

N°d'ordre : 2701

THESE

Présentée à

L'UNIVERSITE BORDEAUX 1

ECOLE DOCTORALE DES SCIENCES CHIMIQUES

PAR CECILE SIMON

POUR OBTENIR LE GRADE DE

DOCTEUR

SPECIALITE : CHIMIE-PHYSIQUE – MENTION CHIMIE ANALYTIQUE

STRUCTURE ET DYNAMIQUE DE PROTEINES DE LA SALIVE HUMAINE EN
INTERACTION AVEC LES TANINS DU VIN DE BORDEAUX

Soutenue le 25 septembre 2003

Après avis de :

Eric Guittet	(Rapporteur)	Directeur de recherche CNRS
Victor de Freitas	(Rapporteur)	Professeur de l'Université de Porto

Devant la commission d'examen formée de :

Eric Guittet	(Rapporteur)	Directeur de recherche CNRS
Victor de Freitas	(Rapporteur)	Professeur de l'Université de Porto
Yves Glories	(Président)	Professeur de l'Université de Bordeaux 2
Jean-Marie Schmitter	(Examinateur)	Professeur de l'Université de Bordeaux 1
Erick Joël Dufourc	(Directeur de thèse)	Directeur de recherche CNRS
Isabelle Pianet	(Directeur de thèse)	Ingénieur de recherche CNRS

Table des matières

LISTE DES ABREVIATIONS.....	11
-----------------------------	----

CHAPITRE I INTRODUCTION : INTERACTION TANIN-PROTEINE : LA CLE DU « GOUT » ? 17

I. 1. LES PROTEINES DE LA SALIVE	20
I. 1. 1. Les Protéines Riches en Proline (PRP).....	20
I. 1. 1. a. La notion de multifamille.....	21
I. 1. 1. b. Rôles des PRPs	21
I. 1. 2. Les PRP basiques.....	22
I. 1. 2. a. Diversités des PRPs basiques.....	22
I. 1. 2. b. Structure primaire des PRP basiques	22
I. 1. 3. Choix d'une protéine et d'un peptide modèles	24
I. 2. LES TANINS	25
I. 2. 1. Les tanins du vin et du raisin	26
I. 2. 2. Les tanins condensés.....	27
I. 2. 2. a. Nature des proanthocyanidines et degré de polymérisation.....	27
I. 2. 3. Choix du tanin modèle	29
I. 3. L'INTERACTION TANINS-PROTEINES.....	30
I. 3. 1. Généralités	30
I. 3. 2. Principes de la complexation	31
I. 3. 3. Capacité d'interaction avec les polyphénols.....	32
I. 3. 4. Nature du site de liaison.....	33
I. 4. OBJECTIFS ET PLAN DE LA THESE	34

CHAPITRE II METHODES PHYSICO-CHIMIQUES 37

II. 1. LE DICHROÏSME CIRCULAIRE	39
II. 1. 1. Principe.....	39
II. 1. 2. Courbes standard et méthode de déconvolution.....	40
II. 1. 2. a. Méthode de déconvolution.....	40

II. 1. 2. b. Obtention des courbes standard	41
II. 2. LA SPECTROMETRIE DE MASSE (SM)	44
II. 2. 1. Le MALDI	44
II. 2. 2. L'ESI	45
II. 2. 3. L'ESI couplé à la trappe à ions.....	46
II. 3. LA SPECTROSCOPIE DE RESONANCE MAGNETIQUE NUCLEAIRE	48
II. 3. 1. Brefs rappels du principe de la RMN	48
II. 3. 2. Détermination de la structure de Protéines.....	50
II. 3. 2. a. Stratégie d'attribution	50
II. 3. 2. b. Première étape : Identification des systèmes de spin	51
II. 3. 2. c. Deuxième étape : Structure Primaire	52
II. 3. 2. d. Troisième étape : Etablissement d'une liste de contraintes	53
II. 3. 2. e. Structure tridimensionnelle.....	54
II. 3. 3. Dynamique.....	55
II. 3. 4. Coefficient de Diffusion translationnel, D	56
II. 3. 5. Spectres de protéines dans l'eau et complexées avec les tanins	57
II. 3. 5. a. Elimination de la résonance de l'eau	57
II. 3. 5. b. HR-MAS.....	58
II. 4. LA MODELISATION MOLECULAIRE.....	60
II. 4. 1. Principes de la dynamique moléculaire	60
II. 4. 2. Minimisation des structures de protéines	62
II. 4. 3. Dynamique.....	63

