement supérieure pour les signaux de BPTI. Cette différence de comportement peut être la conséquence du plus grand nombre d'états de charge d'IB5 au même pH (6 états de charge pour IB5 contre 3 pour BPTI) qui répartit l'intensité des signaux. Elle peut également découler de rendements d'ionisation différents entre les deux protéines.

Dans un deuxième temps, nous avons procédé à la comparaison des interactions BPTI-EgCG et IB5-EgCG. L'impossibilité de déterminer les constantes d'affinité dans le cas des interactions protéine-tanin nous conduit à comparer les pourcentages de protéines libres et de protéines complexées. Cette approche est basée sur l'approximation que les intensités

déduites du spectre de masse sont proportionnelles aux concentrations en solution, ce qui est généralement le cas des systèmes protéine-ligand (cf. chapitre III.3 page 201).

Les mélanges IB5-EgCG et BPTI-EgCG ont été préparés à un rapport molaire protéine-tanin de 1/10 (protéine 5 μM et tanin 50 μM). Les solutions ont été réalisées dans un solvant eau-éthanol 88/12 acidifié avec de l'acide acétique jusqu'à pH 3,2. Aucun trouble ni précipité n'a été observé durant la réalisation de ces solutions indiquant que l'ensemble des partenaires est en solution. Les spectres ESI-MS de ces deux solutions ont ensuite été enregistrés. L'ensemble de ces expériences a été réalisé le même jour sur le même spectromètre de masse.

La figure 84 montre le spectre ESI-MS de la solution BPTI-EgCG (5 μM -50 μM) qui peut être comparé à celui de la solution IB5-EgCG (5 μM -50 μM), présenté dans la figure 4 de la publi n°1 (cf. page 165). La résolution des spectres permet d'attribuer les états de charge de chacun des pics et d'en déduire la masse des ions associés. Des pics correspondant aux protéines libres, aux tanins libres et aux complexes protéine•EgCG avec différentes stœchiométries ont été identifiés. Les stœchiométries observées pour les complexes BPTI•EgCG vont de 1•1 à 1•4 (le signal m/z des complexes 1•4 est de très faible intensité et n'est donc pas légendé sur le spectre présenté) tandis que celles des complexes IB5•EgCG vont de 1•1 à 1•5.

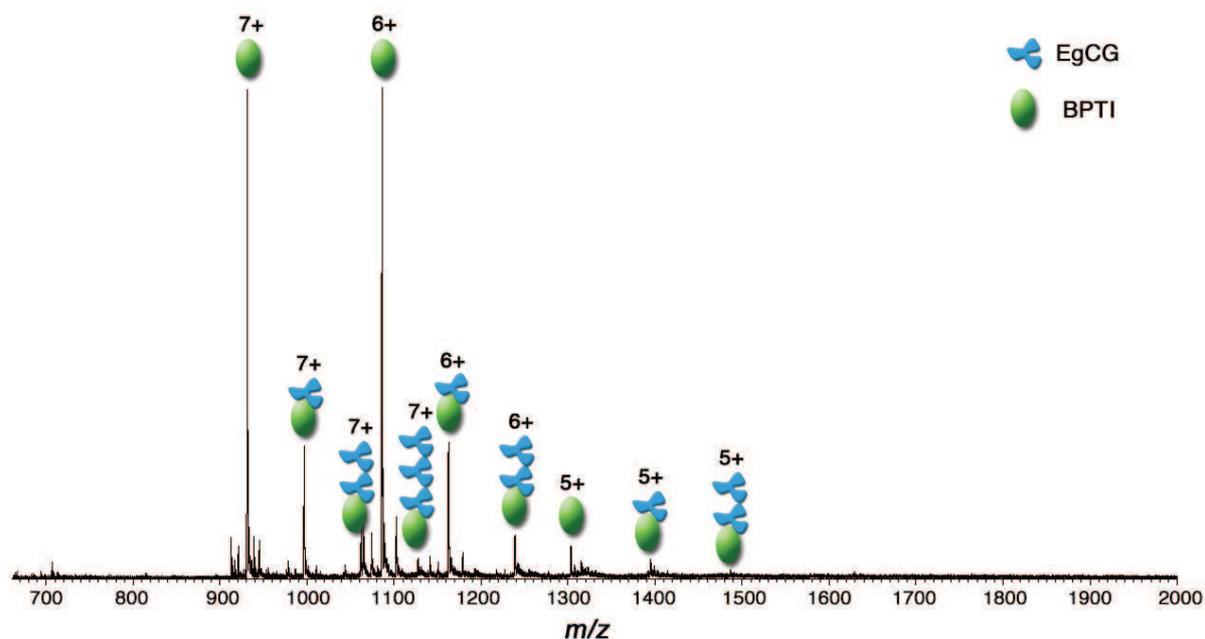


Figure 84. Spectre du mélange BPTI 5 μM - EgCG 50 μM (seuls les pics les plus importants sont légendés).

A partir des spectres et de l'intensité des signaux de chacune des espèces, une estimation du pourcentage de protéine libre et de protéine complexée dans chacun des mélanges est réalisée selon le calcul suivant :

- Pourcentage de $P_{\text{libre}} = \frac{\Sigma \text{ des intensités des signaux de protéine libre}}{(\Sigma \text{ des intensités des signaux de protéine libre} + \Sigma \text{ des intensités des signaux de complexes})} \times 100$.
- Pourcentage de $P_{\text{complexée}} = \frac{\Sigma \text{ intensité des signaux de complexes}}{(\Sigma \text{ des intensités des signaux de protéine libre} + \Sigma \text{ des intensités des signaux de complexes})} \times 100$.

Les résultats de ces calculs sont retranscrits dans le tableau 16.

Tableau 16. Pourcentages de BPTI et d'IB5 libres et complexés calculés à partir des spectres MS des solutions IB5:EgCG et BPTI:EgCG.

	BPTI	IB5
Pourcentage de protéine libre	69,78	39,02
Pourcentage de protéine complexée	30,22	60,98

Le tableau 16 montre que, pour un même rapport molaire protéine/tanin aux mêmes concentrations (5 μM de protéine – 50 μM de tanin), le mélange IB5:tanin présente deux fois plus de protéine complexée que celui BPTI:tanin. Aux concentrations étudiées, IB5 est en majorité complexée, tandis que BPTI est majoritairement présente sous forme libre. IB5 présente donc une affinité pour EgCG supérieure à celle de BPTI. Cette expérience permet de confirmer que les bPRP sont plus affines pour les tanins que les protéines structurées de taille voisine. Ce résultat est en accord avec ceux obtenus précédemment par deux autres études basées:

- 1- sur des expériences de compétition (Hagerman et Butler, 1981), au cours desquelles le taux d'inhibition de la précipitation d'une protéine marquée suite à l'ajout de tanins a été mesuré pour différentes protéines, la protéine marquée et la protéine testée étant en compétition pour interagir avec les tanins ajoutés ;
- 2- sur la comparaison des constantes d'affinité pour différents tanins de protéines immobilisées (Bacon et Rhodes, 1998).

L'origine de l'affinité particulière des PRP pour les tanins a été expliquée tout d'abord par l'originalité de ces protéines. En effet, les fortes teneurs en proline et la présence de

glycine confèrent à ces protéines une structure à la fois ouverte et souple (Hagerman et Butler, 1981 ; Hagerman, 1989 ; Yokotsuka et Singleton, 1995). Leur configuration étendue leur permet de maximiser la surface accessible pour ces interactions (Mehansho *et al.*, 1987a). Le fait que l'atome d'oxygène de la fonction carbonyle lié à l'amine tertiaire de la proline soit un accepteur fort de liaison hydrogène (Hagerman et Klucher, 1986) pourrait également expliquer l'affinité particulière des PRP pour les tanins.

IV.3. Nombre de sites d'interaction et position

IV.3.1. Introduction

Les PRP, et plus particulièrement notre modèle la protéine salivaire humaine IB5, ont donc une affinité particulière pour les tanins mais existe-t-il des sites privilégiés d'interaction des tanins au niveau de leurs séquences ? Et si oui combien ?

Différentes techniques de fragmentation ont été développées depuis les débuts de la spectrométrie de masse (figure 85).

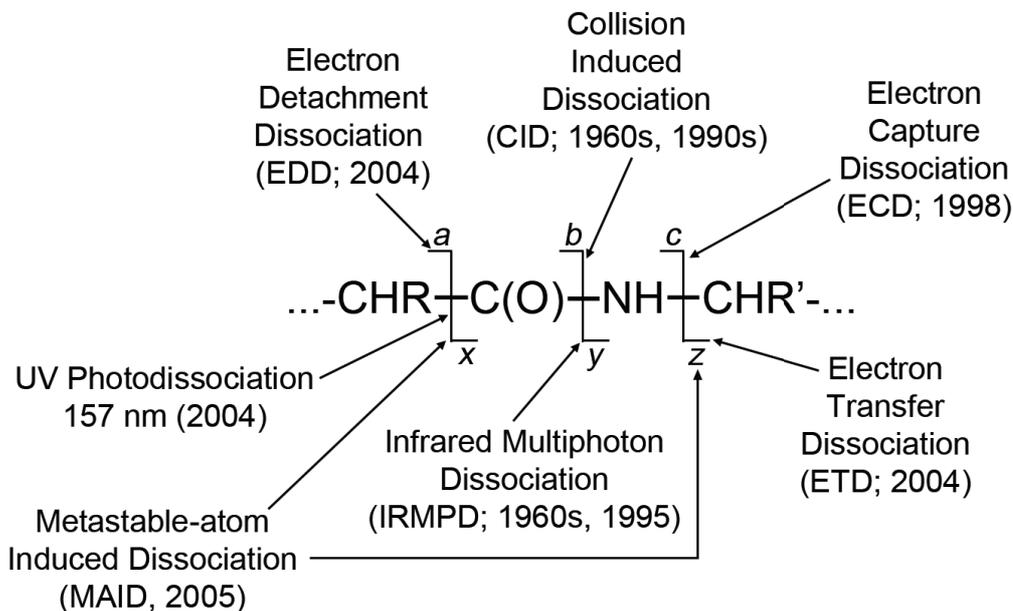


Figure 85. Techniques de fragmentation applicables aux protéines et types de fragments produits.

Au cours des expériences de CID, l'augmentation de l'énergie interne de l'ion sélectionné conduit toujours à la rupture des liaisons les plus faibles en premier. Cette technique permet d'établir les courbes de dissociation de complexes non-covalents en fonction de l'énergie de collision appliquée. L'énergie de demi dissociation des complexes

peut ensuite être comparée. Nous avons ainsi établi les courbes de dissociation des complexes IB5•EgCG de différentes stœchiométries.

L'ECD, contrairement au CID, est basée sur un processus de fragmentation non-ergodique donnant une plus grande diversité de fragments et permettant de rompre la chaîne peptidique sans rompre les liaisons plus labiles (McLafferty *et al.*, 2001). Cette technique a notamment permis la localisation des modifications post-traductionnelles de nombreuses séquences peptidiques (Bakhtiar et Guan, 2006 ; Whitelegge *et al.*, 2007). Elle a également permis de déterminer la localisation des sites d'interaction de diverses protéines avec leur ligand (Xie *et al.*, 2006 ; Jackson *et al.*, 2009 ; Yin et Loo, 2010). Nous avons donc appliqué ce type de fragmentation sur les complexes [IB5•EgCG]⁷⁺ et [IB5•B2 3'OG]⁷⁺ dans l'objectif d'évaluer la présence d'un site d'interaction préférentiel.

IV.3.2. Etude du nombre de sites d'interaction par ESI-MS/MS CID

L'utilisation de fortes concentrations en tanins accroît les stœchiométries des complexes IB5•tanin et le phénomène de « stacking ». Ce terme anglais désigne l'empilement des cycles aromatiques de plusieurs entités flavane. Ces objets peuvent également interagir avec IB5 et conduire à l'augmentation des stœchiométries sans que chaque tanin impliqué dans le complexe interagisse directement avec la protéine. Par ailleurs, on peut aussi imaginer qu'après saturation des sites d'IB5, les tanins excédentaires interagissent avec les tanins déjà fixés. Cela pose la question du nombre de sites de liaison des tanins sur IB5. Pour répondre à cette nouvelle interrogation, nous avons réalisé des expériences de MS/MS CID sur les complexes IB5•EgCG de différentes stœchiométries. La figure 86 présente les courbes de dissociation des complexes de stœchiométries 1•1 à 1•9 présents dans une solution IB5:EgCG 5µM:200µM (rapport molaire de 1/40).

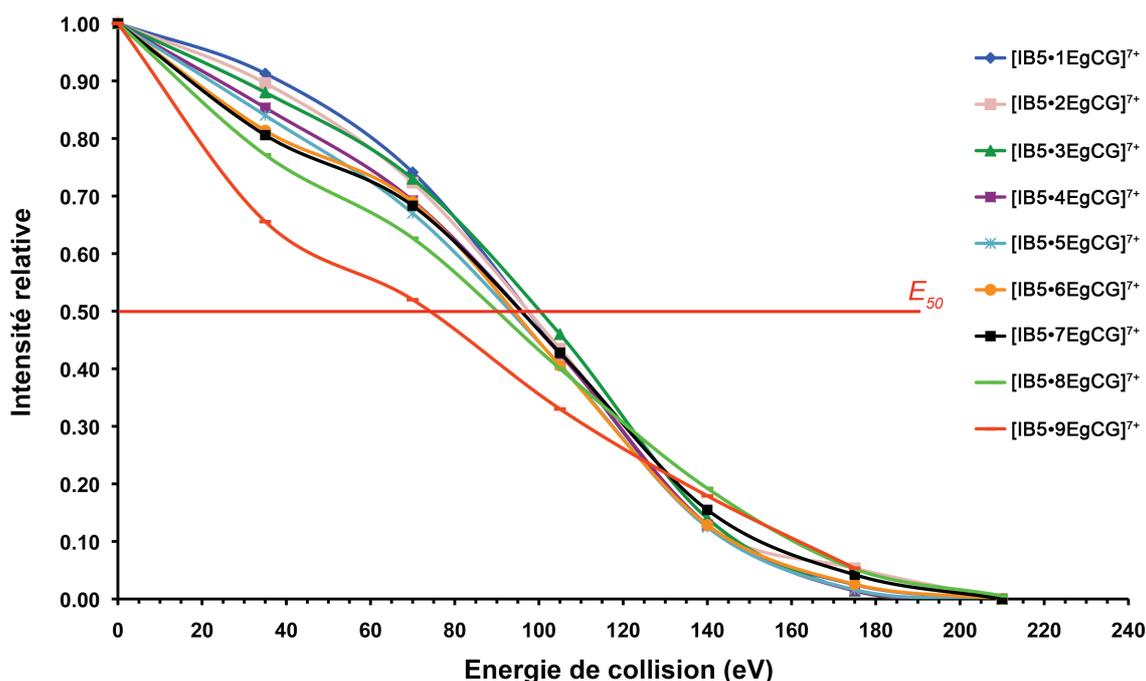


Figure 86. Courbes de dissociation des complexes IB5•EgCG de stœchiométries 1•1 à 1•9 présents dans la solution IB5:EgCG 5 μ M:200 μ M dans le solvant H₂O/EtOH 88/12 à pH 3,2 acidifiée avec de l'acide acétique.

La comparaison des énergies de demi dissociation (E_{50}) des différents complexes montre que les complexes de stœchiométries 1•1 à 1•8 ont des E_{50} proches, alors que le complexe de stœchiométrie 1•9 présente une E_{50} beaucoup plus faible. Ce résultat suggère la présence sur IB5 de huit sites d'interaction dont les énergies de liaison des tanins sont relativement proches. A partir de la saturation des sites d'interaction préférentiels des tanins sur la protéine, un autre type d'interaction interviendrait. Les huit sites équivalents pourraient être interprétés comme les huit clusters de proline présents au sein de la séquence d'IB5, en accord avec les données de la littérature qui décrivent ces motifs comme ayant une affinité particulière pour les tanins (Baxter *et al.*, 1997). La proline serait un constituant majeur d'un site de liaison aux polyphénols à la fois à travers l'oxygène de la fonction carbonyle de ses groupements amides tertiaires qui est un accepteur fort de liaison hydrogène (Haslam, 1996) et par son cycle pyrrolydine formant une surface hydrophobe, ouverte et rigide accessible aux noyaux phénoliques.

Afin de confirmer cette hypothèse, nous avons mis en place des expériences de fragmentations de type ECD pour tenter de localiser les sites d'interaction des tanins sur IB5.

IV.3.3. Etude de la position des sites d'interaction par ESI-MS/MS ECD

L'ECD a été développée uniquement sur les spectromètres de masse de type FT-ICR même si sa variante, la dissociation par transfert d'électrons (ETD, electron-transfer dissociation), a été adaptée depuis à d'autres analyseurs comme les trappes ioniques.

Ce travail a été possible grâce à une collaboration avec le Dr Guillaume Van der Rest du laboratoire des mécanismes réactionnels. Celle-ci nous a permis d'avoir accès aux technologies ECD et FT-ICR, à leur compréhension et à la validation des résultats interprétés.

Trois solutions ont été réalisées dans le solvant H₂O/EtOH 88/12 acidifié à pH 3,2 par de l'acide acétique:

- IB5 à 5 µM,
- IB5:EgCG, 5 µM:50 µM, rapport 1/10 ;
- IB5:B2 3'OG, 5 µM:50 µM, rapport 1/10.

Les spectres MS des ces trois solutions ont été enregistrés. Ils sont similaires aux spectres précédemment obtenus sur les autres spectromètres de masse. Les complexes présentent les mêmes stœchiométries. Les intensités des pics révèlent que l'EgCG a une affinité plus forte pour IB5 que B2 3'OG (figures 87.a, 88.a et 89.a) ce qui est en accord avec les résultats décrits dans le chapitre III.

Les fragmentations MS/MS ECD sont réalisées sur les ions suivants :

- [IB5]⁷⁺ (figure 87.b),
- [IB5•EgCG]⁷⁺ (figure 88.b),
- [IB5•B2 3'OG]⁷⁺ (figure 89.b).

La sélection des ions est effectuée dans la cellule ICR par éjection radio-fréquence des autres ions présents.

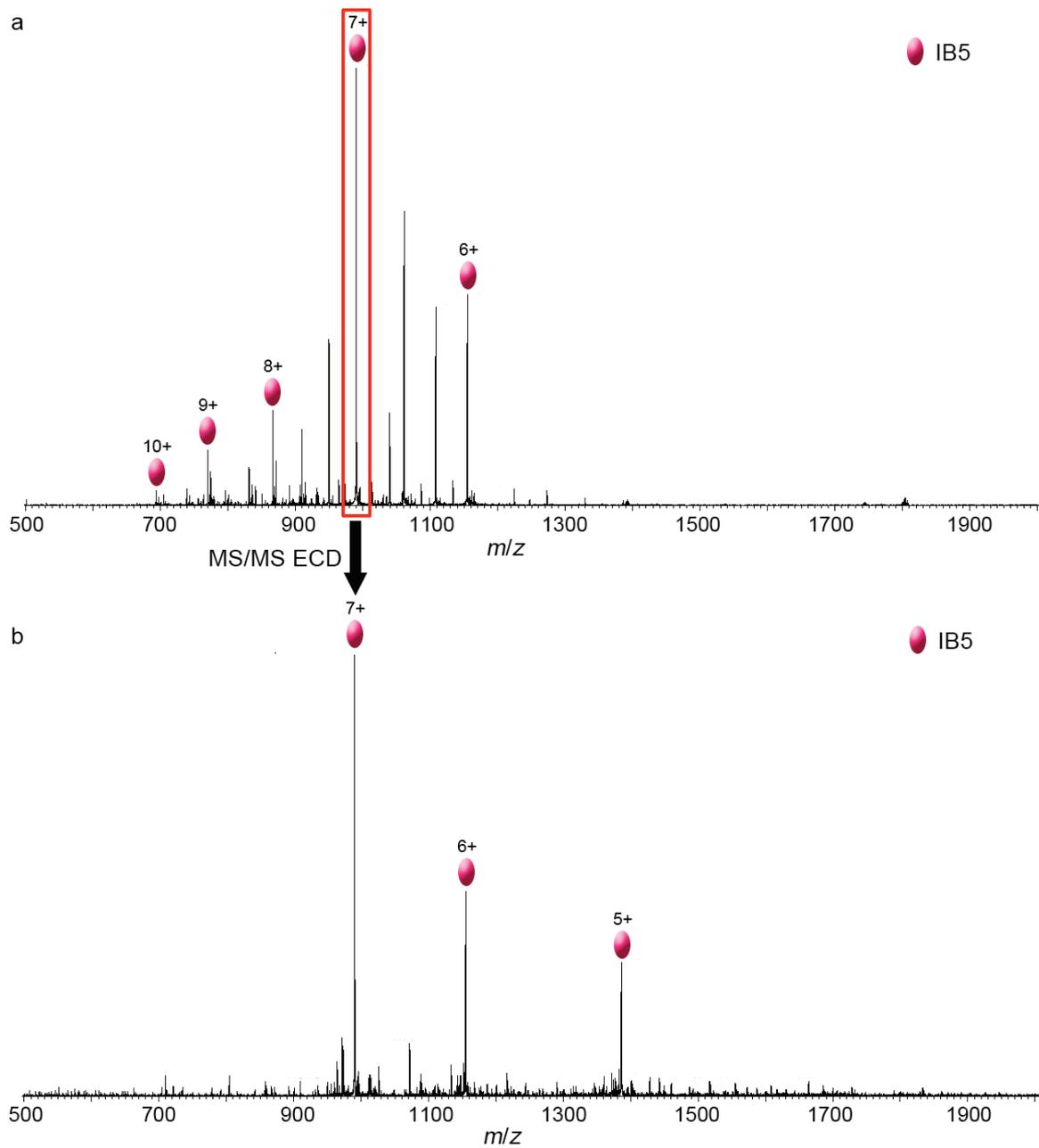


Figure 87. a. Spectre MS de la solution IB5 5 μM ; b. Spectre MS/MS ECD de l'ion parent [IB5]⁷⁺ (m/z 990).