CHAPITRE III SYNTHÈSE D'UNE PRP DE LA SALIVE HUMAINE ET D'UN TANIN DU RAISIN 65

III. 1. PRODUCTION DE PROTEINES ET DE TANINS : SYNTHÈSE OU EXTRACTION DE PRODUITS NATURELS ?.....	67
III. 2. LA PROTEINE IB7 ET LE PEPTIDE IB7 ₁₄	68
III. 2. 1. Synthèse du peptide et de la protéine	68
III. 2. 1. a. Stratégie de la synthèse : Synthèse Peptidique en Phase Solide (SPPS).....	68
III. 2. 1. b. Etapes de la synthèse.....	69
III. 2. 1. c. Isolement et purification.....	74
III. 2. 2. Matériels et Méthodes	74
III. 2. 2. a. Synthétiseur	74

III. 2. 2. b.	Synthèse du peptide IB7 ₁₄ et clivage de la résine	75
III. 2. 2. c.	Synthèse de IB7 et clivage de la résine	76
III. 2. 2. d.	Purification par Chromatographie Liquide à Haute Performance.....	77
III. 2. 2. e.	Dessalage.....	78
III. 2. 2. f.	Vérification de la composition en masse par Spectrométrie de Masse.....	78
III. 2. 3.	Résultats et Discussion.....	79
III. 2. 3. a.	Caractérisation du peptide IB7 ₁₄	79
III. 2. 3. b.	Caractérisation de la protéine IB7.....	81
III. 2. 4.	Conclusion de la synthèse	85
III. 3.	LES TANINS DIMERIQUES	85
III. 3. 1.	Extraction des pépins de raisin.....	86
III. 3. 1. a.	Matériels et Méthodes de séparation	86
III. 3. 1. b.	Résultats et discussion.....	88
III. 3. 2.	Synthèse par couplage cationique	90
III. 3. 2. a.	Etapas de la synthèse	90
III. 3. 2. b.	Matériels et Méthodes	92
III. 3. 2. c.	Synthèse du dimère B3	92
III. 3. 3.	Choix de la méthode retenue.....	94
III. 4.	CONCLUSION DES SYNTHESSES DE PROTEINES ET DE TANINS	95
 CHAPITRE IV DETERMINATION STRUCTURALE DE IB7 ET IB7₁₄.....		97
IV. 1.	INTRODUCTION	99
IV. 1. 1.	Structures de PRPs : pelote ou hélice de type II ?	99
IV. 1. 2.	La Proline et l'hélice de type II.....	100
IV. 1. 3.	Critères déterminants la formation d'hélice de type II	103
IV. 1. 4.	Stratégie de l'étude.....	103
IV. 2.	MATERIELS ET METHODES.....	104
IV. 2. 1.	Dichroïsme Circulaire (DC).....	104
IV. 2. 2.	Résonance Magnétique Nucléaire (RMN).....	105
IV. 2. 3.	Modélisation Moléculaire (MM)	106
IV. 3.	RESULTATS :	108
IV. 3. 1.	Structure secondaire de IB7 et IB7 ₁₄ par DC	108
IV. 3. 1. a.	Spectres dichroïques de IB7 et IB7 ₁₄	108

IV. 3. 1. b.	Auto-association des protéines ?	110
IV. 3. 1. c.	Effet de la présence d'un solvant « structurant » sur la protéine IB7	110
IV. 3. 1. d.	Structure secondaire de IB7 et IB7 ₁₄	111
IV. 3. 2.	RMN	112
IV. 3. 2. a.	Attribution des protons du peptide IB7 ₁₄ par RMN	112
IV. 3. 2. b.	Essais de l'attribution des protons de la protéine IB7	117
IV. 3. 2. c.	Etablissement d'une liste de contraintes pour le peptide IB7 ₁₄	118
IV. 3. 3.	Minimisation et Dynamique Moléculaire sous contraintes RMN : Structure tridimensionnelle du peptide IB7 ₁₄	119
IV. 4.	CONCLUSIONS ET PERSPECTIVES	122

CHAPITRE V INTERACTIONS TANINS-PROTEINES..... 125

V. 1.	INTRODUCTION	127
V. 1. 1.	Nature de l'interaction	127
V. 1. 1. a.	Nature et non-spécificité des sites de liaison	127
V. 1. 1. b.	Types de liaison tanin-protéine	128
V. 1. 1. c.	Affinité : aspects thermodynamique des complexes	128
V. 1. 1. d.	Paramètres contrôlant le complexe tanin-protéine	129
V. 1. 2.	Stratégie de l'étude	131
V. 2.	MATERIELS ET METHODES	131
V. 2. 1.	L'ESI couplée à un détecteur à trappe ionique	131
V. 2. 2.	Dichroïsme Circulaire	132
V. 2. 3.	Dynamique Moléculaire du complexe IB7 ₁₄ -B3	132
V. 2. 3. a.	Dynamique Moléculaire	132
V. 2. 3. b.	Calcul de surface : amphiphilicité	133
V. 2. 3. c.	Détermination du coefficient de diffusion par dynamique moléculaire et taille du complexe formé	133
V. 2. 4.	RMN	134
V. 2. 4. a.	Variation du déplacement chimique du peptide IB7 ₁₄ et de la protéine IB7	134
V. 2. 4. b.	Variation du coefficient de diffusion du peptide IB7 ₁₄ et de la protéine IB7	134
V. 2. 4. c.	Mesures des constantes d'association et de dissociation	135
V. 3.	RESULTATS	137
V. 3. 1.	Mise en évidence d'un complexe IB7 ₁₄ -nB3 par spectrométrie de masse	137

V. 3. 2. Dichroïsme Circulaire : Influence de B3 sur la structure secondaire du peptide IB7 ₁₄	140
V. 3. 3. RMN	142
V. 3. 3. a. Informations structurales	142
V. 3. 3. b. Association IB7 ₁₄ -B3	143
V. 3. 3. c. Association IB7-B3	149
V. 3. 4. Dynamique Moléculaire	153
V. 3. 4. a. Dynamique moléculaire sous contraintes RMN	153
V. 3. 4. b. Dynamique moléculaire sans contraintes RMN	155
V. 3. 4. c. Surfaces de lipophilie du peptide IB7 ₁₄	159
V. 3. 4. d. Détermination du coefficient de diffusion du peptide IB7 ₁₄ et du complexe IB7 ₁₄ /B3 par Dynamique Moléculaire (DM)	160
V. 4. DISCUSSION SUR LA NATURE DU SITE ET L'AFFINITE	161
V. 4. 1. La stœchiométrie et l'affinité du complexe	161
V. 4. 2. Sites de fixation des B3 sur le peptide	163
V. 4. 3. Structure de IB7 ₁₄ et IB7 en présence de tanins B3	164
V. 4. 4. Taille des complexes IB7 ₁₄ /3B3 et IB7/ <i>n</i> B3	165
V. 4. 5. Nature de l'interaction	167
CHAPITRE VI CONCLUSIONS ET PERSPECTIVES	169
VI. 1. CONCLUSION	171
VI. 2. PERSPECTIVES	172
CHAPITRE VII ANNEXES	175
ANNEXE 1 : ABBAQUES DES DEPLACEMENTS CHIMIQUES EN FONCTION DES ACIDES AMINES EN SOLUTION	177
ANNEXES 2 : EFFETS NOES ET CONSTANTES DE COUPLAGE ³ J RAPPORTEES AUX ANGLES DIEDRES	177
ANNEXES 3 : AUTO-ASSOCIATION DU TANIN B3	179
ANNEXES 4 : EFFETS NOES INTERMOLECULAIRES IB7 ₁₄ /B3	180
ANNEXES 5 : VARIATION DU DEPLACEMENT CHIMIQUE DES ¹ H PEPTIDIQUES LORS DE L'ADDITION DE PROCYANIDINE B3	180