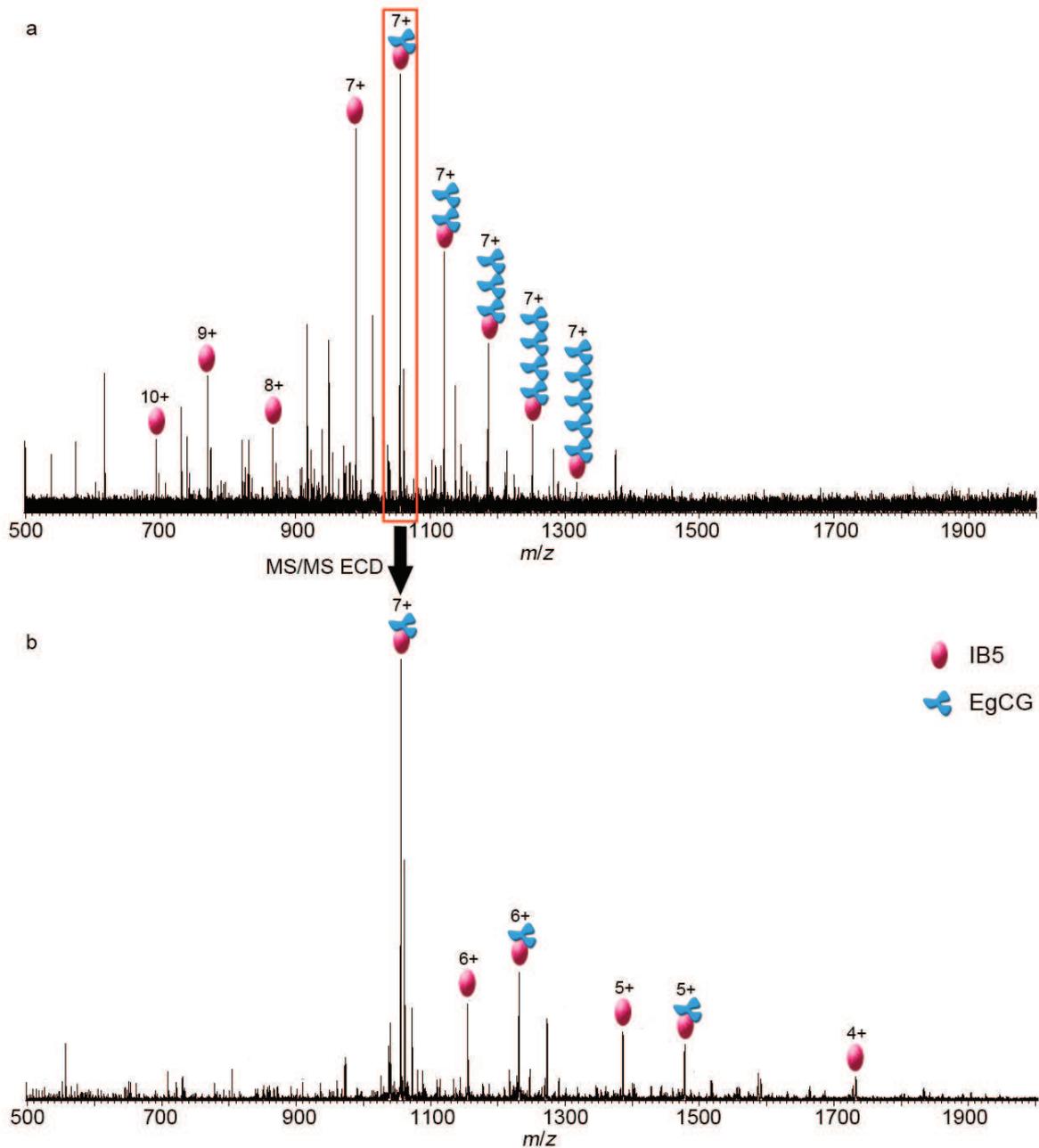


Figure 88. a. Spectre MS de la solution IB5 : EgCG 5 μ M : 50 μ M ; b. spectre MS/MS ECD de l'ion parent $[\text{IB5}\cdot\text{EgCG}]^{7+}$ (m/z 1055).

Sont présents sur les spectres MS/MS ECD:

- l'ion parent à son état de charge initial et n'ayant pas capturé d'électron ;
- des ions parents à des états de charge inférieurs à l'état de charge initial, ceci étant dû à la capture d'un ou de plusieurs électrons,
- des ions fils issus de la fragmentation d'ions parents ayant capturé un ou plusieurs électrons.

La faible intensité des ions fils indique un faible rendement des mécanismes de capture des électrons et de fragmentation de l'ion parent. La haute résolution des analyseurs FT-ICR permet cependant de résoudre les massifs isotopiques d'ions ayant de faibles intensités et ainsi d'identifier précisément les ions fils présents sur les spectres MS/MS ECD.

Le spectre MS/MS ECD du complexe $[IB5 \cdot EgCG]^{7+}$ (figure 88.b) comporte des pics correspondant à la protéine IB5 seule non fragmentée, indiquant une rupture des liaisons non-covalentes liant IB5 et EgCG avant les liaisons covalentes de la chaîne peptidique. La gamme de masse d'acquisition ne permet pas l'observation de l'intensité du pic d'EgCG libre.

Le spectre MS/MS ECD de $[IB5 \cdot B2 \text{ 3'OG}]^{7+}$ (figure 89.b) ne présente pas de pic correspondant à la protéine IB5 libre indiquant, en accord avec les résultats de CID, que les complexes $[IB5 \cdot B2 \text{ 3'OG}]^{7+}$ sont plus stables que les complexes $[IB5 \cdot EgCG]^{7+}$. Toutefois, la présence d'un pic de B2,3'OG libre indique qu'une fraction des molécules de B2 3'OG liées s'est dissociée d'une partie des complexes.

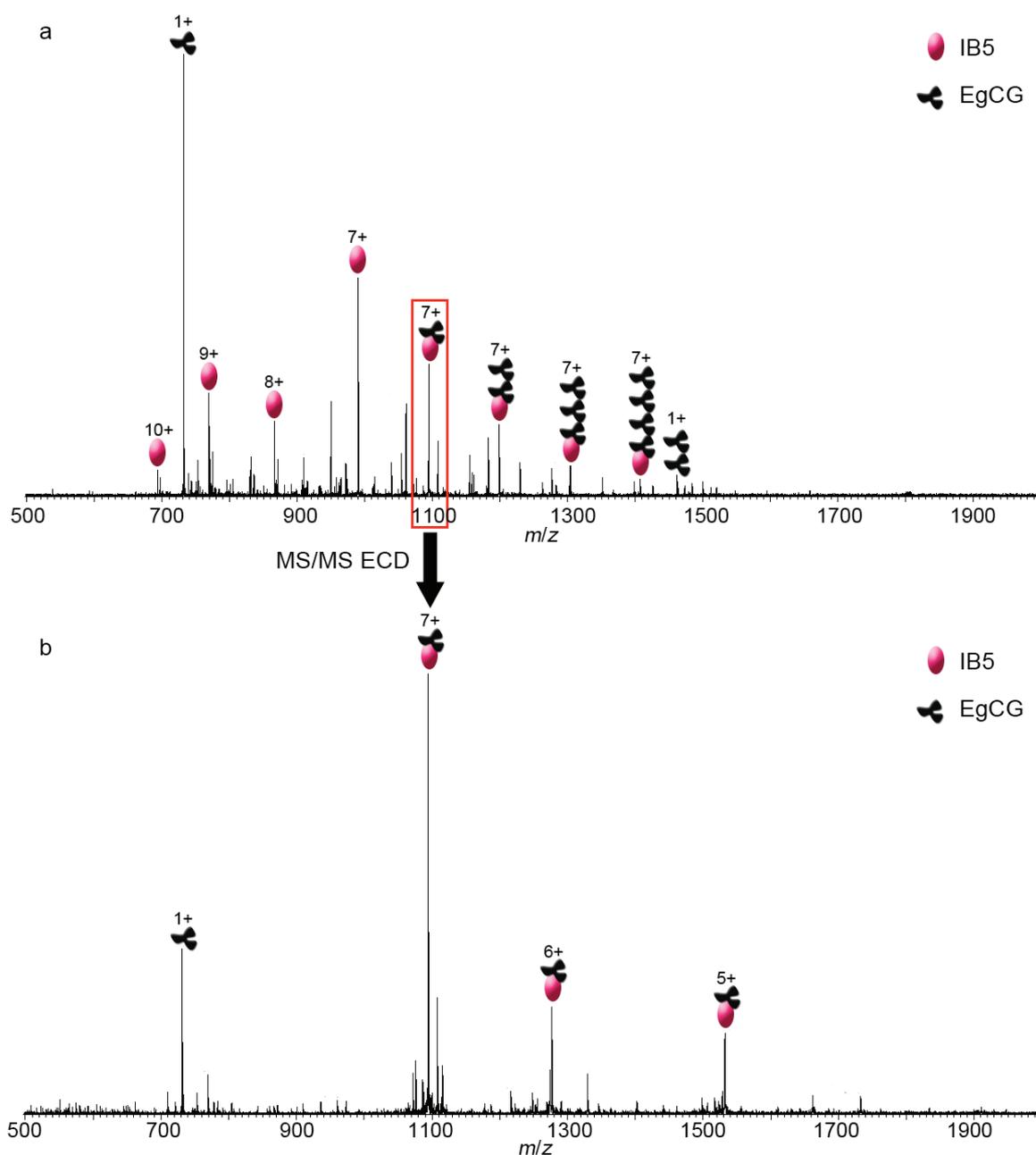


Figure 89. a. Spectre MS de la solution IB5 : B2 3'OG 5 μ M : 50 μ M ; b. Spectre MS/MS ECD de l'ion parent $[IB5 \cdot B2 \text{ 3'OG}]^{7+}$ (m/z 1094).

Les ions fils observés correspondent majoritairement à des fragments de type c et z (figure 85), en accord avec les résultats décrits dans la littérature (Bakhtiar et Guan, 2006). Les fragments obtenus à partir des ions $[IB5]^{7+}$ et $[IB5 \cdot B2 \text{ 3'OG}]^{7+}$ sont décrits sur la figure 90. La couverture de séquence à partir de l'ion $[IB5]^{7+}$ atteint 43%. Cette faible couverture de séquence s'explique par la présence de nombreuses prolines. En effet, le cycle des proline empêche la formation de fragments c et y en position N terminale de ces résidus car la liaison peptidique clivée implique l'azote de ce cycle (figure 91) (Leymarie *et al.*, 2002).

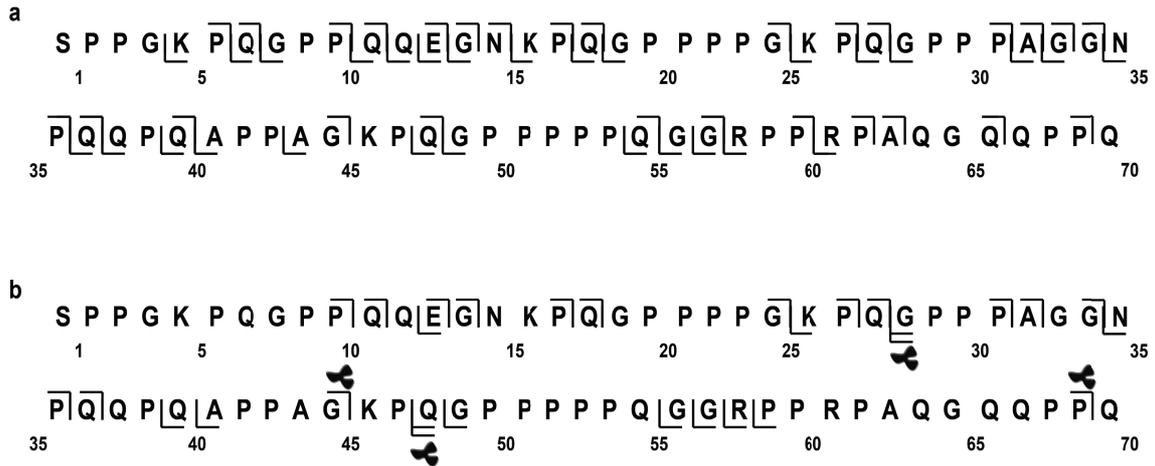


Figure 90. Fragments obtenus en MS/MS ECD pour a. $[IB5]^{7+}$ (m/z 990) et b. $[IB5\cdot 1B2\ 3'OG]^{7+}$. Les fragments ayant retenu le B2 3'OG sont indiqués avec le symbole ⚡.

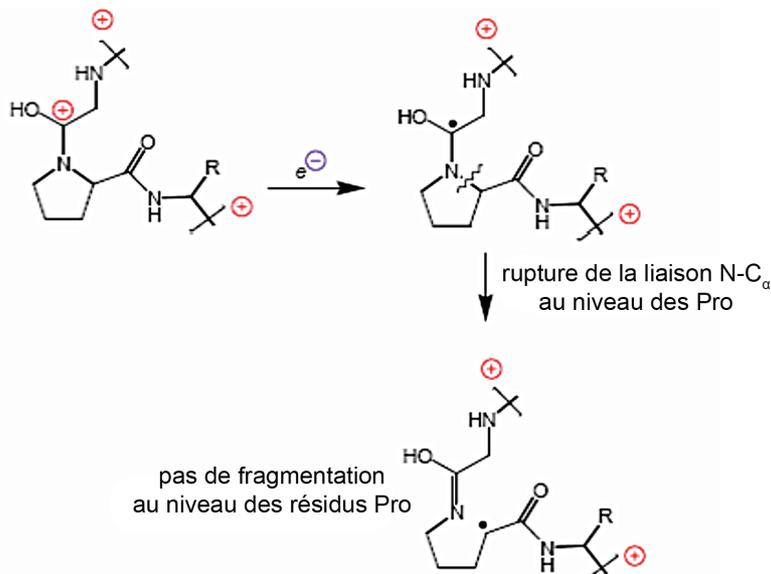


Figure 91. Mécanisme de fragmentation ECD pour un résidu proline.

Le rendement de fragmentation étant faible dans le cas du complexe $[IB5\cdot B2\ 3'OG]^{7+}$, les informations acquises sur IB5 et ses zones d'interactions avec B2 3'OG sont partielles. Néanmoins, des fragments ayant conservé leurs liaisons non-covalentes avec B2 3'OG ont été identifiés. Ces résultats supposent la présence de plusieurs sites de liaison mais ne permettent pas de dégager précisément un site d'interaction privilégié. Ils ne nous ont donc pas permis de valider ou d'invalider l'hypothèse selon laquelle les clusters de proline seraient des sites d'interaction privilégiés des tanins.

Pour expliquer l'affinité particulière des PRP entières pour les tanins, une autre étude a proposé l'hypothèse selon laquelle les PRP entières seraient capables de changements de conformation suite à la fixation des tanins. Ces changements de conformation permettraient de maximiser la surface d'interaction, les auteurs allant jusqu'à invoquer un enroulement des

PRP autour des tanins (Charlton *et al.*, 1996). Notre travail s'est poursuivi par l'étude d'éventuels changements de conformation d'IB5 suite à la fixation des tanins.

IV.4. Etude des modifications de conformation de la protéine IB5 en interaction avec les tanins

IV.4.1. Introduction

Un réarrangement conformationnel d'IB5 pourrait permettre aux tanins fixés sur IB5 d'interagir simultanément avec plusieurs sites de liaison sur la protéine (Baxter *et al.*, 1997). En effet, les tanins sont décrits comme des ligands ayant plusieurs sites d'interaction (Baxter *et al.*, 1997 ; Bennick, 2002) (noyaux aromatiques porteurs de fonctions hydroxyles). Cette hypothèse a été proposée pour expliquer qu'une PRP salivaire entière avait une affinité pour les tanins supérieure à celle d'un peptide constitué d'une seule séquence répétée de la PRP (Charlton *et al.*, 1996).

De plus, des études menées par CD ont montré que la fixation d'EgCG sur IB5 induit chez cette dernière une augmentation de la proportion de structure secondaire indiquant une transition désordre-ordre (Pascal *et al.*, 2007).

La MS couplée à la spectrométrie de mobilité ionique (IMS) a montré sa capacité à révéler les différents états de structuration des protéines (Jarrold, 1998 ; Ruotolo *et al.*, 2007). Ainsi, cette technique devrait permettre de mettre en évidence les éventuels changements de conformation d'IB5 lors de son interaction avec les tanins.

Une collaboration a été développée avec l'équipe de Philippe Dugourd et de Jérôme Lemoine du laboratoire de spectrométrie ionique et moléculaire (LASIM). Cette équipe a conçu et développé un tube de mobilité ionique couplé à un spectromètre de masse de type Q-TOF (microQ-TOF, Bruker, Bremen, Germany).

IV.4.2. Etude des conformations de IB5 et de ses complexes IB5•EgCG par ESI-IMS-MS

Dans un premier temps, la structure de la protéine IB5 seule a été étudiée. Les contraintes de l'appareillage et la sensibilité de l'appareil nécessitent de travailler avec des solutions suffisamment concentrées en analytes. Une solution de la protéine IB5 à une concentration de 50 μ M est analysée.

Cette expérience permet de mesurer les temps de mobilité des différents états de charge de la protéine IB5. Les sections efficaces de collision des états de charge 6+ à 9+ d'IB5 sont ensuite calculées (figure 92).

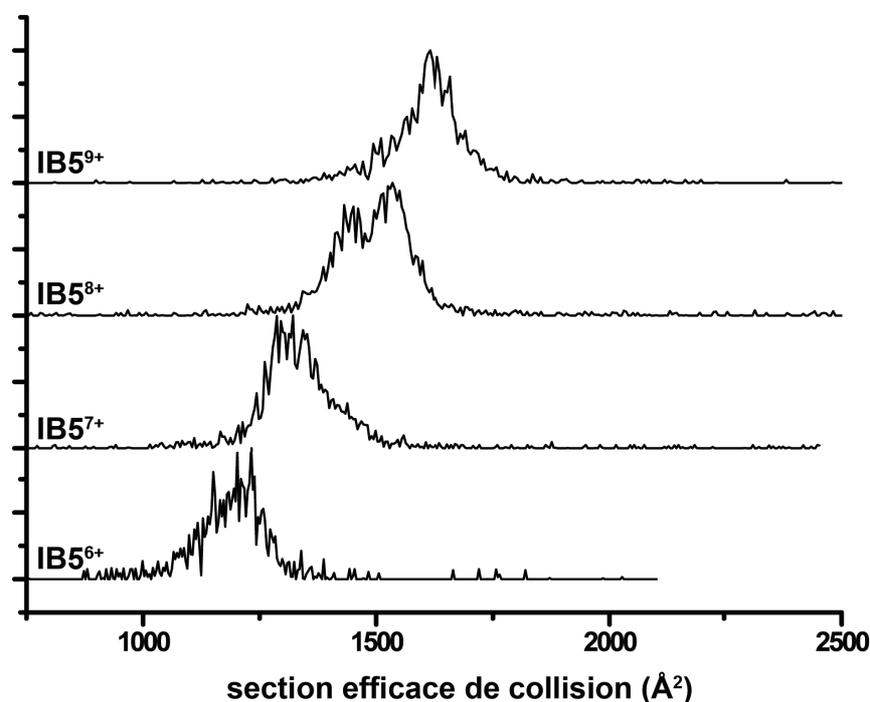


Figure 92. Sections efficaces de collision des états de charge 6+ à 9+ de la protéine IB5.

La section efficace de collision d'IB5 à l'état de charge 7+ est de 1350 Å². Cette valeur est plus importante que celle de la protéine BPTI au même état de charge (1050 Å²). La section efficace d'IB5 obtenue est en accord avec une conformation ouverte et dépliée attendue dans le cas d'une IDP.

La figure 92 met en évidence une augmentation de la section efficace de collision d'IB5 avec l'augmentation de son nombre de charges. Cette évolution est provoquée par l'augmentation proportionnelle des forces de répulsion coulombiennes avec le nombre de charges. Le pic de section efficace de collision de [IB5]⁹⁺ est plus étroit que ceux des autres états de charge indiquant qu'un nombre plus restreint de conformations est accessible à IB5. Ainsi, les forces de répulsions coulombiennes contraignent IB5 dans des conformations étendues.

Les résultats obtenus lors de cette étude ont fait l'objet d'une publication soumise à la revue internationale à comité de lecture *Journal of the American Chemical Society* (publication n°3).

Publication n°3 : Folding of a Salivary Intrinsically Unstructured Protein Upon Binding to Tannins

Francis Canon,³ Renaud Ballivian,^{†‡} Fabien Chirot,[†] Rodolphe Antoine[‡]
 Pascale Sarni-Manchado,¹ Jérôme Lemoine,⁴ Philippe Dugourd⁵

INRA-UMR 1083 Science Pour l'Oenologie, Polyphenol Interaction, Bât 28, 2 place Viala F-34060 Montpellier France and
 Université de Lyon, F-69622, Lyon, France; Université Lyon 1, Villeurbanne; CNRS

RECEIVED DATE (automatically inserted by publisher); dugourd@lasim.univ-lyon1.fr

³ INRA

⁴ Lyon, LSA, UMR5180

⁵ Lyon, LASIM, UMR5579

The basic concept regarding the activity of proteins was dominated by the view that proteins need to have a well-defined three-dimensional structure for their function. However, during the last decade this view had changed,¹ and the sequence-to-structure-to-function paradigm had to be re-assessed. Numerous proteins do have a well-defined function that requires intrinsic disorder, and the term of Intrinsically Unstructured Protein (IUP) has been proposed to describe this group.² IUPs appear to be rather common in living organisms especially in higher eukaryotes: primary sequence analyses indicate that about 10–20% of full-length proteins belong to this class and that 25–40% of all residues fall into such regions. The intrinsic lack of structure can confer functional advantages like the ability to bind several ligands. Moreover, many IUPs undergo coupled binding and folding processes that are intrinsically related to their dedicated function.^{1–3} Now, the ability of IUPs to bind several ligands and in some cases to aggregate in large supramolecular edifices leads to mixtures of complexes with different stoichiometries and heterogeneous conformations. While several strategies are emerging for the characterization of flexible proteins,⁴ the study of mixed IUP-target complexes and of their assembly processes is highly challenging and requires new approaches. Mass spectrometry combined to ion mobility (IMS) allows determination of complex distributions and the qualitative structure of each of the species.⁵ For example, this method was recently applied with success to a number of the amyloid-beta protein isoforms of A beta 40 and A beta 42.⁶ Here, we explore for the first time the potentiality of IMS to study the structure of IUPs and of their complexes.

We investigated the interactions between the human salivary proline-rich protein IB5 and a model of wine and tea tannin: epigallocatechin gallate (EgCG) (IB5 sequence and chemical structure of EgCG are given in the supporting information).^{7,8} IB5 is a human IUPs member, which only known function is to bind and to scavenge tannins. Tannins are phenolic compounds ubiquitous in plant and plant-based food. Their non-covalent interactions with salivary PRPs are thought to be responsible for astringency.⁹ From a biological point of view, tannins act as protecting molecules towards herbivore species. This role has been attributed to their ability to interact with proteins, to precipitate them and to inhibit gastrointestinal enzymes thereby reducing the digestibility of plant proteins.¹⁰ The involvement of PRPs in adaptation to tannin diet either as

astringency mediator or/and scavenger is related to their ability to establish non-covalent interactions with tannins. Tannin-IB5 complexes with several stoichiometries have already been observed.⁸ In the present work, we evidenced an unfolded to folded transition of IB5 protein induced by a structural reorganization upon tannin binding.

Epigallocatechin gallate (EgCG) was purchased from Sigma (Sigma Chemical Co., Poole, Dorset, U.K.). The IB5 human salivary proline rich protein was produced by use of the yeast *Pischia pastoris* as a host organism and purified as previously described.⁸ EgCG and protein stock solutions were prepared in the following medium: water/ethanol, 88:12 (v/v) acidified to pH 3.2 with acetic acid, which corresponds to the mouth conditions in the presence of red wine. MS and IMS spectra were obtained with a home-built ESI ion mobility mass spectrometer.¹¹ Ions passed through a drift cell, a quadrupole, a collision cell and were accelerated into a perpendicular reflectron time-of-flight. 2-D maps (drift time, m/z) are recorded. The mass spectrum obtained by electrospraying the protein solution displayed a series of peaks corresponding to protonated isoforms of IB5 (a 6923.70, b 6642.63 and c 6360.39 Da) with charged states ranging from +5 to +10. For interaction studies, IB5 and EgCG solutions were combined at 1:20 (IB5: EgCG molecular ratio). Though crowded, the mass spectrum reveals free IB5 and IB5 supramolecular complexes with stoichiometries ranging from 1 up to 14 tannins (Figure 1). A Poisson-like distribution is observed which is in favor of a non cooperative sticking process.