ANNEXES 6 : COURBE DE DIFFUSION DU PEPTIDE IB7₁₄ PAR DM, REPRESENTANT LE CARRE DU DEPLACEMENT MOYEN DE LA MOLECULE EN FONCTION DU TEMPS. LA FIGURE DE GAUCHE REPRESENTE LE PEPTIDE COMPLEXE AUX TANINS, CELLE DE DROITE, LE PEPTIDE SEUL..... 187

BIBLIOGRAPHIE 189

PUBLICATIONS..... 199

REMERCIEMENTS

Ce travail de thèse a été réalisé à l'Institut Européen de Chimie et Biologie à Pessac. Je tiens à remercier messieurs Jean-Jacques Toulmé, Léon Ghosez et Jean-Marie Schmitter, co-directeurs de l'institut et directeur de l'UMR pour m'avoir accueilli dans ce laboratoire et pour m'avoir permis de côtoyer l'Excellence IECBienne.

Je remercie le Conseil Interprofessionnel du Vin de Bordeaux pour avoir financé ce projet.

Un directeur de thèse, c'est bien. Mais deux, c'est mieux ! Je remercie mes co-directeurs de thèse : le grand maître Jedi Erick-Joël Dufourc et Zaza, de son petit nom Isabelle Pianet de Boissel Granger. Entre autre de votre passation de connaissances, je rappellerai les grandes aventures Parisiennes, Marocaines ou Baltziennes ! Tant vos connaissances que votre humanité m'ont touchées.

L'IECB ne serait rien sans ce métissage scientifique : je remercie tous mes collaborateurs grâce auxquels un joli travail a été accompli : Mikael The War... enfin Michel Laguerre ([®] Boubou) faisant valser les protéines et dont la disponibilité est légendaire en toute modestie... Allez arrête de râler Michel ! JaMeS, Jean-Marie Schmitter, notre massiste international, dont le temps est bien compté, heureusement qu'il y a l'agenda ! Mais où est-il ? Quelque part sur le bureau, bien sûr ! Mais où est le bureau ? Quelque part sous les dossiers, bien sûr ! Qui dit temps compté dit KB ! Mais non pas la copine à Baygon, notre Karine Barathieu, la pro de la synthèse de polyphénols. Katell, presque toujours porteuse de bonnes nouvelles sur les spectres de masse ! Les membres de l'unité de Biophysique Structurale Philippe Picard et Claude Manigan, je vous remercie pour le temps et les connaissances que vous m'avez apportés.

Je n'oublierai pas l'équipe administrative sans qui tout aurait été bien compliqué : Joseph Wagner, Véronique Chassaing, Sandra et Annie.

Ah, le staff RMN ! Merci Erick pour tes castings de choix : Axounette, notre ingénieuse RMN Axelle Grélard et notre aîné Fabinou lapinou chouchou, dit Fabien Dussenac ou Aussenac perverti par la RMN, que je remercie tous deux d'être simplement eux-mêmes. Je remercie Lulu, autrement dit Lucie Khemtémourian (j'espère ne pas avoir fait de fautes, mais c'est quand-même plus simple de s'appeler Simon !) pour avoir diversifiée le pôle RMN

en un élevage de hamsters ! Je remercie nos post-doc Kiki, Christina Sizun, qui n'est à mon goût pas restée assez longtemps et Régis Nguyen sans qui la synthèse aurait été un calvaire. Je remercie bien sûr tous les stagiaires qui ont permis de rajeunir le bureau : Sylvie Fournier, Cécile Lalanne, Catherine Algisi, Anthonny Couvreur et Marie Garnier.

Puis je remercie l'ensemble de la communauté IECBienne pour faire de ce labo un endroit où il fait bon vivre, dans un ordre tout à fait arbitraire (et il ne faut surtout pas que j'en oublie, comme certains !) : mes compatriotes Boubou dit Loanno dit la grande blonde, Alexis, Victor et son gros rotor ; mes collègues de café-apéro qui m'ont aussi torturé au squash : mon jumeau Jo « Grrrrr », notre hotline internationale Juju, encore un grand méchant Carlos, Thomas le penseur ; Marie, Aurélie, Chrystel, Arnaud & Arnaud (ça devrait rattrapper ?), Laurence, Céline, Stéphane ; nos étrangers : Jez, Peter, Jiang, Conor voire Elise, Jereon ; les p'tits jeunes : Thomas D., Arnaud A., Loïc ; les autres chefs Fred, Jean, Reiko, Ivan; puis bien sûr les dindes, pintades et quiches, hamsters, poissons et le reste. Les anciens que l'on oublie pas bien sûr !

Enfin la mascotte : Loulou, Pipoune, Titi dit P'Yukan Gumo Off Shinsen En.

LISTE DES ABREVIATIONS

Bn : Benzyl

Boc : *tert*-Butoxycarbonyl

CID : Collision induced dissociation

CH₂Cl₂ : Dichlorométhane

COSY : Correlation spectroscopy

COSY-DQF : Correlation spectroscopy-double quantum filtered

C-ter : extension Carboxy-terminale

Da : Dalton

DC : Dichroïsme circulaire

DCC : Dicyclocarbodiimide

DCM : Dichlorométhane

$\Delta\delta$: Variation de déplacement chimique

DDQ : 2,3-dichloro-5,6-dicyano-1,4-benzophenone

DM : Dynamique moléculaire

DMF : N,N-diméthylformamide

DMSO : Diméthylsulfoxyde

DIEA : Diisopropylethylamine

ESI : Electrospray ionisation

FID : Free Induction Decay

J : Constante de couplage (en Hertz)

Fmoc : 9-Fluorenylmethoxycarbonyl

Gln : Glutamine (Q)

Gly : Glycine (G)

HBTU : 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetraméthyluronium hexafluorophosphate

HMBC : Heteronuclear multiple bond connectivity

HMP : 4-Hydroxyméthylphenoxy-méthyle

HOBt : 1-Hydroxybenzotriazole
HRMAS : High Resolution Magic Angle Spinning
Lys : Lysine (K)
MALDI-TOF : Matrix assisted laser desorption/ionisation – Time of flight
mn: minute
nm: nanomètre
NMP : N-méthylpyrrolidone
NOE : Nuclear Overhauser enhancement
NOESY : Nuclear Overhauser enhancement spectroscopy
N-ter : extension amino-terminale
 R_f : (distance parcourue par le soluté)/(distance parcourue par le solvant)
RMN : Résonance Magnétique Nucléaire
RMSD : Root mean squared deviation
ROESY : Rotating frame Overhauser enhancement spectroscopy
RP-HPLC : Chromatographie liquide à haute performance en phase inverse
PEG : Polyéthylèneglycol
Pro : Proline (P)
SM : Spectrométrie de masse
ta : température ambiante
TFA : Acide trifluoroacétique
TFE : Trifluoroethanol
THF : Tétrahydrofurane
TSP : Acide 3-trimethylsilyl-propionique-2,2,3,3-d4
TOCSY : Total correlation spectroscopy
UV : Ultra violet

Les tanins sont impliqués dans le pouvoir gustatif d'un vin de par leur possibilité d'interagir avec certaines protéines. Cette interaction induit la précipitation du complexe tanin-protéine, propriété exploitée lors du collage des vins rouges de Bordeaux. De même, l'interaction des tanins avec des protéines de la salive a pour conséquence de provoquer la sensation d'astringence en bouche. Le but de cette thèse est de déterminer la nature de cette interaction d'un point de vue chimique et physique. Cet objectif nécessite trois étapes clés : la synthèse des deux protagonistes de l'astringence, l'étude structurale des protéines spécifiques, et l'interaction tanin-protéine en milieu aqueux ou hydroalcoolique. Ces divers aspects devraient conduire à une meilleure compréhension de la typicité des vins d'une manière générale ainsi qu'à l'étude d'autres protéines susceptibles d'interagir avec les tanins durant l'étape de clarification des vins, les protéines du collage.

Chapitre I INTRODUCTION :
INTERACTION TANIN-PROTEINE : LA CLE
DU « GOUT » ?

Introduction :

Interaction Tanin-Protéine : la clé du « goût » ?