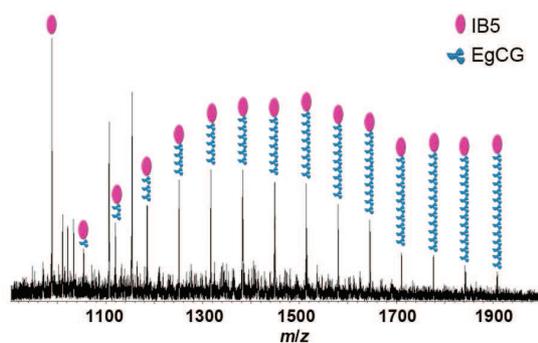


Figure 1. Positive ion mass spectrum of 1:20 IB5:EgCG interaction mixture. Peaks for the +7 charge state of IB5a and its complexes are labelled.

The IMS contribution makes possible the conformational study of each complex and the resolution of eventual conformers showing similar m/z .

For ion mobility experiments, ion packets are injected into the one meter long drift cell containing ≈ 10 Torr of He and travel through the influence of a uniform electric field ($770 \text{ V}\cdot\text{m}^{-1}$). Species are separated in drift-time according to their diffusion cross-section. Compact ions display shorter drift-times than ions with extended structures. Figure 2 displays collision cross sections of selected stoichiometries recorded for charge state +7. For the naked protein, one broad peak is observed, which corresponds to a collision cross section centered at 1350 \AA^2 . This value is larger than the one (1050 \AA^2) measured for the $(M+7H)^{+7}$ ion of BPTI, a globular structured protein with close molecular weight used here as a standard. The collisional cross section of IB5 is then in agreement with an open conformation expected for a proline-rich protein. Binding of 1 to 7 EgCG ligands does not modify the general shape of the cross-section profile. As expected, a regular increase in the collision cross section is observed as the number of ligands increases. From 10 ligands, a second feature appears in the collision cross section profile beside the one following the initial trend from the naked protein. The two features correspond to two families of structures that do not interconvert during the travel through the drift cell (drift time of ~ 60 ms). The second one is shifted to lower collisional cross section values (1200 \AA^2). Thereby, this new feature corresponds to more compact conformations. Though the mass of IB5 complexed with 13 EgCG (12877 Da) is twice the mass of the naked protein, the collision cross section of this more compact oligomer is smaller than the protein one. Interestingly the cross section of the compact state does not increase as the number of bound tannins increases that is different from what is observed for the expanded structures. The ratio of folded to unfolded states increases with the stoichiometry indicating the involvement of several tannins to stabilize the folded state. The conformation transition may allow the efficient binding of numerous tannins by a single IB5 protein. It would then explain the irreversible trapping of tannins that prevent them from interacting with other proteins such as gastrointestinal enzymes.

In summary, IMS was used to get a snapshot of the conformational distribution within a complex mixture of oligomers of one model of IUP protein bound to its natural target and of its conformational adaptability. The data presented here demonstrates that IB5 undergoes a structural transition upon binding with EgCG. This is in agreement with what was inferred from circular dichroism experiments, which have demonstrated conformational transitions during the PRP-tannin interactions in solution.⁷ In the near future, IMS could be part of integrated approaches, which have to be developed to evolve from a static picture of functional IUPs to a dynamic one, in which several conformations are consistent with various aspects of function.

Supporting Information Available: Chemical structure of EgCG and sequence of IB5. Additional cross section distributions.

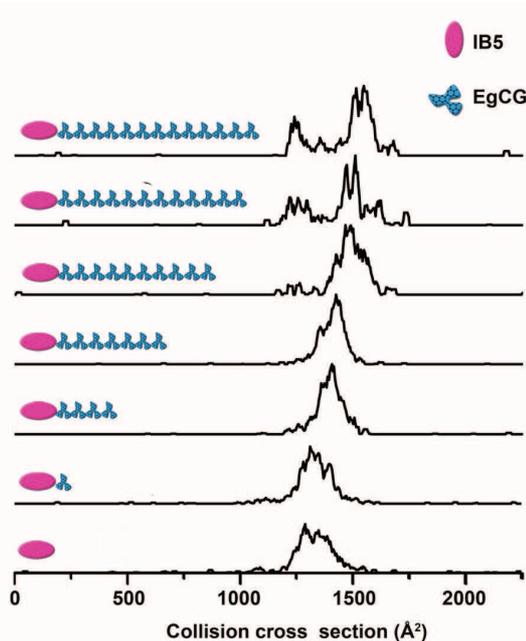
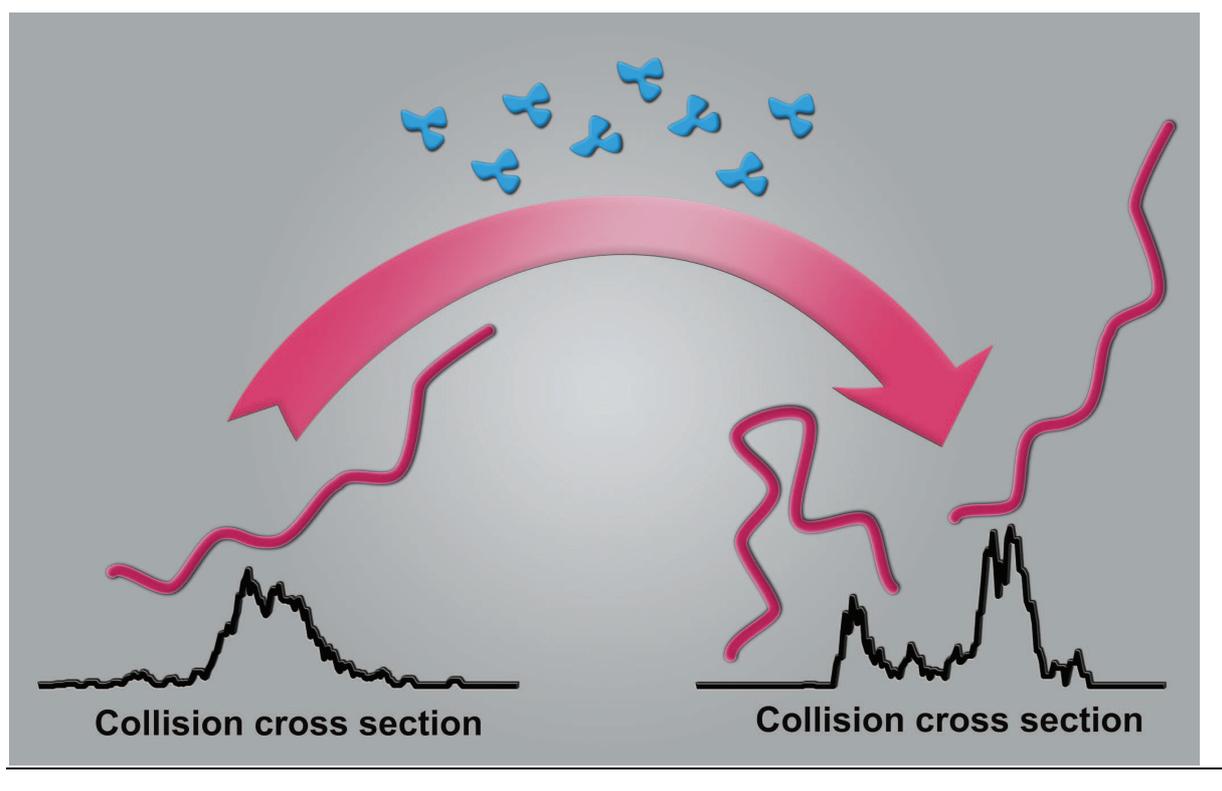


Figure 2. Collision cross section distributions for the +7 charge states of IB5a and IB5a•N EgCG complexes (N=0, 1, 4, 7, 10, 12 and 13)

References

- (1) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321-331.
- (2) Tompa, P. *Trends Biochem. Sci.* **2002**, *27*, 527-533; Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197-208; Cortese, M. S.; Uversky, V. N.; Keith Dunker, A. *Prog. Biophys. Mol. Biol.* **2008**, *98*, 85-106.
- (3) Meszaros, B.; Tompa, P.; Simon, I.; Dosztanyi, Z. *J. Mol. Biol.* **2007**, *372*, 549-561.
- (4) Cszimok, V.; Szollosi, E.; Friedrich, P.; Tompa, P. *Molecular Cellular Proteomics* **2006**, *5*, 265-273; Bernado, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. *J. Am. Chem. Soc.* **2007**, *129*, 5656-5664.
- (5) Bowers, M. T.; Kemper, P. R.; Vonhelden, G.; Vankoppen, P. A. M. *Science* **1993**, *260*, 1446-1451; Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 577-592; Dugourd, P.; Hudgins, R. R.; Clemmer, D. E.; Jarrold, M. F. *Reviews of Scientific Instrument* **1997**, *68*, 1122-1129.
- (6) Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplow, D. B.; Shea, J.-E.; Ruotolo, B. T.; Robinson, C. V.; Bowers, M. T. *Nature Chemistry* **2009**, *1*, 326-331; Murray, M. M.; Bernstein, S. L.; Nyugen, V.; Condron, M. M.; Teplow, D. B.; Bowers, M. T. *J. Am. Chem. Soc.* **2009**, *131*, 6316-6317.
- (7) Pascal, C.; Poncet-Legrand, C.; Imberty, A.; Gautier, C.; Sarni-Manchado, P.; Cheynier, V.; Vernhet, A. *J. Agric. Food Chem.* **2007**, *55*, 4895-4901.
- (8) Canon, F.; Paté, F.; Meudec, E.; Marlin, T.; Cheynier, V.; Giuliani, A.; Sarni-Manchado, P. *Analytical and Bioanalytical Chemistry* **2009**, *395*, 2535-2545.
- (9) Jobstl, E.; O'Connell, J.; Fairclough, J. P. A.; Williamson, M. P. *Biomacromolecules* **2004**, *5*, 942-949.
- (10) Zucker, W. V. *The American Naturalist* **1983**, *121*, 335-365.
- (11) Joly, L.; Antoine, R.; Albrieux, F.; Ballivian, R.; Broyer, M.; Chirot, F.; Lemoine, J.; Dugourd, P.; Greco, C.; Mitric, R.; Bonacic-Koutecky, V. *Journal of Physical Chemistry B* **2009**, *113*, 11293-11300.



We used ion mobility spectrometry to explore conformational adaptability of intrinsically unstructured proteins bound to their targets in complex mixtures. We investigated the interactions between a human salivary proline-rich protein IB5 and a model wine and tea tannin : epigallocatechin gallate (EgCG). The data presented here demonstrate that IB5 undergoes an unfolded to folded structural transition upon binding with EgCG.

IV.5. Conclusions et perspectives

Les tanins sont connus pour être capables d'interagir avec de nombreuses protéines. Cependant, nous avons mis en évidence, au cours des expériences réalisées dans ce chapitre, que la PRP IB5 a une affinité pour EgCG supérieure à celle de la protéine structurée BPTI. Dans le chapitre II, nous avons également montré que les tanins impliqués dans les complexes IB5•tanin de stœchiométrie 1•2 semblent interagir sur des « sites d'interaction » équivalents. Ces résultats associés à l'hypothèse que les séquences répétées des PRP sont des unités fonctionnelles adaptées à la fixation des tanins (Tompa, 2003a) nous ont donc conduits à explorer la possibilité de sites d'interaction privilégiés pour la fixation des tanins sur IB5.

Les expériences de MS/MS CID conduites sur les complexes IB5•EgCG de stœchiométries 1•1 à 1•9 suggèrent la présence de huit sites d'interaction. Les forces d'interaction mises en jeu sont très proches. A partir de la stœchiométrie 1•9, un autre régime d'interaction semble se mettre en place. Les tanins seraient impliqués dans le complexe avec une énergie de liaison plus faible. Les huit sites d'interaction privilégiés saturés, les tanins supplémentaires viendraient interagir avec les tanins déjà fixés sur la protéine.

Les sites d'interaction putatifs sont les séquences répétées « KPQGPP » (Tompa, 2003a) et les clusters de proline (Hagerman et Butler, 1981 ; Charlton *et al.*, 2002a ; Simon *et al.*, 2003). Ils sont respectivement présents au nombre de 5 et de 8 sur la séquence primaire d'IB5. Le motif « KPQGPP » présentant un cluster de deux prolines, les deux catégories de sites se recouvrent partiellement.

L'affinité particulière des tanins pour ces deux types de séquence serait due au fait que l'atome d'oxygène de la fonction carbonyle liée à l'amine tertiaire de la proline est un accepteur fort de liaisons hydrogène (Wolfenden, 1983 ; Hagerman et Klucher, 1986 ; Cheek et Lilley, 1988 ; Fernandez et Lilley, 1992). Ce caractère se trouve renforcé par le groupement méthylène lié à l'atome d'azote de la fonction amide tertiaire qui est donneur d'électrons. De plus, les amides tertiaires sont plus faiblement solvatées que les amides secondaires de la chaîne peptidique. Ainsi, l'établissement d'une liaison hydrogène entre un ligand et l'oxygène d'une fonction carbonyle adjacente à une amide tertiaire demanderait de rompre moins de liaisons hydrogène que si cette liaison s'établissait au niveau d'une fonction carbonyle adjacente à une amide secondaire (Haslam, 1996). Les prolines, en particulier lorsqu'elles sont au moins deux à la suite, fourniraient des sites d'interaction particulièrement bien adaptés à la géométrie des noyaux phénoliques des tanins (Baxter *et al.*, 1997). Les

prolines assureraient alors une double fonction dans l'interaction des PRP avec les tanins. Elles constitueraient des sites d'interaction privilégiés pour les noyaux phénoliques où sont localisés de forts accepteurs de liaisons hydrogène et maintiendraient la structure de la chaîne peptidique étendue maximisant la surface disponible pour interagir avec les tanins (Williamson, 1994).

Les résidus glycine adjacents aux résidus proline sont également susceptibles d'établir des liaisons avec les noyaux phénoliques des tanins (Baxter *et al.*, 1997). Les résultats obtenus par Hatano et Hemingway indiquent que les cycles aromatiques des tanins interagissent de façon préférentielle avec les substituants hydrophobes des autres résidus. (Hatano et Hemingway, 1996). Ces résultats soulignent que les interactions gérées par des effets hydrophobes ont également une grande importance dans la formation ou le maintien des complexes. L'établissement des liaisons PRP-tanin serait du à la fois à des effets hydrophobes et à des liaisons hydrogène.

L'ECD permet de réaliser des fragmentations de la chaîne peptidique tout en maintenant les liaisons non-covalentes. Les résultats que nous avons obtenus en appliquant ces méthodes de fragmentation ont confirmé l'existence de plusieurs sites d'interaction dont l'affinité pour les tanins est voisine. Toutefois, la localisation précise des sites d'interaction n'a pu être établie avec certitude.

Une meilleure localisation des sites d'interaction des tanins sur IB5 pourrait passer par l'étude des complexes de stœchiométries plus élevées pour lesquels la protéine serait « saturée » en tanins. Cela est réalisable par l'emploi de solutions plus concentrées en tanin pour lesquelles les hautes stœchiométries sont favorisées. L'utilisation d'une autre configuration d'appareil est également à envisager. Par exemple, un appareil FT-ICR équipé d'un filtre de masse quadripolaire entre la source d'ionisation et la cellule ICR permet d'accumuler un plus grand nombre d'ions parents dans la cellule ICR. Outre une meilleure statistique au niveau des fragments obtenus, cette configuration éliminerait la nécessité d'une sélection des ions d'intérêt par éjection radiofréquence des autres ions tout en augmentant l'énergie interne des ions parents. L'augmentation de l'énergie interne (énergie vibrationnelle) des complexes fragilise les liaisons non-covalentes avant l'irradiation des ions par le faisceau d'électrons. L'utilisation combinée de solutions plus concentrées et de la configuration d'appareil décrite peut permettre de délimiter plus précisément les différents sites d'interactions.

Les expériences de mobilité ionique ont permis de montrer des modifications de la conformation d'IB5 suite à sa liaison avec au moins dix molécules d'EgCG. Le complexe formé adopte une structure plus compacte que celle de la protéine seule. Ce résultat est en

accord avec les données de dichroïsme circulaire indiquant l'acquisition d'un plus haut degré de structuration d'IB5 suite à son interaction avec EgCG (Pascal *et al.*, 2007). L'adoption de formes plus compactes pourrait permettre à IB5 de lier plus efficacement les tanins. Cela pourrait expliquer le piégeage irréversible des tanins par les PRP. Les PRP en piégeant de cette façon les tanins au niveau de la cavité buccale empêcheraient ces derniers d'interagir avec les PRP glycosylées puis avec les enzymes du tractus gastrointestinal, permettant à ces protéines de remplir leurs fonctions.

La suite de cette étude sur les conformations et les contraintes structurales liées à l'interaction avec des tanins devra envisager l'étude de l'interaction avec le B2 3'OG. Dans une première étape, les expériences de mobilité ionique avec le B2 3'OG devraient permettre de savoir si l'interaction d'un dimère avec IB5 conduit au même phénomène. De plus, nous avons montré que l'encombrement stérique supérieur du dimère conduit à des complexes IB5•tanin ayant des stœchiométries moins élevées par rapport à celles des complexes impliquant l'EgCG. Cela pourrait se traduire par des modifications de la conformation d'IB5 plus ou moins marquées et cela à partir de stœchiométries plus faibles. Il serait également intéressant d'étudier les éventuels changements de conformation d'IB5 suite à la fixation de tanins ayant de plus hauts degrés de polymérisation.

Conclusion générale et perspectives

L'étude des interactions entre les protéines riches en proline et les tanins s'insère dans une problématique générale dont l'objectif final est d'arriver à une meilleure compréhension des phénomènes à l'origine de la sensation d'astringence. Les tanins, molécules capables de générer cette sensation, sont des métabolites secondaires ubiquitaires au sein du règne végétal. Ainsi, l'astringence est une des principales caractéristiques organoleptiques de nombreux produits agro-alimentaires d'origine végétale. Il est donc important de disposer de connaissances plus approfondies sur les phénomènes intervenant dans cette sensation afin de pouvoir les contrôler au cours des procédés d'élaboration des produits. Le collage fait partie de ces traitements technologiques et permet une modulation de l'astringence des vins. Il met en jeu des gélamines porcines définies comme des PRP archaïques. Caractériser les interactions qui s'établissent entre les PRP et les tanins permettra de maîtriser cette modulation, ce qui représente un enjeu fort pour la filière agro-alimentaire et plus particulièrement pour la filière viti-vinicole.

Différentes observations indiquent que les PRP salivaires sont sécrétées en réponse à une alimentation riche en tanins (Mehansho *et al.*, 1987b) et présentent une affinité particulière pour ces derniers (Hagerman et Butler, 1981). Elles appartiennent à une classe de protéines récemment décrite, les protéines intrinsèquement déstructurées (PID) (Dunker *et al.*, 2000b ; Dunker *et al.*, 2001). La découverte de cette catégorie de protéines a conduit à une révision du paradigme « séquence-structure-fonction » (Dunker et Obradovic, 2001), l'absence d'une structure tridimensionnelle définie peut également être à l'origine de la (des) fonction(s) d'une protéine. De par leurs caractéristiques atypiques, ces protéines posent un certain nombre de problèmes aux approches classiquement utilisées en biologie structurale. Par exemple, la cristallographie est incapable de fournir des informations sur ces protéines très mobiles et réfractaires à la cristallisation. « Les expérimentateurs qui s'intéressent à ces dernières doivent déployer mille ruses expérimentales et combiner plusieurs techniques biophysiques pour arriver à les caractériser avec précision. C'est à l'issue de nombreuses adaptations expérimentales seulement que les protéines ou leurs régions désordonnées ont pu être étudiées » (Receveur-Bréchet et Karlin, 2005). L'étude de ces protéines nécessite donc une approche multi-technique. C'est sur cette base qu'a été élaboré le projet PROTANIN, soutenu par l'ANR. Ce projet inclut les approches suivantes : la RMN, le CD, le SAXS, la DLS, la MS et la modélisation moléculaire. Le travail décrit dans ce mémoire étudie l'interaction entre les tanins et la PRP salivaire non structurée IB5 par spectrométrie

de masse et montre les apports de cette technique à l'étude des phénomènes impliquant ces molécules flexibles que sont les PRP et les tanins.

L'apparition des sources d'ionisation douces (ESI et MALDI) a permis l'entrée de la MS dans les laboratoires de biologie. La MS s'est révélée être une technique de choix pour l'étude des échantillons biologiques. En effet, couplée à ces sources d'ionisation, la MS permet l'obtention des spectres de masse des macromolécules biologiques intactes : protéines, ADN, ARN, glucides, lipides, polymères... (McLafferty, 2008). Mais elle a aussi contribué à l'essor de la protéomique, en permettant l'identification et la caractérisation des protéines (Aebersold, 1993). Son utilisation s'est étendue progressivement à la biologie supramoléculaire et à l'étude des complexes non-covalents (Daniel *et al.*, 2002 ; Aebersold et Mann, 2003 ; Sanglier *et al.*, 2005 ; Benesch *et al.*, 2007). Elle autorise aujourd'hui l'étude de complexes ayant des tailles allant jusqu'à l'ordre du mégadalton (Zhou *et al.*, 2008). Elle est capable de fournir des informations sur leur structure, leur composition, leur(s) stœchiométrie(s) et leur dynamique d'interaction. Le couplage de la spectrométrie de mobilité ionique et de la MS (IMS-MS) apporte des informations sur les conformations des protéines et leur dynamique (Clemmer et Jarrold, 1997 ; Ruotolo *et al.*, 2008). L'utilisation de ces différents outils basés sur la MS nous a permis de rassembler, à travers les résultats obtenus, de nouvelles informations concernant les partenaires des interactions et les complexes non-covalents qu'ils sont susceptibles de former.

A l'issue de leur production par la souche *Pichia Pastoris-PRB4S* et de leur purification, les PRP salivaires humaines, IB5, II-1 et II-1NG ont été caractérisées par MS. Les données collectées, chaînes polypeptidiques présentes, nombre et position des glycosylations portées par la protéine II-1, sont essentielles aux études que nous voulions développer sur leurs interactions avec les tanins. Ce travail a donc constitué la base de l'étude par MS des interactions de la protéine IB5 avec différents tanins modèles. Il a également apporté des informations nécessaires à la bonne interprétation des données acquises par les approches mises en œuvre par les autres partenaires du projet PROTANIN. C'est dans ce contexte que les travaux de caractérisation concernant les protéines, II-1, II-1NG et IB5 ont fait l'objet d'une publication commune avec l'équipe en charge de l'étude des PRP par SAXS (Boze *et al.*, 2010).

L'ESI-MS a démontré l'existence de complexes solubles résultant de l'interaction entre IB5 et l'EgCG, un monomère de tanin. Elle s'est révélée capable d'analyser les solutions d'interactions en détectant ces complexes IB5•tanin dont la composition s'est avérée variable. En effet, ils sont présents en solution sous la forme de plusieurs stœchiométries. Au sein des complexes observés, le nombre de tanins présents augmente avec leur concentration. Nous avons également montré par MS/MS CID que les tanins se

fixeraient sur des sites distincts de la protéine. Ces premiers résultats ont fait l'objet d'une publication dans la revue internationale à comité de lecture « Analytical and Bioanalytical Chemistry » (Canon *et al.*, 2009).

Ceci nous a amené à aborder la question de l'impact de la structure des tanins sur leur interaction avec IB5. Pour cela, nous avons utilisé des expériences de compétition entre tanins pour IB5 qui ont été analysées par ESI-MS et poursuivi l'approche par MS/MS CID qui évalue la stabilité des différents complexes. L'effet de la structure des tanins sur la stabilité des complexes solubles IB5•tanin est similaire à celui démontré pour son impact sur leur astringence. En effet, la galloylation et le degré de polymérisation des tanins, deux caractéristiques structurales importantes pour leur caractère astringent (Vidal *et al.*, 2003), se sont révélés favorables à la stabilité des complexes formés (Canon *et al.*, 2010a). Une corrélation a été obtenue entre le nombre d'hydroxyles portés par les tanins et la stabilité de leurs complexes avec IB5 en phase gazeuse. Ce résultat démontre l'implication de liaisons hydrogène dans ces interactions en accord avec les résultats d'autres équipes. Une deuxième publication a été acceptée dans la revue internationale à comité de lecture « Analytical and Bioanalytical Chemistry » (Canon *et al.*, 2010a). Les résultats des expériences de compétition et de MS/MS CID ont montré une différence entre l'affinité des tanins pour IB5 et la stabilité de leurs complexes avec cette protéine. Alors que les complexes impliquant les dimères apparaissent plus stables que ceux impliquant les monomères, l'affinité des dimères pour IB5 serait plus faible. Les expériences de collision permettent d'apprécier la stabilité des complexes et donnent des informations essentiellement cinétiques sur les constantes de dissociation (k_d) et non sur l'affinité. Si l'affinité des dimères en solution est faible, ceci pourrait indiquer que a) ce qui se passe en phase gazeuse n'est pas représentatif de ce qui se passe en solution, ou que b) la constante d'association (k_a) est faible.

Les différences existant entre la phase gazeuse et la phase aqueuse proviennent du fait qu'en phase gazeuse une partie des contributions d'origine entropique ne sont pas conservées alors que les contributions enthalpiques apportées par les liaisons électrostatiques (liaisons hydrogène incluses) sont renforcées par rapport au milieu aqueux (Chen *et al.*, 1998). L'augmentation des contributions enthalpiques est plus importante pour les complexes impliquant les dimères que pour ceux impliquant les monomères, à cause du plus grand nombre de liaisons hydrogène que les dimères peuvent établir avec la protéine. Cette plus forte contribution énergétique des forces de liaisons de nature enthalpique en phase gazeuse pourrait compenser une pénalité entropique plus importante dans le cas des dimères par rapport à celui des monomères. La contribution entropique qui est probablement défavorable pour les dimères conduit à un k_a faible, ce qui conduirait à un k_D élevé (donc à

une faible affinité). Cette pénalité entropique pourrait avoir pour origine une restriction conformationnelle plus importante de la protéine lors de son interaction avec les dimères par rapport à son interaction avec les monomères.

Cette hypothèse peut être évaluée par l'étude des aspects structuraux de l'interaction. Après avoir démontré qu'IB5 a une affinité pour les tanins supérieure à celle d'une protéine témoin structurée, BPTI, nous avons recherché les sites d'interactions sur IB5 et étudié les éventuels changements structuraux liés à sa complexation. Les résultats obtenus par MS/MS CID sur les complexes IB5•EgCG de stœchiométries 1•1 à 1•9 semblent indiquer la présence de 8 sites d'interaction privilégiés des tanins sur IB5. Ces huit sites pourraient correspondre aux huit clusters de proline présents au sein des séquences répétées d'IB5 en accord avec la littérature (Murray *et al.*, 1994 ; Charlton *et al.*, 1996 ; Baxter *et al.*, 1997). Cependant, les résidus prolines ne sont probablement pas les seuls acides aminés impliqués dans cette interaction, les résidus glycine et arginine pourraient également intervenir (Luck *et al.*, 1994 ; Murray *et al.*, 1994 ; Hatano *et al.*, 1999 ; Charlton *et al.*, 2002a). Nous avons essayé de détecter ces sites d'interaction privilégiés par MS/MS ECD. Quoiqu'en accord avec la présence de plusieurs zones d'interaction sur la protéine, les résultats obtenus n'ont pas permis d'identifier de zones précises d'interaction. Les modifications conformationnelles d'IB5 induites par la fixation des tanins constituent un point encore soumis à questionnement. Le couplage de la spectrométrie de masse avec la mobilité ionique a permis de mettre en évidence la formation de complexes IB5•tanins, pour lesquels IB5 acquiert des conformations plus compactes à partir de la liaison d'une dizaine de tanins. Une diminution de la largeur du pic de mobilité d'IB5 est à noter à partir de la fixation de deux tanins sur sa chaîne peptidique. Cette observation indique une diminution du nombre de conformations explorées par la protéine et donc une pénalité de nature entropique due à la fixation des tanins.

Selon le modèle d'interaction de certaines IUP décrit par Fuxreiter *et al.* (Fuxreiter *et al.*, 2004) (cf introduction générale, partie 4.4.), il est raisonnable de penser que l'interaction s'établit au niveau des clusters proline qui constituent des structures rigides et accessibles (Charlton *et al.*, 1996). Autour de ces clusters polyproline rigides, la présence de glycine assurerait des zones de liaisons flexibles (Charlton *et al.*, 1996 ; Chatterjee *et al.*, 2005) qui permettraient au reste de la chaîne d'explorer l'espace conformationnel. Ceci autoriserait les clusters libres à établir de nouvelles liaisons avec d'autres groupements fonctionnels des tanins (Charlton *et al.*, 2002a). De plus, l'établissement d'un premier point de liaison réduit l'espace conformationnel exploré par le reste de la chaîne peptidique et rend possible une fixation plus rapide des suivants. Cela pourrait être particulièrement utile pour fixer les tanins de haut degré de polymérisation. Des phénomènes similaires impliquant une transition

structurale de type « désordre-ordre » et qui conduiraient à une augmentation de l'affinité de la protéine pour son partenaire au cours de l'interaction ont été observés chez de nombreuses PID (Dyson et Wright, 2002 ; Uversky, 2002b ; Fuxreiter *et al.*, 2004). Ces processus d'interaction conduisent à des états plus ordonnés, sans qu'il y ait obligatoirement acquisition de structures secondaires bien définies (Permyakov *et al.*, 2003 ; Bourhis *et al.*, 2005). La sélection d'un conformère de la protéine parmi les différentes conformations qui lui sont accessibles aboutit à une réduction de son entropie conformationnelle (Receveur-Brechot *et al.*, 2006). Ce phénomène nous apparaît supporté par les résultats de IMS-MS acquis dans le cas de l'interaction entre IB5 et l'EgCG.

Les résultats obtenus ont permis de démontrer que la SM est aujourd'hui un outil capable d'apporter des informations pertinentes dans l'élucidation des problématiques posées à la biologie structurale. Elle procure des données complémentaires à celles délivrées par les méthodes d'analyses classiquement utilisées pour l'étude des complexes supramoléculaires impliquant les protéines. On retiendra plus particulièrement que la MS est capable de fournir individuellement pour chacun des objets présents en solution

- leur masse (mesure du rapport m/z),
- leur composition et leur stœchiométrie,
- la séquence de la protéine partenaire,
- la localisation et le nombre des modifications post-traductionnelles portées par la protéine partenaire,
- les sites d'interactions au niveau de la séquence peptidique de la protéine partenaire (utilisation des techniques de fragmentation non ergotique),
- l'affinité des partenaires (expérience de compétition),
- leur stabilité (MS/MS CID),
- leur structure grâce au couplage avec la spectrométrie de mobilité ionique (mesure de la section efficace de collision).

Les résultats obtenus sont en accord avec ceux acquis par les autres techniques d'études (figure 93) (Boze *et al.*, 2010). Les sections efficaces de collision d'IB5 et de II-1NG indiquent que ces protéines ont une conformation étendue.

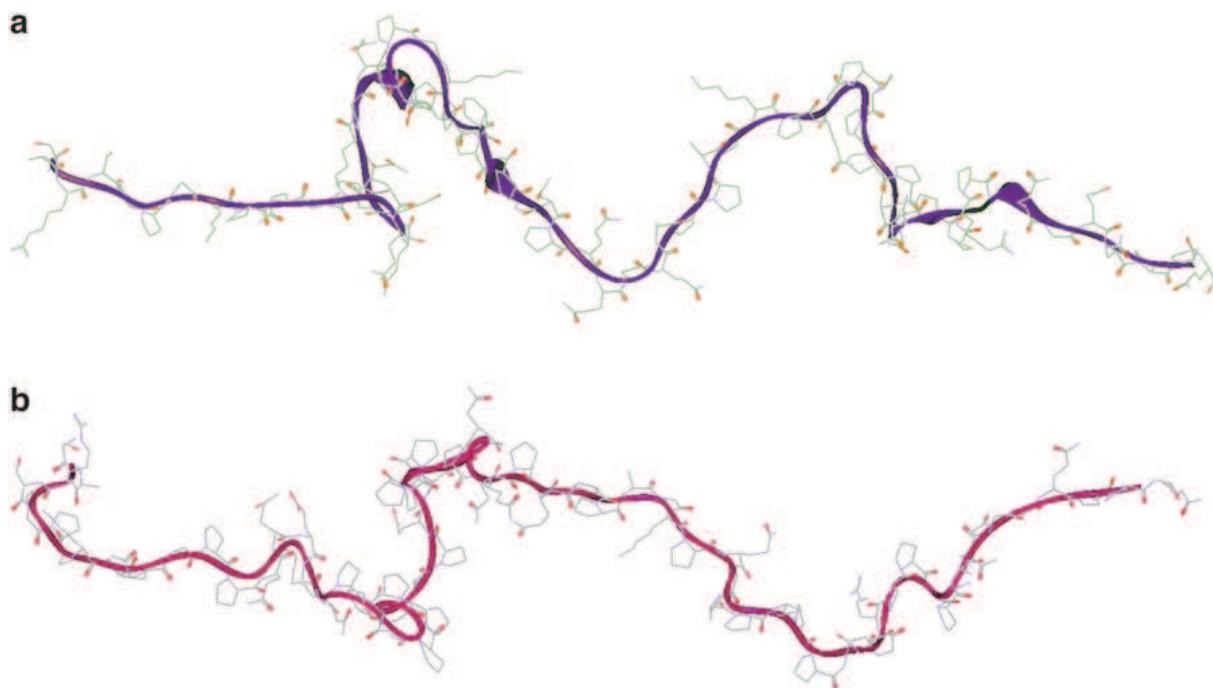


Figure 93. Deux exemples de conformation de la protéine IB5 compatibles avec les données de SAXS ; d'après (Boze *et al.*, 2010).

Les résultats d'IMS-MS indiquent l'acquisition d'un plus haut degré de structuration d'IB5 lors de son interaction avec les tanins en accord avec ceux obtenus par CD.

Les données de RMN ont permis l'attribution des signaux des atomes de carbones et d'hydrogène de la protéine IB5. Ces études se poursuivent actuellement sur les interactions IB5-EgCG. Ils devraient permettre de corréler les résultats de futures expériences de MS dont le but serait de déterminer les sites d'interactions préférentiels des tanins sur IB5.

En effet, des expériences supplémentaires de MS/MS visant à fragmenter la chaîne peptidique sans briser les liaisons non-covalentes pourraient fournir des informations sur la localisation des sites d'interactions. Ces expériences pourraient cibler des complexes avec des stœchiométries plus élevées que celle étudiée dans le chapitre IV. Ces complexes de hautes stœchiométries devraient présenter une saturation des sites d'interaction préférentiels des tanins sur la chaîne peptidique d'IB5. Ces expériences pourraient être à nouveau réalisées avec la fragmentation de type ECD sur un FT-ICR. Toutefois, l'approche qui à ce jour nous apparaît la plus prospective est relative à une nouvelle méthode d'activation en phase gazeuse apte à réaliser ce type de fragmentation. Cette technologie fait partie de la famille des fragmentations induites par photons en faisant intervenir des rayonnements dont les longueurs sont comprises dans l'ultraviolet. Les photons utilisés ont des énergies comprises entre 10 et 20eV, et permettent la formation d'espèces radicalaires par photoionisation d'un ion précurseur. Les processus de fragmentation mis en jeu

semblent de nature non ergodique, et pourraient se révéler capables de cliver la chaîne peptidique tout en conservant les liaisons faibles, telles que les liaisons non-covalentes.

La résolution au niveau de la mesure de la section efficace des molécules étudiées par IMS-MS est encore à améliorer. Ainsi, différentes équipes, dont la plus active est celle de Clemmer, développent à l'heure actuelle des solutions innovantes pour palier ce problème (Koeniger *et al.*, 2006 ; Kurulugama *et al.*, 2008 ; Merenbloom *et al.*, 2008 ; Merenbloom *et al.*, 2009). De plus, le couplage de l'IMS avec la MS/MS s'avère d'ores et déjà prometteur (Hyung *et al.*, 2009 ; Leary *et al.*, 2009). Un spectromètre de masse présentant une cellule de collision en amont de la cellule de mobilité et une en aval est actuellement proposé sur le marché. Les expériences de CID/IMS, IMS/CID ou de CID/IMS/CID réalisées sur ce type d'appareil permettent d'obtenir des informations à la fois sur la fragmentation et la structure des ions parents et de leurs ions produits. En effet, la prochaine étape de l'étude de ces interactions pourrait concerner l'étude des interactions entre PRP et tanins de haut degré de polymérisation. Cependant, les techniques de purification actuelles peinent encore à isoler précisément ces derniers. La capacité de l'IMS-MS à étudier individuellement chacune des espèces présentes en solution pourrait s'avérer être un avantage crucial dans l'étude de ces partenaires seuls et en interaction.

La MS s'avère une technique prometteuse dans l'identification et la caractérisation des objets supramoléculaires impliquant les protéines au moment où la protéomique et la biologie structurale se rejoignent (Zhou et Robinson, 2010) mais ne peut se concevoir comme la seule démarche dans ce contexte particulier des PID et des tanins.

En effet, la question relative à la relation qui peut exister entre ces interactions et l'astringence est importante. L'astringence est définie comme étant l'ensemble des sensations comprenant le dessèchement de la bouche, la rugosité des tissus buccaux ainsi qu'une impression d'étirement des joues et des muscles du visage (ASTM, 1989). Ces sensations pourraient résulter de différents phénomènes :

- le « dysfonctionnement » des protéines chargées de la lubrification de la cavité buccale, telles que les PRP salivaires glycosylées,
- l'interaction des tanins avec les protéines exposées à la surface de l'épithélium buccal,
- la précipitation de complexes protéine•tanin sur l'épithélium buccal.

Les PRP salivaires interviendraient à différents niveaux en fonction de leur nature. Les PRP glycosylées, telles que II-1, qui ont un rôle de lubrification de la cavité buccale pourraient être gênées dans leur activité par la fixation de tanin. Cette diminution de leur capacité de lubrification pourrait être la cause de l'augmentation de la friction ressentie au

niveau de la bouche. On peut émettre l'hypothèse que les *bPRP*, telles que IB5, interviennent comme une « première ligne de défense » contre les tanins. Leur plus forte affinité pour les tanins par rapport aux autres protéines démontrée par nos expériences avec la protéine BPTI, leur permettrait de protéger ces dernières en fixant un maximum de tanins, en accord avec les hautes stœchiométries observées. Les aliments peu concentrés en tanins ne permettraient pas de saturer le pouvoir de piégeage des *bPRP*, ainsi ils ne seraient pas perçus comme astringent. A l'inverse, la consommation d'aliments très riches en tanins aboutirait à une saturation de la capacité de fixation des *bPRP* et finalement à leur précipitation. Les tanins surnuméraires interagiraient alors avec les autres protéines dont les protéines présentent à la surface de l'épithélium buccal et celles ayant pour fonction la lubrification de la cavité buccale. La diminution des capacités de lubrification de ces dernières accompagnée de la précipitation des autres protéines salivaires serait alors à l'origine de l'ensemble des sensations prenant part à l'astringence. En faveur de cette hypothèse, il a été récemment montré que la diminution de la quantité de protéines salivaires au niveau de la cavité buccale augmentait la sensation d'astringence (Nayak et Carpenter, 2008). Dans cette étude, la distinction entre deux catégories de protéines salivaires, les protéines salivaires libres et les protéines salivaires formant la pellicule. Les résultats de cette étude suggèrent que les protéines libres protègent les protéines de la pellicule en formant des complexes solubles avec les tanins. Lorsque les protéines salivaires libres ne sont plus présentes, l'interaction des tanins avec les protéines de la pellicule entraîne la formation d'un précipité et la sensation d'astringence est plus fortement ressentie. Cette sensation induirait alors un comportement d'évitement vis-à-vis d'aliments trop riches en tanins.

Afin de confirmer cette hypothèse, une stratégie d'étude pourrait être l'élaboration d'un milieu modèle composé d'un épithélium buccal, de PRP glycosylées, d'un ensemble de macromolécules représentant la pellicule, de *bPRP* représentant les protéines libres et d'une solution reflétant les caractéristiques physico-chimiques de la salive. Au niveau de ce milieu des solutions de différentes concentrations en tanins pourraient être ajoutées. Une approche ESI-IMS-MS permettrait l'analyse et la caractérisation des complexes formés en solution, tandis qu'une approche d'imagerie par MALDI-MS pourrait permettre d'analyser les molécules et macromolécules adsorbées au niveau de l'épithélium. Cette approche pourrait être combinée à la microscope à force atomique (AFM, Atomic Force Microscopy) qui permettrait de topographier la surface de l'épithélium et de mettre en évidence la précipitation de gros agrégats. Le suivi de l'évolution de la composition des deux milieux (salive et surface de l'épithélium) en fonction de la concentration en tanins permettrait de vérifier l'hypothèse concernant le rôle protecteur des *bPRP* pour les autres protéines vis-à-

vis des tanins. L'évolution de la viscosité à la surface de l'épithélium en fonction de la concentration en tanins est un autre paramètre qu'il pourrait être important de suivre aux cours de ces expériences.

Références bibliographiques

- Aebersold R.** (1993) Mass spectrometry of proteins and peptides in biotechnology. *Current Opinion in Biotechnology*, 4, 412-419.
- Aebersold R. et Mann M.** (2003) Mass spectrometry-based proteomics. *Nature*, 422, 198-207.
- Amir D. et Haas E.** (1988) Reduced bovine pancreatic trypsin inhibitor has a compact structure. *Biochemistry*, 27, 8889-8893.
- Anfinsen C.B.** (1973) Principles that govern the folding of protein chains. *Science*, 181, 223-230.
- Ann D.K. et Carlson D.M.** (1985) The structure and organization of a proline-rich protein gene of a mouse multigene family. *Journal of Biological Chemistry*, 260, 15863-15872.
- Ann D.K., Gadbois D. et Carlson D.M.** (1987) Structure, organization, and regulation of a hamster proline-rich protein gene. *The Journal of Biological Chemistry*, 262, 3958-3963.
- Ann D.K., Lin H.H. et Kousvelary E.** (1997) Regulation of salivary-gland-specific gene expression. *Critical Reviews in Oral Biology and Medicine*, 8, 244-252.
- Argos P.** (1990) An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *Journal of Molecular Biology*, 211, 943-958.
- Arnold R.A., Noble A.C. et Singleton V.L.** (1980) Bitterness and astringency of phenolic fractions in wine. *Journal of Agricultural and Food Chemistry*, 28, 675-678.
- Artz W.E., Bishop P.D., Dunker A.K., Schanus E.G. et Swanson B.G.** (1987) Interaction of synthetic proanthocyanidin dimer and trimer with bovine serum albumin and purified bean globulin fraction G-1. *Journal of Agricultural and Food Chemistry*, 35, 417-421.
- Asquith T.N. et Butler L.** (1985) Use of dye-labeled protein as spectrophotometric assay for protein precipitants such as tannin. *Journal of Chemical Ecology*, 11, 1535-1544.
- Asquith T.N., Uhlig J., Mehansho H., Putnam L., Carlson D.M. et Butler L.** (1987) Binding of condensed tannins to salivary proline-rich glycoproteins: the role of carbohydrates. *Journal of Agricultural and Food Chemistry*, 35, 331-334.
- ASTM** (1989) Standard definitions of terms relating to sensory evaluation of materials and products., Annual Book of ASTM Standards, vol. 15.070. American Society of Testing and Materials, Philadelphia. pp. 2.
- Austin P.J., Suchar L.A., Robbins C.T. et Hagerman A.E.** (1989) Tannin-binding proteins in saliva of deer and their absence in saliva of sheep and cattle. *Journal of Chemical Ecology*, 15, 1335-1347.
- Azen E., Prakobphol A. et Fisher S.J.** (1993) PRB3 null mutations result in absence of the proline-rich glycoprotein GI and abolish *Fusobacterium nucleatum* interactions with saliva *in vitro*. *Infection and Immunity*, 61, 4434-4439.

- Azen E.A., Amberger E., Fisher S., Prakobphol A. et Niece R.L.** (1996) PRB1, PRB2, and PRB4 coded polymorphisms among human salivary concanavalin-A binding, II-1, and Po proline-rich proteins. *American Journal of Human Genetics*, 58, 143-153.
- Bacon J.R. et Rhodes M.J.C.** (1998) Development of a competition assay for the evaluation of the binding of human parotid salivary proteins to dietary complex phenols and tannins using a peroxidase-labeled tannin. *Journal of Agricultural and Food Chemistry*, 46, 5083-5088.
- Bakhtiar R. et Guan Z.** (2006) Electron capture dissociation mass spectrometry in characterization of peptides and proteins. *Biotechnology Letters*, 28, 1047-1059.
- Bate-Smith E. et Swain T.** (1962). In: H. Mason and A. Florkin (eds.), *Comparative biochemistry*, vol. 3. Academic Press, New York. pp. 764.
- Bate-Smith E.C.** (1973) Haemanalysis of tannins : the concept of relative astringency. *Phytochemistry*, 12, 907-912.
- Baxter N.J., Lilley T.H., Haslam E. et Williamson M.P.** (1997) Multiple interactions between polyphenols and a salivary proline-rich protein repeat results in complexation and precipitation. *Biochemistry*, 36, 5566-5577.
- Beeley J.A., Khoo K.S. et Lamey P.J.** (1991a) Two-dimensional electrophoresis of human parotid salivary proteins from normal and connective tissue disorder subjects using immobilised pH gradients. *Electrophoresis*, 12, 493-499.
- Beeley J.A., Sweeney D., Lindsay J.C.B., Buchanan M.L., Sarna L. et Khoo K.S.** (1991b) Sodium dodecyl sulphate-polyacrylamide gel electrophoresis of human parotid salivary proteins. *Electrophoresis*, 12, 1032-1041.
- Beeley J.A., Newman F., Wilson P.H. et Shimmin I.C.** (1996) Sodium dodecyl sulphate-polyacrylamide gel electrophoresis of human parotid salivary proteins: Comparison of dansylation, Coomassie Blue R-250 and silver detection methods. *Electrophoresis*, 17, 505-506.
- Beeley J.A.** (2001) Basic proline-rich proteins: multifunctional defence molecules? *Oral Diseases*, 7, 69-70.
- Benesch J.L.P., Ruotolo B.T., Simmons D.A. et Robinson C.V.** (2007) Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chemical Reviews*, 107, 3544-3567.
- Bennick A., Cannon M. et Madapallimattam G.** (1979) The nature of the hydroxyapatite-binding site in salivary acidic proline-rich proteins. *Biochemical Journal*, 183, 115-126.
- Bennick A., McLaughlin A.C., Grey A.A. et Madapallimattam G.** (1981) The location and nature of calcium-binding sites in salivary acidic proline-rich phosphoproteins. *Journal of Biological Chemistry*, 256, 4741-4746.
- Bennick A.** (1982) Salivary proline-rich proteins. *Molecular and Cellular Biochemistry*, 45, 83-99.
- Bennick A., Kells D. et Madapallimattam G.** (1983) Interaction of calcium ions and salivary acidic proline-rich proteins with hydroxyapatite. *Biochemical Journal*, 213, 11-20.

- Bennick A.** (2002) Interaction of plant polyphenol with salivary proteins. *Critical Reviews in Oral Biology & Medicine*, 13, 184-196.
- Berg J.M., Tymoczko J.L. et Stryer L.** (2002) *Biochemistry*, fifth ed. W. H. Freeman and Company, New York, pp. 1514.
- Biemann K.** (1988) Contributions of mass spectrometry to peptide and protein structure. *Biological Mass Spectrometry*, 16, 99-111.
- Blytt, Guscar et Butler** (1988) Antinutritional effects and ecological significance of dietary condensed tannins may not be due to binding and inhibiting digestive enzymes. *Journal of Chemical Ecology*, 14, 1455-1463.
- Borch J., Jorgensen T.J.D. et Roepstorff P.** (2005) Mass spectrometric analysis of protein interactions. *Current Opinion in Chemical Biology*, 9, 509-516.
- Bourhis J.-M., Receveur-Bréchet V., Oglesbee M., Zhang X., Buccellato M., Darbon H., Canard B., Finet S. et Longhi S.** (2005) The intrinsically disordered C-terminal domain of the measles virus nucleoprotein interacts with the C-terminal domain of the phosphoprotein via two distinct sites and remains predominantly unfolded. *Protein Science*, 14, 1975-1992.
- Boze H., Marlin T., Durand D., Pérez J., Vernhet A., Canon F., Sarni-Manchado P., Cheynier V. et Cabane B.** (2010) Proline-rich salivary proteins have extended conformations. *Biophysical Journal*, 99, 656-665.
- Breslin P.A., Gilmore M., Beauchamp G.K. et Green B.G.** (1993) Psychophysical evidence that oral astringency is a tactile sensation. *Chemical Senses*, 18, 405-417.
- Bretthauer R.K. et Castellino F.J.** (1999) Glycosylation of *Pichia pastoris* - derived proteins. *Biotechnology and Applied Biochemistry*, 30, 193-200.
- Bruins A.P.** (1998) Mechanistic aspects of electrospray ionization. *Journal of Chromatography A*, 794, 345-357.
- Burritt E.A. et Provenza F.D.** (2000) Role of toxins in intake of varied diets by sheep. *Journal of Chemical Ecology*, 26, 1991-2005.
- Caillet M.-M.** (1994) Stabilisation et clarification des vins par le collage. *Revue des Oenologues*, 74, 15-18.
- Canon F., Paté F., Meudec E., Marlin T., Cheynier V., Giuliani A. et Sarni-Manchado P.** (2009) Characterization, stoichiometry and stability of salivary protein-tannin complexes by ESI-MS and ESI-MS/MS. *Analytical and Bioanalytical Chemistry*, 395, 2535-2545.
- Canon F., Giuliani A., Paté F. et Sarni-Manchado P.** (2010a) Ability of a salivary intrinsically unstructured protein to bind different tannin targets revealed by mass spectrometry. *Analytical and Bioanalytical Chemistry*, 398, 815-822.
- Chalmers M.J., Håkansson K., Johnson R., Smith R., Shen J., Emmett M.R. et Marshall A.G.** (2004) Protein kinase A phosphorylation characterized by tandem Fourier transform ion cyclotron resonance mass spectrometry. *PROTEOMICS*, 4, 970-981.
- Chan M. et Bennick A.** (2001) Proteolytic processing of a human salivary proline-rich protein precursor. *European Journal of Biochemistry*, 268, 3423-3431.

Chandler D. (2005) Interfaces and the driving force of hydrophobic assembly. *Nature*, 437, 640-647.

Charlton A.J., Baxter N.J., Lilley T.H., Haslam E., McDonald C.J. et Williamson M.P. (1996) Tannin interactions with a full-length human salivary proline -rich protein display a stronger affinity than with single proline-rich repeats. *FEBS Letters*, 382, 289-292.

Charlton A.J., Baxter N.J., Lokman Khan M., Moir A.J.G., Haslam E., Davies A.P. et Williamson M.P. (2002a) Polyphenol/peptide binding and precipitation. *Journal of Agricultural and Food Chemistry*, 50, 1593-1601.

Charlton A.J., Haslam E. et Williamson M.P. (2002b) Multiple conformations of the proline-rich-protein/epigallocatechin gallate complex determined by time-averaged nuclear overhauser effects. *Journal of the American Chemical Society*, 124, 9899-9905.

Chatterjee A., Kumar A., Chugh J., Srivastava S., Bhavesh N.S. et Hosur R.V. (2005) NMR of unfolded proteins. *Journal of Chemical Sciences*, 117, 3-21.

Chazin W. et Veenstra T.D. (1999) Determination of the metal-binding cooperativity of wild-type and mutant calbindin D9K by electrospray ionization mass spectrometry. *Rapid Communications in Mass Spectrometry*, 13, 548-555.

Cheek P.J. et Lilley T.H. (1988) The enthalpies of interaction of some amides with urea in water at 25 °C. *Journal of the Chemical Society-Faraday Transactions I*, 84, 1927-1940.

Chen J., McAllister M.A., Lee J.K. et Houk K.N. (1998) Short, strong hydrogen bonds in the gas phase and in solution: theoretical exploration of pKa matching and environmental effects on the strengths of hydrogen bonds and their potential roles in enzymatic catalysis. *The Journal of Organic Chemistry*, 63, 4611-4619.

Cheyrier V., Prieur C., Guyot S., Rigaud J. et Moutounet M. (1997) The structures of tannins in grapes and wines and their interactions with proteins. In: T. R. Watkins (ed), *Wine. Nutritional and therapeutic benefits*, vol. 661. American Chemical Society, Washington, DC. pp. 81-93.

Cheyrier V. et Sarni-Manchado P. (2006) Structures phénoliques et goût. In: P. Sarni-Manchado and V. Cheyrier (eds.), *Les polyphénols en agroalimentaire*. Tec & Doc, Lavoisier, Londres, Paris, New York. pp. 89-134.

Chung-MacCoubrey A.L., Hagerman A.E. et Kirkpatrick R.L. (1997) Effects of tannins on digestion and detoxification activity in gray squirrels (*Sciurus carolinensis*). *Physiological Zoology*, 70, 270-277.

Clausen T.P., Provenza F.D., Burritt E.A., Reichardt P.B. et Bryant J.P. (1990) Ecological implication of condensed tannin structure: a case study. *Journal of Chemical Ecology*, 16, 2381-2392.

Clauss M., Gehrke J., Hatt J.M., Dierenfeld E.S., Flach E.J., Hermes R., Castell J., Streich W.J. et Fickel J. (2005) Tannin-binding salivary proteins in three captive rhinoceros species. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 140, 67-72.

Clemmer D.E. et Jarrold M.F. (1997) Ion mobility measurements and their applications to clusters and biomolecules. *Journal of Mass Spectrometry*, 32, 577-592.

- Clifford M.N.** (1997) Astringency. In: F. Tomas-Barberan and R. Robins (eds.), *Phytochemistry of fruits and vegetables*. Clarendon Press, Oxford. pp. 87-108.
- Cole R.B.** (2003) Focus on noncovalent interactions. *Journal of the American Society for Mass spectrometry*, 14, 417-418.
- Colley K.J.** (1997) Golgi localization of glycosyltransferases: more questions than answers. *Glycobiology*, 7, 1-13.
- Cooper H.J., Håkansson K. et Marshall A.G.** (2005) The role of electron capture dissociation in biomolecular analysis. *Mass Spectrometry Reviews*, 24, 201-222.
- Cooper S.M. et Owen-Smith N.** (1985) Condensed tannins deter feeding by browsing ruminants in a South African savanna. *Oecologia*, 67, 142-146.
- Creese A.J. et Cooper H.J.** (2007) Liquid Chromatography Electron Capture Dissociation Tandem Mass Spectrometry (LC-ECD-MS/MS) versus Liquid Chromatography Collision-induced Dissociation Tandem Mass Spectrometry (LC-CID-MS/MS) for the Identification of Proteins. *Journal of the American Society for Mass spectrometry*, 18, 891-897.
- Creese A.J. et Cooper H.J.** (2008) The effect of phosphorylation on the electron capture dissociation of peptide ions. *Journal of the American Society for Mass spectrometry*, 19, 1263-1274.
- Cunningham L., Clouse R. et Ford J.** (1963) Heterogeneity of the carbohydrate moiety of crystalline ovalbumin. *Biochimica et Biophysica Acta*, 78, 379.
- Dadlez M. et Kim P.S.** (1995) A third native one-disulphide intermediate in the folding of bovine pancreatic trypsin inhibitor. *Nature structural biology*, 2, 674-679.
- Daniel J.M., Friess S.D., Rajagopalan S., Wendt S. et Zenobi R.** (2002) Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *International Journal of Mass Spectrometry*, 216, 1-27.
- Daniel J.M., McCombie G., Wendt S. et Zenobi R.** (2003) Mass spectrometric determination of association constants of adenylate kinase with two noncovalent inhibitors. *Journal of American Chemical Society*, 14.
- Dass C.** (2007) *Fundamentals of Contemporary Mass Spectrometry*. John Wiley & Sons, Hoboken, New Jersey, pp. 585.
- Daughdrill G.W., Chadsey M.S., Karlinsey J.E., Hughes K.T. et Dahlquist F.W.** (1997) The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, σ 28. *Nature Structural Biology*, 4, 285-291.
- de Freitas V. et Mateus N.** (2001) Structural features of procyanidin interactions with salivary proteins. *Journal of Agricultural and Food Chemistry*, 49, 940-945.
- de Hoffmann E. et Stroobant V.** (2005) *Spectrométrie de masse. Cours et exercices corrigés.*, Dunod ed, vol. Third edition, Paris, pp. 425.
- Dearing M.D. et Cork S.** (1999) Role of detoxification of plant secondary compounds on diet breadth in a mammalian herbivore, *Trichosurus vulpecula*. *Journal of Chemical Ecology*, 25, 1205-1219.

- Dixon R., Xie D. et Sharma S.** (2005) Proanthocyanidins - a final frontier in flavonoid research? *New Phytologist*, 165, 9-28.
- Dodds M.W.J., Johnson D.A. et Yeh C.-K.** (2005) Health benefits of saliva: a review. *Journal of Dentistry*, 33, 223-233.
- Dole M., Mack L.L., Hines R.L., Mobley R.C., Ferguson L.D. et Alice M.B.** (1968) Molecular beams of macroions. *Journal of Chemical Physics*, 49, 2240.
- Dreisewerd K., Schürenberg M., Karas M. et Hillenkamp F.** (1995) Influence of the laser intensity and spot size on the desorption of molecules and ions in matrix-assisted laser desorption/ionization with a uniform beam profile. *International Journal of Mass Spectrometry and Ion Processes*, 141, 127-148.
- Drickamer K. et Taylor M.E.** (1998) Evolving views of protein glycosylation. *Trends in Biochemical Sciences*, 23, 321-324.
- Dugourd P., Hudgins R.R., Clemmer D.E. et Jarrold M.F.** (1997) High-resolution ion mobility measurements. *Reviews of Scientific Instrument*, 68, 1122-1129.
- Dunker A.K., Obradovic Z., Romero P., Garner E.C. et Brown C.J.** (2000a) Intrinsic protein disorder in complete genomes. *Genome Informatics Series. Workshop on Genome Informatics*, 11, 161-171.
- Dunker A.K., Obradovic Z., Romero P., Garner E.C. et Brown C.J.** (2000b) Intrinsic protein disorder in complete genomes. *Genome Informatics*, 11, 161-171.
- Dunker A.K., Lawson J.D., Brown C.J., Williams R.M., Romero P., Oh J.S., Oldfield C.J., Campen A.M., Ratliff C.M. et Hipps K.W.** (2001) Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling*, 19, 26-59.
- Dunker A.K. et Obradovic Z.** (2001) The protein trinity linking function and disorder. *Nature Biotechnology*, 19, 805-806.
- Dunker A.K., Brown C.J., Lawson J.D., Iakoucheva L.M. et Obradovic Z.** (2002) Intrinsic disorder and protein function. *Biochemistry*, 41, 6573-6582.
- Dunker A.K., Silman I., Uversky V.N. et Sussman J.L.** (2008) Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology*, 18, 756-764.
- Dyson H.J. et Wright P.E.** (2002) Coupling of folding and binding for unstructured proteins. *Current Opinion in Structural Biology*, 12, 54-60.
- Dyson H.J. et Wright P.E.** (2005) Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology*, 6, 197-208.
- Edgar W.M.** (1992) Saliva: its secretion, composition and functions. *British Dental Journal*, 172, 305-312.
- Ezaki-Furuichi E., Nonaka G.I., Nishioka I. et Hayashi K.** (1987) Affinity of procyanidins (condensed tannins) from the bark of *Rhaphiolepis umbellata* for proteins. *Agricultural and Biological Chemistry*, 51, 115-120.
- Fagarasan S. et Honjo T.** (2003) Intestinal IgA synthesis: regulation of front-line body defences. *Nature Reviews Immunology*, 3, 63-72.

- Fenn J.B., Mann M., Meng C.K., Wong S.F. et Whitehouse C.M.** (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246, 64-71.
- Ferguson D.B.** (1999) The salivary glands and their secretions. In: D. B. Ferguson (ed), *Oral Bioscience*. Churchill Livingstone, Edinburgh, London, New York, Philadelphia, St Louis. pp. 117-150.
- Fernandez J. et Lilley T.H.** (1992) Aqueous solutions containing amino acids and peptides. Part 28.—Enthalpy of interaction of some amides with glycine and -alanine : interactions of the zwitterionic group of -amino acids with hydrophobic groups and peptide groups. *Journal of the Chemical Society, Faraday Transactions*, 88, 2503-2509.
- Fickel J., Göritz F., Joest B.A., Hildebrandt T., Hofmann R.R. et Breves G.** (1998) Analysis of parotid and mixed saliva in Roe deer (*Capreolus capreolus L.*). *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 168, 257-264.
- Fink A.L.** (2005) Natively unfolded proteins. *Current Opinion in Structural Biology*, 15, 35-41.
- Fischer E.** (1894) Einfluss der configuration auf die wirkung derenzyme. *Berichte der deutschen chemischen Gesellschaft*, 27, 2985–2993.
- Fischer U., Boulton R. et Noble A.** (1994) Physiological factors contributing to the variability of sensory assessments : relationship between salivary flow-rate and temporal perception of gustatory stimuli. *Food Quality and Preference*, 5, 55-64.
- Foley W.J. et Moore B.D.** (2005) Plant secondary metabolites and vertebrate herbivores - from physiological regulation to ecosystem function. *Current Opinion in Plant Biology*, 8, 430-435.
- Fowler M.E. et Richards W.P.C.** (1965) Acorn poisoning in a cow and a sheep. *Journal of the American Veterinary Medical Association*, 147, 1215-1220.
- Fuxreiter M., Simon I., Friedrich P. et Tompa P.** (2004) Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *Journal of Molecular Biology*, 338, 1015-1026.
- Gabelica V. et De Pauw E.** (2001) Comparison between solution-phase stability and gas-phase kinetic stability of oligodeoxynucleotide duplexes. *Journal of Mass Spectrometry*, 36, 397-402.
- Gabelica V., Galic N., Rosu F., Houssier C. et Pauw E.D.** (2003) Influence of response factors on determining equilibrium association constants of non-covalent complexes by electrospray ionization mass spectrometry. *Journal of Mass Spectrometry*, 38, 491-501.
- Gabelica V., Rosu F. et De Pauw E.** (2009) A simple method to determine electrospray response factors of noncovalent complexes. *Analytical Chemistry*, 81, 6708-6715.
- Ganem B., Li Y.T. et Henion J.D.** (1991) Detection of non-covalent receptor-ligand complexes by mass spectrometry. *Journal of the American Chemical Society*, 113, 6294-9296.
- Ganem B.** (1993) Detecting non-covalent interactions: new frontiers for mass spectrometry. *American Biotechnology Laboratory*, 11, 32-34.

- Ganem B. et Henion J.D.** (2003) Going gently into flight : analyzing noncovalent interactions by mass spectrometry. *Bioorganic and Medicinal Chemistry*, 11, 311-314.
- Gemmill T.R. et Trimble R.B.** (1999) Overview of N- and O-linked oligosaccharide structures found in various yeast species. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1426, 227-237.
- Gillece-Castro B.L., Prakobphol A., Burlingame A.L., Leffler H. et Fisher S.J.** (1991) Structure and bacterial receptor activity of a human salivary proline-rich glycoprotein. *Journal of Biological Chemistry*, 266, 17358-17368.
- Glückmann M., Pfenninger A., Krüger R., Thierolf M., Karasa M., Horneffer V., Hillenkamp F. et Strupat K.** (2001) Mechanisms in MALDI analysis: surface interaction or incorporation of analytes? *International Journal of Mass Spectrometry*, 210-211, 121-132.
- Gomez A. et Tang K.** (1994) Charge and fission of droplets in electrostatic sprays. *Physics of Fluid*, 6, 404-414.
- Gonzalez-de-Colmenares N., Ramirez-Martinez J.R., Aldana J.O., Ramos-Nino M.E., Clifford M.N., Pekerar S. et Mendez B.** (1998) Isolation, characterisation and determination of biological activity of coffee proanthocyanidins. *Journal of the Science of Food and Agriculture*, 77, 368-372.
- Green B.G.** (1993) Oral astringency : a tactile component of flavor. *Acta Psychologica*, 84, 119-125.
- Griffiths W.J., Jonsson A.P., Liu S., Rai D.K. et Wang Y.** (2001) Electrospray and tandem mass spectrometry in biochemistry. *Biochemical Journal*, 355, 545-61.
- Gross J.H.** (2004) Mass spectrometry: a textbook. Springer Verlag, Berlin Heidelberg New York, pp. 536.
- Gsponer J., Futschik M.E., Teichmann S.A. et Babu M.M.** (2008) Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science*, 322, 1365-1368.
- Guinard J.X., Pangborn R.M. et Lewis M.J.** (1986a) Preliminary studies on acidity-astringency interactions in model solutions and wines. *Journal of the Science of Food and Agriculture*, 37, 811-817.
- Guinard J.X., Pangborn R.M. et Lewis M.J.** (1986b) The time-course of astringency in wine upon repeated ingestion. *American Journal of Enology and Viticulture*, 37, 184-189.
- Guinard J.X. et Mazzucchelli R.** (1996) The sensory perception of texture and mouthfeel. *Trends in Food Science & Technology*, 7, 213-219.
- Gunasekaran K., Tsai C., Kumar S., Zanuy D. et Nussinov R.** (2003) Extended disordered proteins: targeting function with less scaffold. *Trends in Biochemical Sciences*, 28, 81 - 85.
- Gustavson K.H.** (1956) The chemistry of tanning processes. Academic Press, New York, pp. 403.

- Guyot S., Pellerin P., Brillouet J.-M. et Cheynier V.** (1996) Inhibition of β -glucosidase (*Amygdalae dulces*) by (+)-catechin oxidation products and procyanidin dimers. *Bioscience, Biotechnology and Biochemistry*, 60, 1131-1135.
- Hager J.W.** (2002) A new linear ion trap mass spectrometer. *Rapid Communications in Mass Spectrometry*, 16, 512-526.
- Hagerman A., E. et Butler L.G.** (1978) Protein precipitation method for the quantitative determination of tannins. *Journal of Agricultural and Food Chemistry*, 26, 809-812.
- Hagerman A.E. et Butler L.G.** (1980) Condensed tannin purification and characterization of tannin-associated proteins. *Journal of Agricultural and Food Chemistry*, 28, 947-952.
- Hagerman A.E. et Butler L.G.** (1981) The specificity of proanthocyanidin-protein interactions. *The Journal of Biological Chemistry*, 256, 4494-4497.
- Hagerman A.E. et Klucher K.M.** (1986) Tannin-protein interactions. In: V. M. Cody, E.; Harborne, J. (ed), *Plant flavonoids in biology and medicine: biochemical, pharmacological and structure-activity relationships*. Alan R. Liss, Inc., New York. pp. 67-76.
- Hagerman A.E.** (1989) Chemistry of tannin-protein complexation. In: R. W. Hemingway and J. J. Karchesy (eds.), *Chemistry and significance of condensed tannins*. Plenum Press, New York, London. pp. 323-331.
- Hagerman A.E. et Robbins C.T.** (1993) Specificity of tannin-binding salivary proteins relative to diet selection by mammals. *Canadian Journal of Zoology*, 71, 628-633.
- Hagerman A.E., Rice M.E. et Richard N.T.** (1998) Mechanisms of protein precipitation for two tannins, pentagalloylglucose and epicatechin 16 (4-8)catechin (procyanidin). *Journal of Agricultural and Food Chemistry*, 46, 2590-2595.
- Haghighat M., Moetamed A., Vaseghi T. et Aminlari M.** (1996) Isoprenaline induces biosynthesis of proline-rich proteins in the salivary glands of rat but not in sheep. *Comparative Biochemistry and Physiology - Part C: Toxicology & Pharmacology*, 115, 165-168.
- Haslam E.** (1974) Polyphenol-protein interactions. *Biochemical Journal*, 139, 285-288.
- Haslam E. et Lilley T.H.** (1988) Natural astringency in foodstuffs. A molecular interpretation. *Critical Reviews in Food Science and Nutrition*, 27, 1-40.
- Haslam E., Lilley T.H., Warminski E., Liao H., Cai Y., Martin R., Gaffney S.H., Goulding P.N. et Luck G.** (1992) Polyphenol complexation. A study in molecular recognition. In: C.-T. Ho, C. Y. Lee and M.-T. Huang (eds.), *Phenolic compounds in food and their effects on health*, vol. 506. American Chemical Society, New York. pp. 8-50.
- Haslam E.** (1996) Natural polyphenols (vegetable tannins) as drugs: possible modes of action. *Journal of Natural Products*, 59, 205-215.
- Hatano T. et Hemingway R.W.** (1996) Association of (+)-catechin and catechin-(4 α →8)-catechin with oligopeptides. *Journal of Chemical Society. Chemical Communications*, 2537-2538.
- Hatano T., Yoshida T. et Hemingway R.W.** (1999) Interaction of flavonoids with peptides and proteins and conformations of dimeric flavonoids in solution. In: G. G. Gross, R. W.

Hemingway, T. Yoshida and S. J. Branham (eds.), Plant polyphenols 2. Chemistry, biology, pharmacology, ecology. Kluwer Academic/Plenum Publisher, New York, Boston, Dordrecht, London, Moscow. pp. 509-526.

Hatton M.N., Loomis R.E., Levine M.J. et Tabak L.A. (1985) Masticatory lubrication. The role of carbohydrate in the lubricating property of a salivary glycoprotein-albumin complex. *Biochemical Journal*, 230, 817-820.

Hauer J.A., Taylor S.S. et Johnson D.A. (1999) Binding-dependent disorder-order transition in PKI α : a fluorescence anisotropy study. *Biochemistry*, 38, 6774-6780.

Hay D., Carlson E., Schluckebier S., Moreno E. et Schlesinger D. (1987) Inhibition of calcium phosphate precipitation by human salivary acidic proline-rich proteins: Structure-activity relationships. *Calcified Tissue International*, 40, 126-132.

Hebrard E., Bessin Y., Michon T., Longhi S., Uversky V., Delalande F., Van Dorselaer A., Romero P., Walter J., Declerck N. et Fargette D. (2009) Intrinsic disorder in viral proteins genome-linked : experimental and predictive analyses. *Virology Journal*, 6, 23.

Heck A.J.R. et van den Heuvel R.H.H. (2004) Investigation of intact protein complexes by mass spectrometry. *Mass Spectrometry Reviews*, 23, 368-389.

Helenius A. et Aebi M. (2001) Intracellular functions of N-linked glycans. *Science*, 291, 2364-2369.

Hemingway R.W., Foo L.Y. et Porter L.J. (1982) Linkage isomerism in trimeric and polymeric 2,3-cis-procyanidins. *Journal of Chemical Society Perkin Transactions I*, 1209-1216.

Hernandez H. et Robinson C.V. (2001) Dynamic protein complexes : insights from mass spectrometry. *The Journal of Biological Chemistry*, 276, 46685-46688.

Horne J., Hayes J. et Lawless H.T. (2002) Turbidity as a measure of salivary protein reactions with astringent substances. *Chemical Senses*, 27, 653-659.

Hu C.C. et Ghabrial S.A. (1995) The conserved, hydrophilic and arginine-rich N-terminal domain of cucumovirus coat proteins contributes to their anomalous electrophoretic mobilities in sodium dodecylsulfate-polyacrylamide gels. *Journal of Virological Methods*, 55, 367-379.

Hu P. et Loo J.A. (1995) Determining calcium-binding stoichiometry and cooperativity of parvalbumin and calmodulin by mass spectrometry. *Journal of Mass Spectrometry*, 30, 1076-1082.

Humphrey S.P. et Williamson R.T. (2001) A review of saliva: normal composition, flow, and function. *The Journal of Prosthetic Dentistry*, 85, 162-169.

Hunt D.F., Yates J.R., Shabanowitz J., Winston S. et Hauer C.R. (1986) Protein sequencing by tandem mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, 83, 6233-6237.

Hyung S.-J., Robinson C.V. et Ruotolo B.T. (2009) Gas-phase unfolding and disassembly reveals stability differences in ligand-bound multiprotein complexes. *Chemistry & Biology*, 16, 382-390.

- Iakoucheva L.M., Kimzey A.L., Masselon C.D., Smith R.D., Dunker A.K. et Ackerman E.J.** (2001) Aberrant mobility phenomena of the DNA repair protein XPA. *Protein Science*, 10, 1353-1362.
- Iribarne J.V. et Thomson B.A.** (1976) On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics*, 64, 2287.
- Iribarne J.V., Dziedzic P.J. et Thomson B.A.** (1983) Atmospheric pressure ion evaporation-mass spectrometry. *International Journal of Mass Spectrometry and Ion Physics*, 50, 331-347.
- Jackson S.N., Dutta S. et Woods A.S.** (2009) The use of ECD/ETD to identify the site of electrostatic interaction in noncovalent complexes. *Journal of the American Society for Mass Spectrometry*, 20, 176-179.
- Jarrold M.F.** (1998) Unfolding, refolding, and hydration of proteins in the gas phase. *Accounts of Chemical Research*, 32, 360-367.
- Jiang S.T., Chen G.H., Tang S.J. et Chen C.S.** (2002) Effect of glycosylation modification (N-Q-108I → N-Q-108T) on the freezing stability of recombinant chicken cystatin overexpressed in *Pichia pastoris* X-33. *Journal of Agricultural and Food Chemistry*, 50, 5313-5317.
- Jorgensen T.J.D., Roepstorff P. et Heck A.J.R.** (1998) Direct determination of solution binding constants for noncovalent complexes between bacterial cell wall peptide analogues and vancomycin group antibiotics by electrospray ionisation mass spectrometry. *Analytical Chemistry*, 70, 4427-4432.
- Jorgensen T.J.D., Hvelplund P., Andersen J.U. et Roepstorff P.** (2002) Tandem mass spectrometry of specific vs. nonspecific noncovalent complexes of vancomycin antibiotics and peptide ligands. *International Journal of Mass Spectrometry*, 219, 659-670.
- Jørgensen T.J.D., Delforge D., Remacle J., Bojesen G. et Roepstorff P.** (1999) Collision-induced dissociation of noncovalent complexes between vancomycin antibiotics and peptide ligand stereoisomers : evidence for molecular recognition in the gas phase. *International Journal of Mass Spectrometry*, 188, 63-85.
- Kallithraka S., Bakker J. et Clifford M.N.** (1997) Evaluation of bitterness and astringency of (+)-catechin and (-)-epicatechin in red wine and in model solutions. *Journal of Sensory Studies*, 12, 25-37.
- Karas M. et Hillenkamp F.** (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*, 60, 2299-2301.
- Kauffman D.L. et Keller P.J.** (1979) The basic proline-rich proteins in human parotid saliva from a single subject. *Archives of Oral Biology*, 24, 249-256.
- Kebarle P. et Peschke M.** (2000) On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions. *Analytica Chimica Acta*, 406, 11-35.
- Kempen E.C. et Brodbelt J.S.** (2000) A method for the determination of binding constants by electrospray ionization mass spectrometry. *Analytical Chemistry*, 72, 5411-5416.
- Kim H.-S., Smithies O. et Maeda N.** (1990) A physical map of the human salivary proline-rich protein gene cluster covers over 700kbp of DNA. *Genomics*, 6, 260-267.

- Kim H.S. et Maeda N.** (1986) Structures of two Haell-types genes in the human salivary proline-rich protein multigene family. *The Journal of Biological Chemistry*, 261, 6712-6718.
- Kim H.S., Lyons K.M., Saitoh E., Azen E.A., Smithies O. et Maeda N.** (1993) The structure and evolution of the human salivary proline-rich protein gene family. *Mammalian Genome*, 4, 3-14.
- Koeniger S.L., Merenbloom S.I., Valentine S.J., Jarrold M.F., Udseth H.R., Smith R.D. et Clemmer D.E.** (2006) An IMS-IMS analogue of MS-MS. *Analytical Chemistry*, 78, 4161-4174.
- Kriwacki R.W., Hengst L., Tennant L., Reed S.I. et Wright P.E.** (1996) Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 11504-11509.
- Kukuruzinska M.A. et Lennon-Hopkins K.** (1999) ALG gene expression and cell cycle progression. *Biochimica Et Biophysica Acta-General Subjects*, 1426, 359-372.
- Kurulugama R.T., Valentine S.J., Sowell R.A. et Clemmer D.E.** (2008) Development of a high-throughput IMS-IMS-MS approach for analyzing mixtures of biomolecules. *Journal of Proteomics*, 71, 318-331.
- Laemmli U.K.** (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature*, 227, 680-685.
- Lawson C., Walker C., Awford J., Biffen M., Mallinder P. et Jackson A.** (2002) Purification and characterization of recombinant rat mast cell protease 7 expressed in *Pichia pastoris*. *Protein Expression and Purification*, 25, 256-262.
- Lea A.G.H. et Arnold G.M.** (1978) The phenolics of ciders : bitterness and astringency. *Journal of the Science of Food and Agriculture*, 29, 478-483.
- Leary J.A., Schenauer M.R., Stefanescu R., Andaya A., Ruotolo B.T., Robinson C.V., Thalassinos K., Scrivens J.H., Sokabe M. et Hershey J.W.B.** (2009) Methodology for measuring conformation of solvent-disrupted protein subunits using T-WAVE ion mobility MS : an investigation into eukaryotic initiation factors. *Journal of the American Society for Mass spectrometry*, 20, 1699-1706.
- Lee C.B. et Lawless H.T.** (1991) Time-course of astringent sensations. *Chemical Senses*, 16, 225-238.
- Lesschaeve I. et Noble A.C.** (2005) Polyphenols: factors influencing their sensory properties and their effects on food and beverage preferences. *American Journal of Clinical Nutrition*, 81, 330-335.
- Leymarie N., Berg E.A., McComb M.E., O'Connor P.B., Grogan J., Oppenheim F.G. et Costello C.E.** (2002) Tandem mass spectrometry for structural characterization of proline-rich proteins : application to salivary PRP-3. *Analytical Chemistry*, 74, 4124-4132.
- Li Y.-T., Hsieh Y.-L., Henion J.D., Senko M.W., McLafferty F.W. et Ganem B.** (1993) Mass spectrometric studies on noncovalent dimers of leucine zipper peptides. *Journal of the American Chemical Society*, 115, 8409-8413.

Liao P.-C. et Allison J. (1995) Ionization processes in matrix-assisted laser desorption/ionization mass spectrometry: Matrix-dependent formation of $[M + H]^+$ vs $[M + Na]^+$ ions of small peptides and some mechanistic comments. *Journal of Mass Spectrometry*, 30, 408-423.

Liu H. et Håkansson K. (2006) Electron capture dissociation of tyrosine O-sulfated peptides complexed with divalent metal cations. *Analytical Chemistry*, 78, 7570-7576.

Londry F.A. et Hager J.W. (2003) Mass selective axial ion ejection from a linear quadrupole ion trap. *Journal of the American Society for Mass Spectrometry*, 14, 1130-1147.

Loo J.A. et Ogorzalek-Loo R.R. (1997) Electrospray ionization mass spectrometry of peptides and proteins. In: R. B. Cole (ed), *Electrospray ionization mass spectrometry*. John Wiley and Sons, Inc., New York. pp. 385-419.

Loo J.A. et Sannes-Lowery K.A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews*, 16, 1-23.

Loo J.A. (2000) Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. *International Journal of Mass Spectrometry*, 200, 175-186.

Lu Y. et Bennick A. (1998) Interaction of tannin with human salivary proline-rich proteins. *Archives of Oral Biology*, 43, 717-728.

Luck G., Liao H., Murray N.J., Grimmer H.R., Warminski E.E., Williamson M.P., Lilley T.H. et Haslam E. (1994) Polyphenols, astringency and proline-rich proteins. *Phytochemistry*, 37, 357-371.

Lyman B.J. et Green B.G. (1990) Oral astringency: effects of repeated exposure and interactions with sweeteners. *Chemical Senses*, 15, 151-164.

Lyons K.M., Azen E.A., Goodman P.A. et Smithies O. (1988a) Many protein products from a few loci: assignment of human salivary proline-rich proteins to specific loci. *Genetics*, 120, 255-265.

Lyons K.M., Stein J.H. et Smithies O. (1988b) Length polymorphisms in human proline-rich protein genes generated by intragenic unequal crossing over. *Genetics*, 120, 267-278.

Maeda N. (1985) Inheritance of the human salivary proline-rich proteins : a reinterpretation in terms of six loci forming two subfamilies. *Biochemical Genetics*, 23, 455-464.

Maeda N., Kim H.S., Azen E.A. et Smithies O. (1985) Differential RNA splicing and post-translational cleavages in the human salivary proline-rich protein gene system. *The Journal of Biological Chemistry*, 260, 1123-1130.

March R.E. (1997) An introduction to quadrupole ion trap mass spectrometry. *Journal of Mass Spectrometry*, 32, 351-369.

Martinet W., Saelens X., Deroo T., Neiryneck S., Contreras R., Jou W.M. et Fiers W. (1997) Protection of mice against a lethal influenza challenge by immunization with yeast-derived recombinant influenza neuraminidase. *European Journal of Biochemistry*, 247, 332-338.

Matsuo R. (2000) Role of saliva in the maintenance of taste sensitivity. *Critical Reviews in Oral Biology and Medicine*, 11, 216-229.

Maury C., Sarni-Manchado P., Lefebvre S., Cheynier V. et Moutonet M. (2001a) Influence of fining with different molecular weight gelatins on proanthocyanidin composition and perception of wines. *American Journal of Enology and Viticulture*, 52, 140-145.

Maury C., Sarni-Manchado P., Lefebvre S., Cheynier V. et Moutonet M. (2001b) Proanthocyanidin composition of wines after various kinds of fining treatments, In *Vino Analytica Scientia*, Bordeaux, France. pp. P-063.

Maury C., Sarni-Manchado P., Lefebvre S., Cheynier V. et Moutonet M. (2003) Influence of fining with plant proteins on proanthocyanidin composition of red wines. *American Journal of Enology and Viticulture*, 54, 105-111.

McArthur C., Sanson G.D. et Beal A.M. (1995) Salivary proline-rich proteins in mammals: role in oral homeostasis and counteracting dietary tannin. *Journal of Chemical Ecology*, 21, 663-691.

McLafferty F.W. (1981) Tandem mass spectrometry. *Science*, 214, 280-7.

McLafferty F.W., Horn D.M., Breuker K., Ge Y., Lewis M.A., Cerda B., Zubarev R.A. et Carpenter B.K. (2001) Electron capture dissociation of gaseous multiply charged ions by Fourier-transform ion cyclotron resonance. *Journal of the American Society for Mass Spectrometry*, 12, 245-249.

McLafferty F.W. (2008) Mass spectrometry across the sciences. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 18088-18089.

McLean J.A., Ruotolo B.T., Gillig K.J. et Russell D.H. (2005) Ion mobility-mass spectrometry : a new paradigm for proteomics. *International Journal of Mass Spectrometry*, 240, 301-315.

McLuckey S.A., Van Berkel G.J., Goeringer D.E. et Glish G.L. (1994) Ion trap mass spectrometry using high-pressure ionization. *Analytical Chemistry*, 66, 737A-743A.

McManus J.P., Davis K.G., Beart J.E., Galfney S.H., Lilley T.H. et Haslam E. (1985) Polyphenol interactions. Part 1. Introduction; some observations on the reversible complexation of polyphenols with proteins and polysaccharides. *Journal of Chemical Society Perkin Transactions II*, 1429-1438.

Medina N., Huth-Fehre T., Westman A. et Sundqvist B.U.R. (1994) Matrix-assisted laser desorption : dependence of the threshold fluence on analyte concentration. *Organic Mass Spectrometry*, 29, 207-209.

Mehansho H. et Carlson D.M. (1983) Induction of protein and glycoprotein synthesis in rat submandibular glands by isoproterenol. *The Journal of Biological Chemistry*, 258, 6616-6620.

Mehansho H., Clements S., Sheares B.T., Smith S. et Carlson D.M. (1985) Induction of proline-rich glycoprotein synthesis in mouse salivary glands by isoproterenol and by tannins. *The Journal of Biological Chemistry*, 260, 4418-4423.

- Mehansho H., Ann D.K., Butler L.G., Rogler J. et Carlson D.M.** (1987a) Induction of proline-rich proteins in hamster salivary glands by isoproterenol treatment and unusual growth inhibition by tannins. *The Journal of Biological Chemistry*, 262, 12344-12350.
- Mehansho H., Butler L.G. et Carlson D.M.** (1987b) Dietary tannins and salivary proline-rich proteins : interactions, induction, and defense mechanisms. *Annual Review of Nutrition*, 7, 423-440.
- Mehansho H., Asquith T.N., Butler L.G., Rogler J.C. et Carlson D.M.** (1992) Tannin-mediated induction of proline-rich protein synthesis. *Journal of Agricultural and Food Chemistry*, 40, 93-97.
- Merenbloom S.I., Koeniger S.L., Bohrer B.C., Valentine S.J. et Clemmer D.E.** (2008) Improving the efficiency of IMS-IMS by a combing technique. *Analytical Chemistry*, 80, 1918-1927.
- Merenbloom S.I., Glaskin R.S., Henson Z.B. et Clemmer D.E.** (2009) High-resolution ion cyclotron mobility spectrometry. *Analytical Chemistry*, 81, 1482-1487.
- Mirgorodskaya E., Roepstorff P. et Zubarev R.A.** (1999) Localization of O-glycosylation sites in peptides by electron capture dissociation in a Fourier transform mass spectrometer. *Analytical Chemistry*, 71, 4431-4436.
- Mitaru B.N., Reichert R.D. et Blair R.** (1984) The binding of dietary proteins by Sorghum tannins in the digestive tract of pigs. *Journal of Nutrition*, 114, 1787-1796.
- Mole S., Butler L.G. et Iason G.** (1990) Defense against dietary tannins in herbivores: a survey for proline-rich salivary proteins in mammals. *Biochemical Systematics and Ecology*, 18, 287-293.
- Mouls L., Aubagnac J.-L., Martinez J. et Enjalbal C.** (2007) Low energy peptide fragmentations in an ESI-Q-tof type mass spectrometer. *Journal of Proteome Research*, 6, 1378-1391.
- Murray N.J. et Williamson M.P.** (1994) Conformational study of a salivary proline-rich protein repeat sequence. *European Journal of Biochemistry*, 219, 915-921.
- Murray N.J., Williamson M.P., Lilley T.H. et Haslam E.** (1994) Study of the interaction between salivary proline-rich proteins and a polyphenol by ¹H-NMR spectroscopy. *European Journal of Biochemistry*, 219, 923-935.
- Murray P.A., Prakobphol A., Lee T., Hoover C.I. et Fisher S.J.** (1992) Adherence of oral streptococci to salivary glycoproteins. *Infection and Immunity*, 60, 31-38.
- Naish M., Clifford M.N. et Birch G.G.** (1993) Sensory astringency of 5-O-caffeoylquinic acid, tannic acid and grape seed tannin by a time-intensity procedure. *Journal of the Science of Food and Agriculture*, 61, 57-64.
- Nayak A. et Carpenter G.H.** (2008) A physiological model of tea-induced astringency. *Physiology & Behavior*, 95, 290-294.
- Niho N., Shibutani M., Tamura T., Toyoda K., Uneyama C., Takahashi N. et Hirose M.** (2001) Subchronic toxicity study of gallic acid by oral administration in F344 rats. *Food and Chemical Toxicology*, 39, 1063-1070.

- O'Neill R.A.** (1996) Enzymatic release of oligosaccharides from glycoproteins for chromatographic and electrophoretic analysis. *Journal of Chromatography A*, 720, 201-215.
- Oppenheim F.G., Hay D.I., Smith D.J., Offner G.D. et Troxler R.F.** (1987) Molecular basis of salivary proline-rich protein and peptide synthesis : cell-free translations and processing of human and macaque statherin mRNAs and partial amino acid sequence of their signal peptides. *Journal of Dental Research*, 66, 462-466.
- Pangborn R.M., Lewis M.J. et Yamashita J.F.** (1983) Comparison of time-intensity with category scaling of bitterness of iso- α -acids in model systems and in beer. *Journal of the Institute of Brewing*, 89, 349-355.
- Paramasivam M., Saravanan K., Uma K., Sharma S., Singh T.P. et Srinivasan A.** (2002) Expression, purification, and characterization of equine lactoferrin in *Pichia pastoris*. *Protein Expression and Purification*, 26, 28-34.
- Parker D., Rivera M., Zor T., Henrion-Caude A., Radhakrishnan I., Kumar A., Shapiro L.H., Wright P.E., Montminy M. et Brindle P.K.** (1999) Role of secondary structure in discrimination between constitutive and inducible activators. *Molecular and Cellular Biology*, 19, 5601-5607.
- Pascal C., Bigey F., Ratomahenina R., Boze H., Moulin G. et Sarni-Manchado P.** (2006) Overexpression and characterization of two human salivary proline rich proteins. *Protein Expression and Purification*, 47, 524-532.
- Pascal C., Poncet-Legrand C., Imberty A., Gautier C., Sarni-Manchado P., Cheynier V. et Vernhet A.** (2007) Interactions between a non glycosylated human proline-rich protein and flavan-3-ols are affected by protein concentration and polyphenol/protein ratio. *Journal of Agricultural and Food Chemistry*, 55, 4895-4901.
- Peleg H., Gacon K., Schlich P. et Noble A.C.** (1999) Bitterness and astringency of flavan-3-ol monomers, dimers and trimers. *Journal of The Science of Food and Agriculture*, 79, 1123-1128.
- Permyakov S.E., Millett I.S., Doniach S., Permyakov E.A. et Uversky V.N.** (2003) Natively unfolded C-terminal domain of caldesmon remains substantially unstructured after the effective binding to calmodulin. *Proteins: Structure, Function, and Genetics*, 53, 855-Na.
- Plet B. et Schmitter J.M.** (2009) Wine astringency approached by ESI-IT-MS. In: R. E. March and J. F. J. Todd (eds.), Practical aspects of trapped ion mass spectrometry, vol. 5. CRC Press, Boca Raton, FL. pp. 153-167.
- Poncet-Legrand C., Cartalade D., Putaux J.-L., Cheynier V. et Vernhet A.** (2003) Flavan-3-ol aggregation in model ethanolic solutions : incidence of polyphenol structure, concentration, ethanol content and ionic strength. *Langmuir*, 19, 10563-10572.
- Pontius B.W.** (1993) Close encounters : why unstructured, polymeric domains can increase rates of specific macromolecular association. *Trends in Biochemical Sciences*, 18, 181-186.
- Porter L.J.** (1988) Flavans and proanthocyanidins. In: J. B. Harborne (ed), The flavonoids : advances in research since 1980. Chapman and Hall, New York. pp. 21-60.
- Porter L.J.** (1989) Chap. 11: Tannins. In: J. B. Harborne (ed), Methods in plant biochemistry: Plant phenolics, vol. 1. Academic Press Limited. pp. 389-419.

- Pramanik B.N., Bartner P.L., Mirza U.A., Liu Y.H. et Ganguly A.K.** (1998) Electrospray ionization mass spectrometry for the study of non-covalent complexes : an emerging technology. *Journal of Mass Spectrometry*, 33, 911-920.
- Prieur C., Cheynier V. et Moutounet M.** (1994) Turbidimetric determination of protein-tannin interactions, Journées Internationales d'Etudes du Groupe Polyphénols, Palma de Majorque, Espagne.
- Puretzky A.A., Geohegan D.B., Hurst G.B., Buchanan M.V. et Luk'yanchuk B.S.** (1999) Imaging of vapor plumes produced by matrix assisted laser desorption : a plume sharpening effect. *Physical Review Letters*, 83, 444-447.
- Rademacher T.W., Parekh R.B. et Dwek R.A.** (1988) Glycobiology. *Annual Review of Biochemistry*, 57, 785-838.
- Receveur-Brechot V., Bourhis J.M., Uversky V.N., Canard B. et Longhi S.** (2006) Assessing protein disorder and induced folding. *PROTEINS: structure, fonction, and bioinformatics*, 62, 24-45.
- Receveur-Bréchot V. et Karlin D.** (2005) Le désordre clé de voûte des protéines La recherche, vol. 387. pp. 52-55.
- Renfrow M.B., Cooper H.J., Tomana M., Kulhavy R., Hiki Y., Toma K., Emmett M.R., Mesteccky J., Marshall A.G. et Novak J.** (2005) Determination of aberrant O-glycosylation in the IgA1 hinge region by electron capture dissociation Fourier transform-ion cyclotron resonance mass spectrometry. *Journal of Biological Chemistry*, 280, 19136-19145.
- Ricardo da Silva J.M., Cheynier V., Souquet J.M., Moutounet M., Cabanis J.C. et Bourzeix M.** (1991a) Interaction of grape seed procyanidins with various proteins in relation to wine fining. *Journal of the Science of Food and Agriculture*, 57, 111-125.
- Ricardo da Silva J.M., Rigaud J., Cheynier V., Cheminat A. et Moutounet M.** (1991b) Procyanidin dimers and trimers from grape seeds. *Phytochemistry*, 30, 1259-1264.
- Robbins C.T., Mole S., Hagerman A.E. et Hanley T.A.** (1987) Role of tannins in defending plants against ruminants; reduction of dry matter digestion? *Ecology*, 68, 1606-1615.
- Robbins C.T., Hagerman A.E., Austin P.J., Arthur C.M. et Hanley T.A.** (1991) Variation in mammalian physiological responses to a condensed tannin and its ecological implications. *Journal of Mammalogy*, 72, 480-486.
- Robichaud J.L. et Noble A.C.** (1990) Astringency and bitterness of selected phenolics in wine. *Journal of the Science of Food and Agriculture*, 53, 343-353.
- Robinson C.V., Chung E.W., Kragelund B.B., Knudsen J., Aplin R.T., Poulsen F.M. et Dobson C.M.** (1996) Probing the nature of noncovalent interactions by mass spectrometry. A study of protein-CoA ligand binding and assembly. *Journal of the American Chemical Society*, 118, 8646-8653.
- Robinson C.V.** (2002) Protein complexes take flight. *Nature Structural Biology*, 9, 505-506.
- Roepstorff P.** (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomedical Mass Spectrometry*, 11, 601.

Rogniaux H., Sanglier S., Strupat K., Azza S., Roitel O., Ball V., Tritsch D., Branlant G. et Van Dorsselaer A. (2001) Mass spectrometry as a novel approach to probe cooperativity in multimeric enzymatic systems. *Analytical Biochemistry*, 291, 48-61.

Rossi J.A. et Singleton V.L. (1966) Flavor effect and adsorptive properties of purified fractions of grape-seed phenols. *American Journal of Enology and Viticulture*, 17, 240-246.

Ruitenbergh K.M., Gilkerson J.R., Wellington J.E., Love D.N. et Whalley J.M. (2001) Equine herpesvirus 1 glycoprotein D expressed in *Pichia pastoris* is hyperglycosylated and elicits a protective immune response in the mouse model of EHV-1 disease. *Virus Research*, 79, 125-135.

Ruotolo B.T., Giles K., Campuzano I., Sandercock A.M., Bateman R.H. et Robinson C.V. (2005) Evidence for macromolecular protein rings in the absence of bulk water. *Science*, 310, 1658-1661.

Ruotolo B.T. et Robinson C.V. (2006) Aspects of native proteins are retained in vacuum. *Current Opinion in Chemical Biology*, 10, 402-408.

Ruotolo B.T., Hyung S.-J., Robinson P.M., Giles K., Bateman R.H. et Robinson C.V. (2007) Ion mobility-mass spectrometry reveals long-lived, unfolded intermediates in the dissociation of protein complexes¹³. *Angewandte Chemie International Edition*, 46, 8001-8004.

Ruotolo B.T., Benesch J.L.P., Sandercock A.M., Hyung S.-J. et Robinson C.V. (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nature Protocols*, 3, 1139-1152.

Salgues M. et Razungles A. (1983) Le collage des vins. *Vititechniques*, 64, 32-35.

Sanglier S., Van Dorsselaer A., Vivat-Hannah V. et Renaud J.-P. (2005) La spectrométrie de masse des assemblages non-covalents en biologie : aspects fondamentaux et applications à la recherche pharmaceutique. *Spectra Analyse*, 245.

Sarni-Manchado P. et Cheynier V. (1999) Phenolic structure and astringency. *Journal International de la Vigne et du Vin*, N° Hors série: Wine tasting, 111-118.

Sarni-Manchado P., Cheynier V. et Moutounet M. (1999a) Interactions of grape seed tannins with salivary proteins. *Journal of Agricultural and Food Chemistry*, 47, 42-47.

Sarni-Manchado P., Deleris A., Avallone S., Cheynier V. et Moutounet M. (1999b) Analysis and characterization of wine condensed tannins precipitated by protein used as fining agent in enology. *American Journal of Enology and Viticulture*, 50, 81-86.

Sarni-Manchado P. et Cheynier V. (2002) Study of noncovalent complexation between catechin derivatives and peptide by electrospray ionization-mass spectrometry (ESI-MS). *Journal of Mass Spectrometry*, 37, 609-616.

Sarni-Manchado P., Canals-Bosch J., Mazerolles G. et Cheynier V. (2008) Influence of the glycosylation of human salivary proline-rich proteins on their interactions with condensed tannins. *Journal of Agricultural and Food Chemistry*, 56, 9563-9569.

Scharbert S., Jezussek M. et Hofmann T. (2004) Evaluation of the taste contribution of theaflavins in black tea infusions using the taste activity concept. *European Food Research And Technology*, 218, 442-447.

- Shakin-Eshleman S.H., Spitalnik S.L. et Kasturi L.** (1996) The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core-glycosylation efficiency. *Journal of Biological Chemistry*, 271, 6363-6366.
- Sharon M. et Robinson C.V.** (2007) The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annual Review of Biochemistry*, 76, 167-193.
- Shelikoff M., Sinskey A.J. et Stephanopoulos G.** (1996) A modeling framework for the study of protein glycosylation. *Biotechnology and Bioengineering*, 50, 73-90.
- Shi S.D.H., Hemling M.E., Carr S.A., Horn D.M., Lindh I. et McLafferty F.W.** (2000) Phosphopeptide/phosphoprotein mapping by electron capture dissociation mass spectrometry. *Analytical Chemistry*, 73, 19-22.
- Shimada T. et Saitoh T.** (2003) Negative effects of acorns on the wood mouse *Apodemus speciosus*. *Population Ecology*, 45, 7-17.
- Shimada T.** (2006) Salivary proteins as a defense against dietary tannins. *Journal of Chemical Ecology*, 32, 1149-1163.
- Siebert K.J.** (1999) Effects of protein-polyphenol interactions on beverage haze, stabilization and analysis. *Journal of Agricultural and Food Chemistry*, 47, 353-362.
- Siebert K.J.** (2006) Haze formation in beverages. *LWT - Food Science and Technology European Symposium on Apple Processing*, 39, 987-994.
- Simon C., Barathieu K., Laguerre M., Schmitter J.M., Fouquet E., Pianet I. et Dufourc E.J.** (2003) Three-dimensional structure and dynamics of wine tannin-saliva protein complexes. A multitechnique approach. *Biochemistry*, 42, 10385-10395.
- Skopec M.M., Hagerman A.E. et Karasov W.H.** (2004) Do salivary proline-rich proteins counteract dietary hydrolyzable tannin in laboratory rats? *Journal of Chemical Ecology*, 30, 1679-1692.
- Spencer C.M., Cai Y., Martin R., Gaffney S.H., Goulding P.N., Magnolato D., Lilley T.H. et Haslam E.** (1988) Polyphenol complexation-some thoughts and observations. *Phytochemistry*, 27, 2397-2409.
- Spengler B. et Cotter R.J.** (1990) Ultraviolet laser desorption/ionization mass spectrometry of proteins above 100,000 daltons by pulsed ion extraction time-of-flight analysis. *Analytical Chemistry*, 62, 793-796.
- Stubbs M., Chan J., Kwan A., So J., Barchynsky U., Rassouli-Rahsti M., Robinson R. et Bennick A.** (1998) Encoding of human basic and glycosylated proline-rich proteins by the PRB gene complex and proteolytic processing of their precursor proteins. *Archives of Oral Biology*, 43, 753-770.
- Tabak L.A.** (1995) In defense of the oral cavity: structure, biosynthesis, and function of salivary mucins. *Annual Review of Physiology*, 57, 547-564.
- Tanaka K., Waki H., Ido Y., Akita S., Yoshida Y., Yoshida T. et Matsuo T.** (1988) Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 2, 151-153.

- Taylor G.** (1964) Disintegration of water drops in an electric field. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 280, 383-397.
- Thomson B.A. et Iribarne J.V.** (1979) Field induced ion evaporation from liquid surfaces at atmospheric pressure. *The Journal of Chemical Physics*, 71, 4451.
- Thorngate J.H. et Noble A.C.** (1995) Sensory evaluation of bitterness and astringency of 3R(-)-epicatechin and 3S(+)-catechin. *Journal of the Science of Food and Agriculture*, 67, 531-535.
- Tompa P.** (2002) Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, 27, 527-533.
- Tompa P.** (2003a) Intrinsically unstructured proteins evolve by repeat expansion. *Bioessays*, 25, 847-855.
- Tompa P.** (2003b) The functional benefits of protein disorder. *Journal of Molecular Structure (Theochem)*, 361-371.
- Tompa P.** (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Letters*, 579, 3346-3354.
- Trautschold I., Werle E. et Zickgraf-Rüdel G.** (1967) Trasylol. *Biochemical Pharmacology*, 16, 59-72.
- Tsai H. et Bobek L.A.** (1998) Human salivary histatins: promising anti-fungal therapeutic agents. *Critical Reviews in Oral Biology and Medicine*, 9, 480-497.
- Uversky V.N.** (2002a) Natively unfolded proteins: a point where biology waits for physics. *Protein Science*, 11, 739-756.
- Uversky V.N.** (2002b) What does it mean to be natively unfolded? *European Journal of Biochemistry*, 269, 2-12.
- Van den Heuvel R.H.H. et Heck A.J.R.** (2004) Native protein mass spectrometry: from intact oligomers to functional machineries. *Current Opinion in Chemical Biology*, 8, 519-526.
- Van Valen L.** (1973) A new evolutionary law. *Evolutionary Theory*, 1, 1-30.
- Vernhet A., Pellerin P., Prieur C., Oszmianski J. et Moutounet M.** (1996) Charge properties of some grape and wine polysaccharide and polyphenolic fractions. *American Journal of Enology and Viticulture*, 47, 25-30.
- Vidal S., Francis L., Guyot S., Marnet N., Kwiatkowski M., Gawel R., Cheynier V. et Waters E.J.** (2003) The mouth-feel properties of grape and apple proanthocyanidins in a wine-like medium. *Journal of the Science of Food and Agriculture*, 83, 564-573.
- Wallon G., Rappsilber J., Mann M. et Serrano L.** (2000) Model for stathmin/OP18 binding to tubulin. *EMBO Journal*, 19, 213-222.
- Ward P., Paz E. et Conneely O.** (2005) Multifunctional roles of lactoferrin : a critical overview. *Cellular and Molecular Life Sciences*, 62, 2540-2548.
- Waterman P.G. et Mole S.** (1994) Analysis of phenolic plant metabolites. In: B. Scientific (ed), *Analysis of Phenolic Plant Metabolites*, Oxford, Boston. pp. 238.

- Whitelegge J.P., Zabrouskov V., Halgand F., Souda P., Bassilian S., Yan W., Wolinsky L., Loo J.A., Wong D.T.W. et Faull K.F.** (2007) Protein-sequence polymorphisms and post-translational modifications in proteins from human saliva using top-down Fourier-transform ion cyclotron resonance mass spectrometry. *International Journal of Mass Spectrometry*, 268, 190-197.
- Williamson M.P.** (1994) The structure and function of proline-rich regions in proteins. *Biochemical Journal*, 297, 249-260.
- Wlodawer A., Walter J., Huber R. et Sjölin L.** (1984) Structure of bovine pancreatic trypsin inhibitor : results of joint neutron and X-ray refinement of crystal form II. *Journal of Molecular Biology*, 180, 301-329.
- Wolfenden R.** (1983) Waterlogged molecules. *Science*, 222, 1087-1093.
- Wright P.E. et Dyson H.J.** (1999) Intrinsically unstructured proteins : re-assessing the protein structure-function paradigm. *Journal of Molecular Biology*, 293, 321-331.
- Xie Y., Zhang J., Yin S. et Loo J.A.** (2006) Top-down ESI-ECD-FT-ICR mass spectrometry localizes noncovalent protein-ligand binding sites. *Journal of American Chemical Society*, 128, 14432-14433.
- Yan B., Zhang W., Ding J. et Gao P.** (1999) Sequence pattern for the occurrence of N-glycosylation in proteins. *Journal of Protein Chemistry*, 18, 511-521.
- Yan Q. et Bennick A.** (1995) Identification of histatins as tannin-binding proteins in human saliva. *Biochemical Journal*, 311, 341-347.
- Yin S. et Loo J.A.** (2010) Elucidating the site of protein-ATP binding by top-down mass spectrometry. *Journal of the American Society for Mass spectrometry*, 21, 899-907.
- Yokotsuka K. et Singleton V.L.** (1995) Interactive precipitation between phenolic fractions and peptides in wine-like model solutions : turbidity, particle size, and residual content as influenced by pH, temperature and peptide concentration. *American Journal of Enology and Viticulture*, 46, 329-338.
- Zhang S., VanPelt C.K. et Wilson D.B.** (2003) Quantitative determination of noncovalent binding interactions using automated nano-electrospray mass spectrometry. *Analytical Chemistry*, 75, 3010-3018.
- Zhou H.-X.** (2003) Quantitative account of the enhanced affinity of two linked scFvs specific for different epitopes on the same antigen. *Journal of Molecular Biology*, 329, 1-8.
- Zhou M., Sandercock A.M., Fraser C.S., Ridlova G., Stephens E., Schenauer M.R., Yokoi-Fong T., Barsky D., Leary J.A., Hershey J.W., Doudna J.A. et Robinson C.V.** (2008) Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 18139-18144.
- Zhou M. et Robinson C.V.** (2010) When proteomics meets structural biology. *Trends in Biochemical Sciences*, In Press, Corrected Proof.

Zhu M., Phillipson D., Greengrass P.M., Bowery N.E. et Cai Y. (1997) Plant polyphenols : biologically active compounds or non selective binders to protein? *Phytochemistry*, 44, 441-447.

Zubarev R.A., Kelleher N.L. et McLafferty F.W. (1998) Electron capture dissociation of multiply charged protein cations. A Non-ergodic process. *Journal of the American Chemical Society*, 120, 3265-3266.

Zubarev R.A. (2004) Electron-capture dissociation tandem mass spectrometry. *Current Opinion in Biotechnology*, 15, 12-16.

Zucker W.V. (1983) Tannins: Does structure determine function? An ecological perspective. *The American Naturalist*, 121, 335-365.

Annexe 1

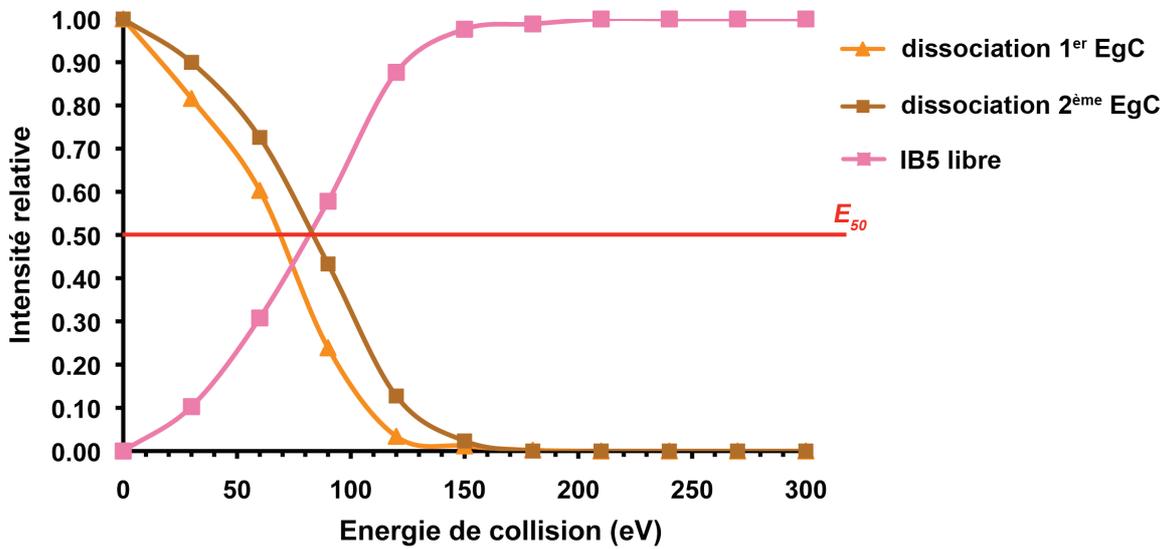


Figure 94. Courbes de dissociation des tanins du complexe $[\text{IB5}\cdot\text{EgC}\cdot\text{EgC}]^{6+}$ et de libération d'IB5.

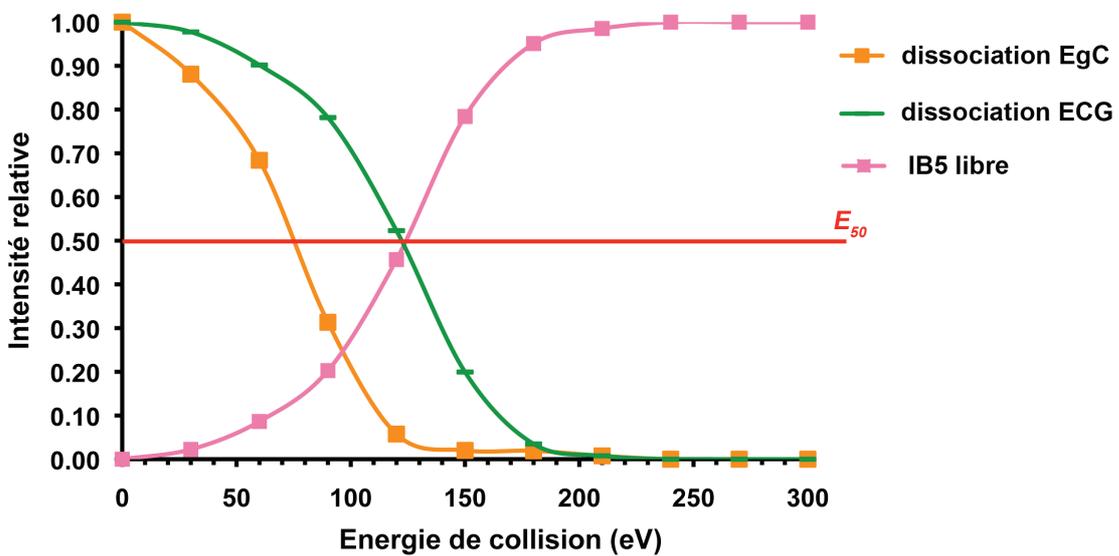


Figure 95. Courbes de dissociation des tanins du complexe $[\text{IB5}\cdot\text{EgC}\cdot\text{ECG}]^{6+}$ et de libération d'IB5.

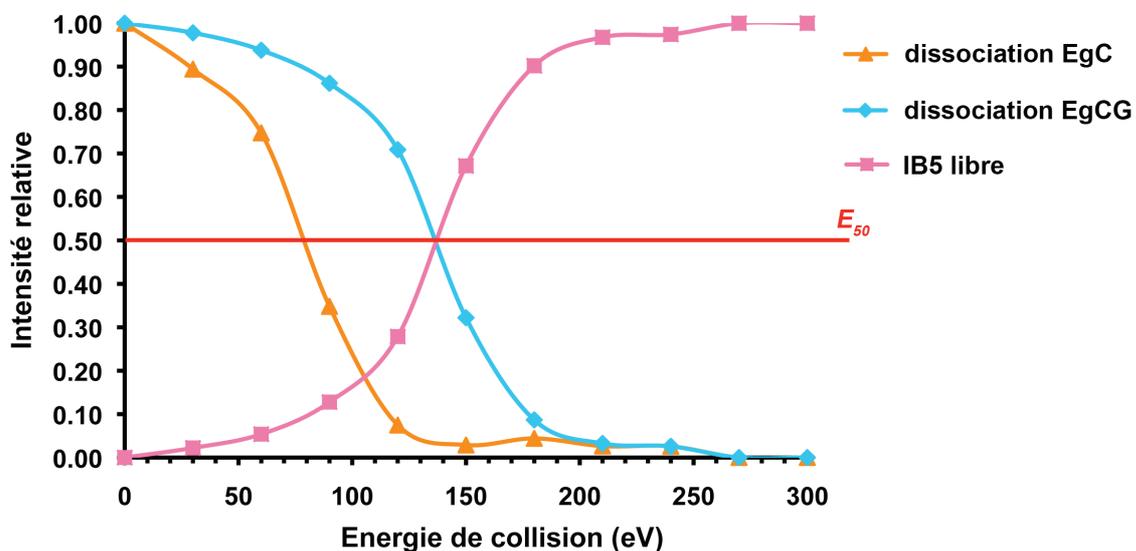


Figure 96. Courbes de dissociation des tanins du complexe $[IB5 \cdot EgC \cdot EgCG]^{6+}$ et de libération d'IB5.

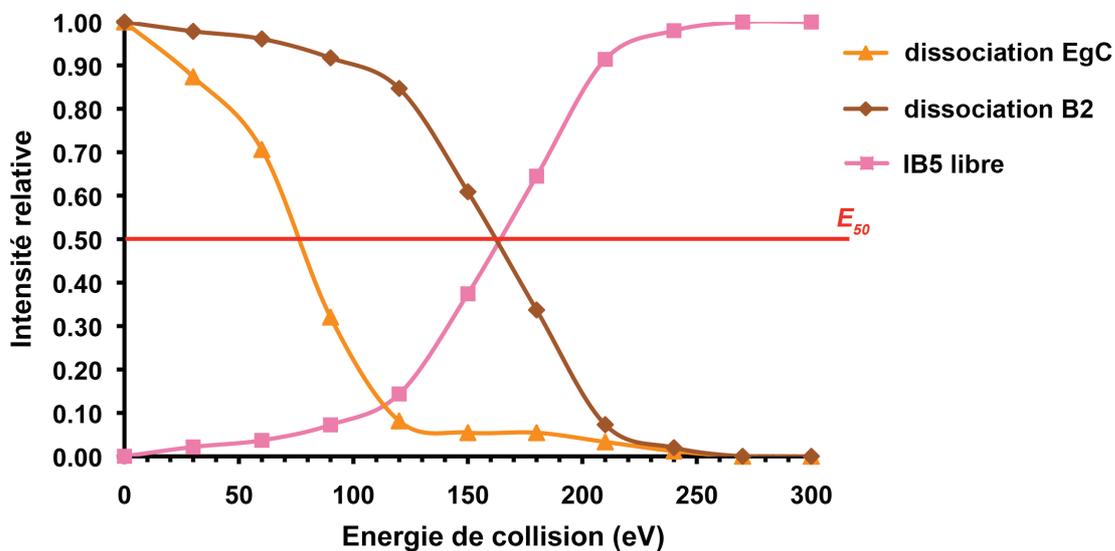


Figure 97. Courbes de dissociation des tanins du complexe $[IB5 \cdot EgC \cdot B2]^{6+}$ et de libération d'IB5.

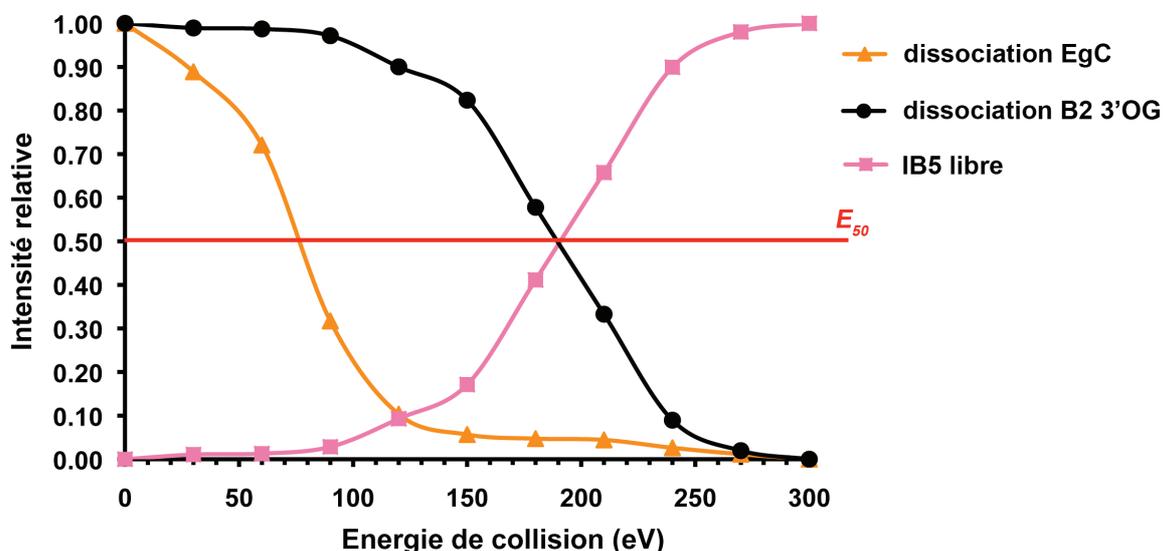


Figure 98. Courbes de dissociation des tanins du complexe [IB5•EgC•B2 3'OG]⁶⁺ et de libération d'IB5.

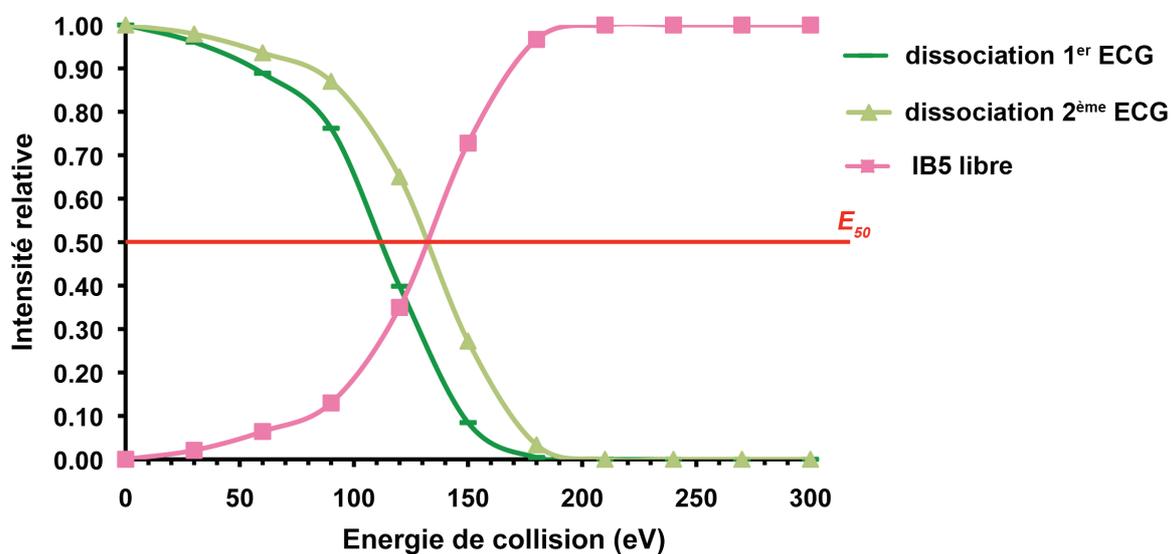


Figure 99. Courbes de dissociation des tanins du complexe [IB5•ECG•ECG]⁶⁺ et de libération d'IB5.

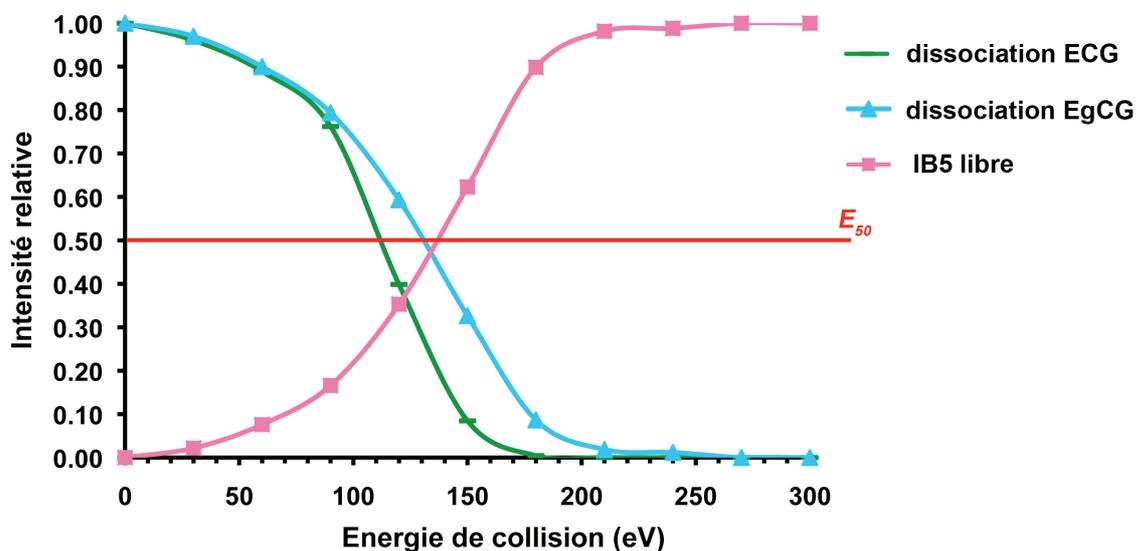


Figure 100. Courbes de dissociation des tanins du complexe [IB5•ECG•EgCG]⁶⁺ et de libération d'IB5.

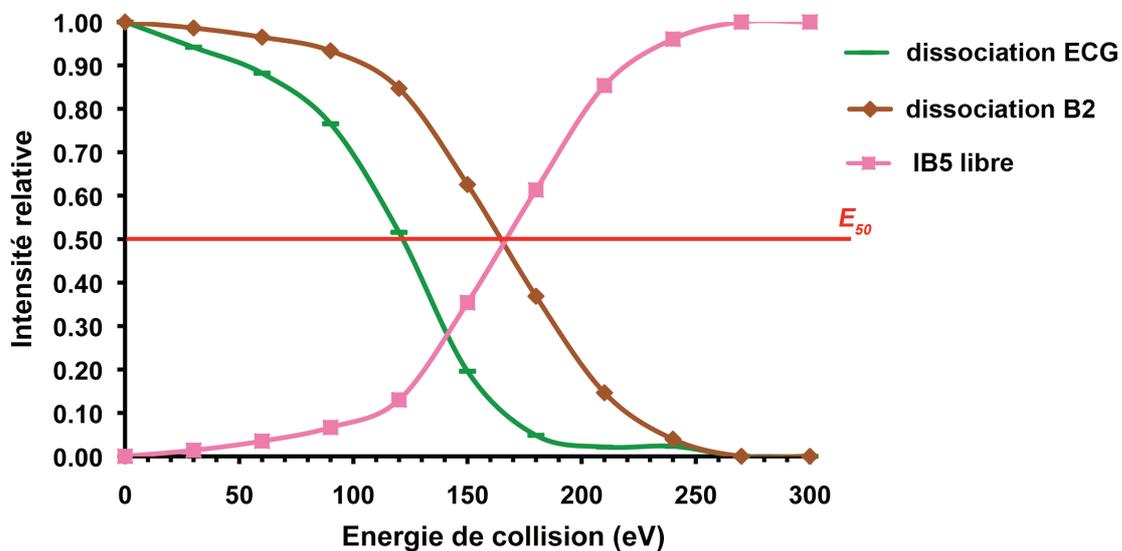


Figure 101. Courbes de dissociation des tanins du complexe [IB5•ECG•B2]⁶⁺ et de libération d'IB5.

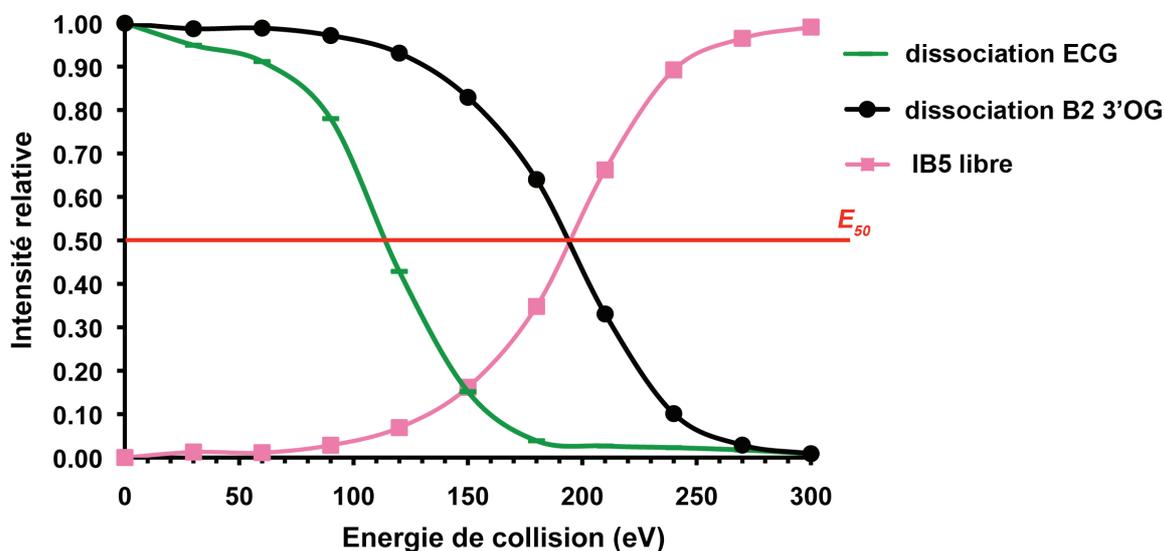


Figure 102. Courbes de dissociation des tanins du complexe $[IB5 \cdot ECG \cdot B2 \text{ 3'OG}]^{6+}$ et de libération d'IB5.

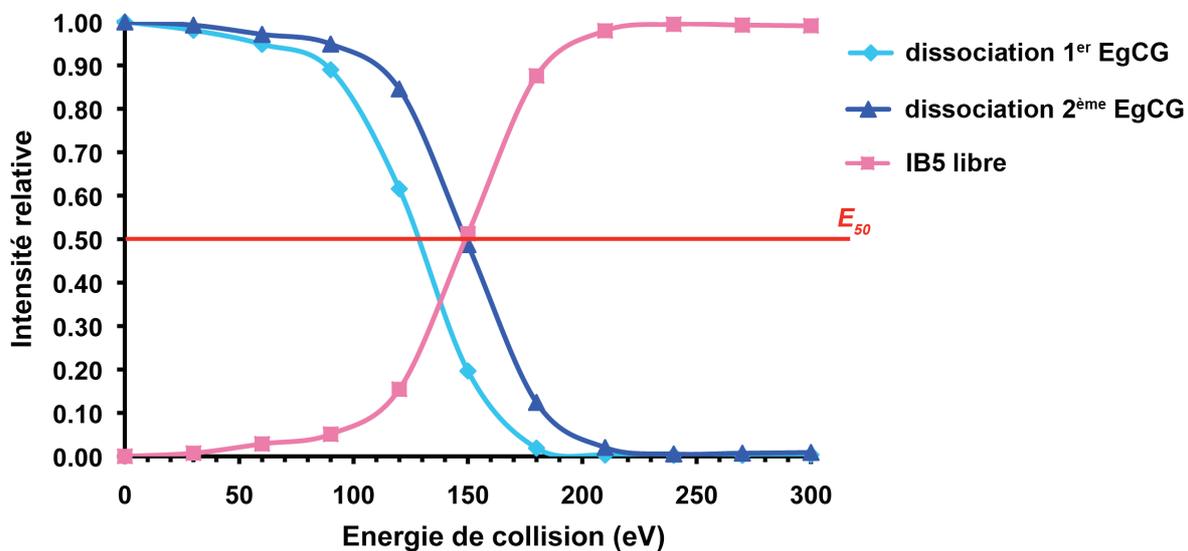


Figure 103. Courbes de dissociation des tanins du complexe $[IB5 \cdot EgCG \cdot EgCG]^{6+}$ et de libération d'IB5.

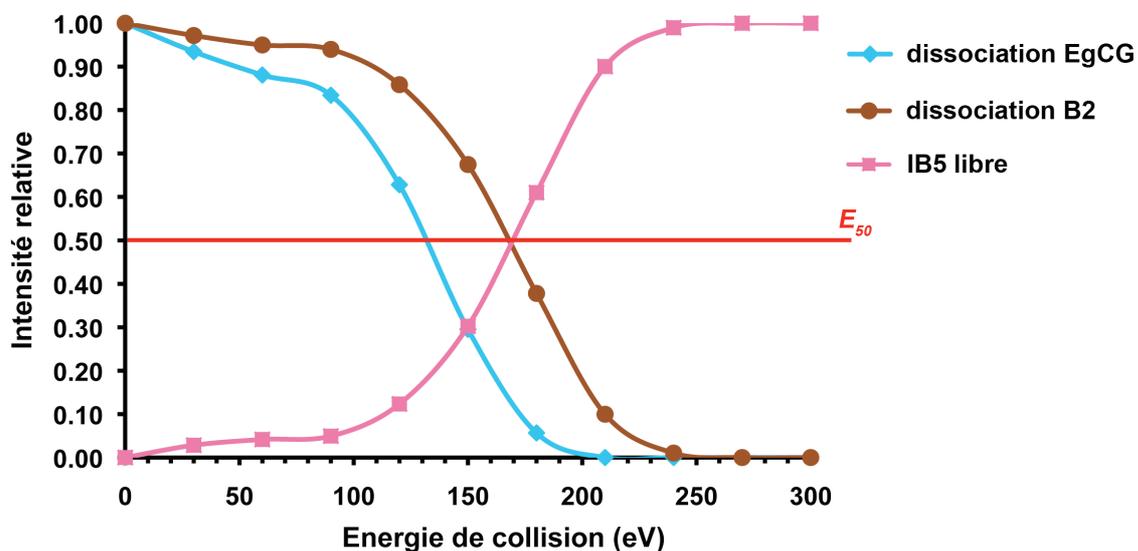


Figure 104. Courbes de dissociation des tanins du complexe $[IB5 \cdot EgCG \cdot B2]^{6+}$ et de libération d'IB5.

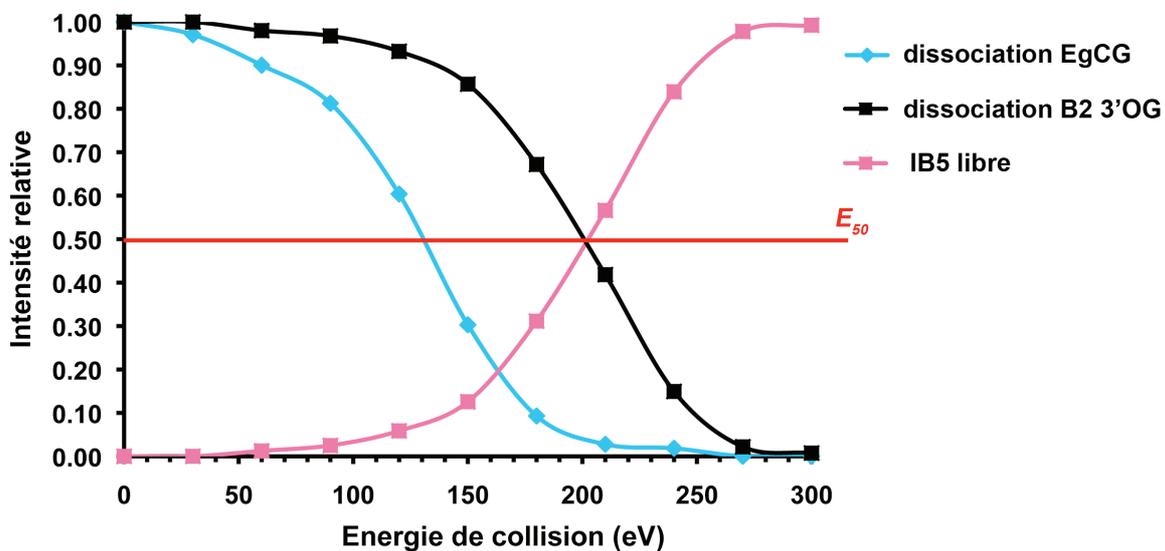


Figure 105. Courbes de dissociation des tanins du complexe $[IB5 \cdot EgCG \cdot B2 \text{ 3'OG}]^{6+}$ et de libération d'IB5.

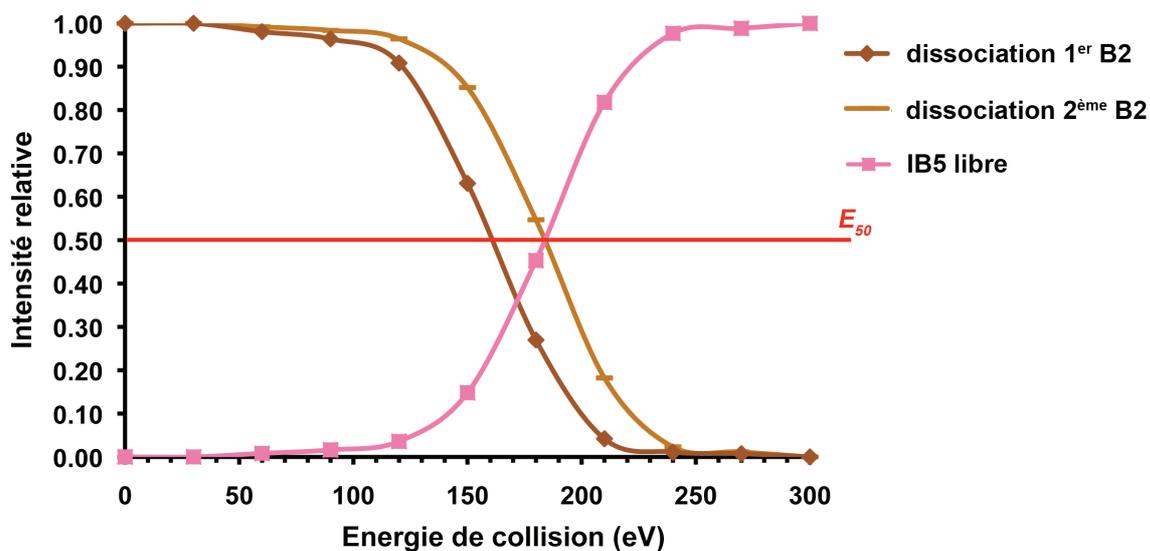


Figure 106. Courbes de dissociation des tanins du complexe $[IB5 \cdot B2 \cdot B2]^{6+}$ et de libération d'IB5.

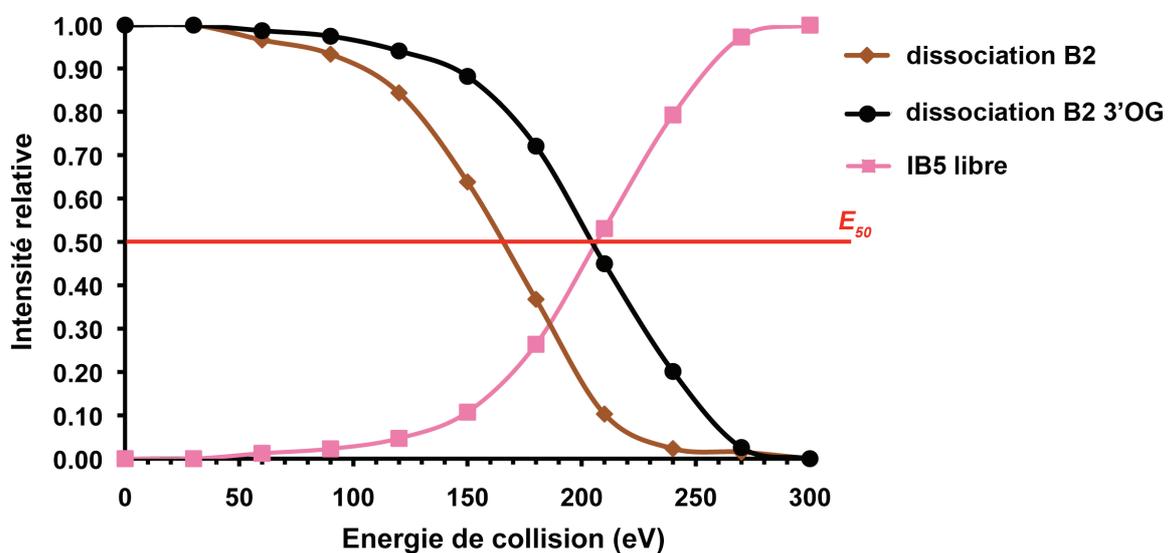


Figure 107. Courbes de dissociation des tanins du complexe $[IB5 \cdot B2 \cdot B2 \text{ 3'OG}]^{6+}$ et de libération d'IB5.

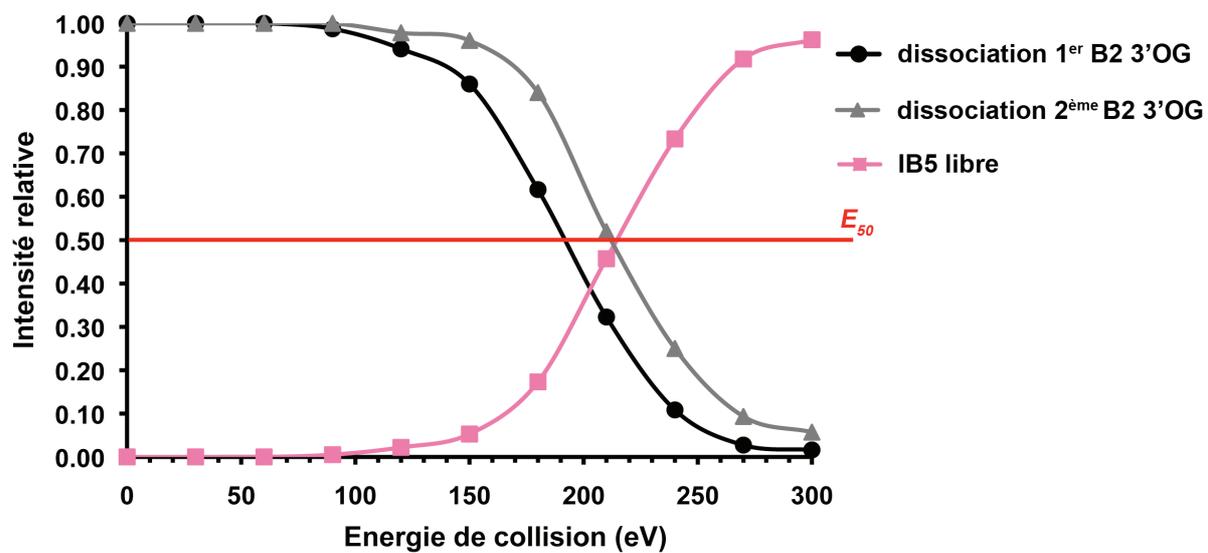


Figure 108. Courbes de dissociation des tanins du complexe [IB5•EgCG•EgCG]⁶⁺ et de libération d'IB5.