Lors de la dégustation d'un vin rouge diverses caractéristiques sont prises en compte : comme la robe, la couleur, la texture et la souplesse, mais au-delà de ces aspects visuels, les sensations tactiles et gustatives en bouche apparaissent comme primordiales afin de définir le caractère propre d'un vin. La reconnaissance, au niveau buccal, d'un vin rouge est ainsi fortement basée sur son potentiel « tannique ». La communauté des œnologues s'attache donc à relier la qualité d'« un tanin » à ses propriétés d'auto-agrégation et d'association avec les protéines, les polysaccharides et les alcaloïdes¹⁻³. Dans le vin, ces propriétés sont en relation avec les qualités appelées corps, structure et souplesse d'un vin. Les tanins s'agrègent avec les protéines, donc ont un rôle dans la stabilité colloïdale et le goût. En effet, ils sont capables de précipiter les protéines originaires du raisin ou de la levure, mais aussi les protéines rajoutées lors du collage telles que l'albumine ou la gélatine. La sensation de sécheresse en bouche lors de la dégustation d'un vin rouge, serait liée, quant à elle, à l'interaction entre les polyphénols issus du vin avec les protéines de la salive⁴ : cette sensation est appelée l'astringence. Aristote (384-322 av. JC) fut le premier à ouvrir le débat concernant l'astringence perçue comme une sensation tactile ou gustative. Aujourd'hui, l'ensemble de la communauté scientifique associe l'astringence à une sensation tactile⁵⁻⁷ tant d'un point de vue physique⁸ que perceptif^{9,10}.

Ce phénomène d'astringence serait la somme de plusieurs effets tels que la sécheresse (perte de lubrification), la rugosité (différence de texture dans la cavité orale) et la rétraction des muqueuses¹¹⁻¹³. Ces notions sont tout à fait subjectives étant donné que l'astringence est usuellement définie par des tests de dégustation, or il est évident que toute étude sensorielle dépend de la perception de chacun. Les individus possèdent des taux de sécrétions salivaires plus ou moins importants et donc ont une perception différente¹⁴.

L'astringence est liée à l'interaction entre certains tanins et les protéines de la salive^{4,15,16}. La salive est un fluide aqueux contenant protéines, glycoprotéines, glycolipides, polysaccharides et ions organiques¹⁷. L'une des principales fonctions de ces composés est de lubrifier la surface muqueuse¹⁸. 70% des protéines sécrétées par les glandes parotide et submandibulaire appartiennent à la famille des PRPs^{19,20}, constituées majoritairement par 3 résidus (proline, glycine et glutamine jusqu'à 88%). Les PRPs sont les protéines salivaires ayant le plus d'affinité pour les tanins, en raison de leurs multiples séquences riches en proline¹. L'astringence serait principalement due à l'interaction entre les tanins et les PRPs^{7,21} car celles-ci constituent un lubrifiant buccal. D'ailleurs, Kallithraka *et al.*²² observent un changement de la distribution des protéines salivaires après dégustation d'un vin rouge.

L'acteur causal de l'astringence est donc le Tanin, appartenant à la grande famille des polyphénols. Ainsi, la notion d'astringence est reliée à la notion de tanin qui a été utilisée pour la première fois par Seguin (1796) pour décrire le matériel dans les barriques de chêne. Les flavan-3-ols mono et polymériques sont les principaux constituants phénoliques du vin²³. Ils sont issus des parties dures des grappes de raisin (pépins et pellicule) durant la fermentation²⁴ mais pas de la pulpe²⁵.

Dans ce qui suit nous allons donc rappeler ce qui est connu sur les deux partenaires de l'astringence, les protéines de la salive et les tanins.

I. 1. LES PROTEINES DE LA SALIVE

D'après les dernières informations des banques de données, recherche effectuée jusqu'en 2002, plus de 50 protéines et peptides, distinctes par leur structure et leur rôle, ont été mis en évidence dans la salive humaine. Deux types de protéines constituent les acteurs forts de la lubrification salivaire, les mucines et les protéines riches en proline (PRPs). Les mucines constituent 26% des protéines totales de la salive, principalement sécrétées par la glande sublinguale²⁶. La grande majorité des protéines salivaires appartient à la super famille des Protéines Riche en Proline, appelées PRPs. Les PRPs ont la capacité de précipiter les tanins. D'autres protéines ont une certaine affinité pour les tanins comme les histatines - dont 12 ont été identifiées²⁷ qui sont de petits peptides cationiques et antifongiques riches en résidus histidines ayant des rôles de défense contre des agents pathogènes salivaires - ou encore les statherines, cystatines, mucines et α -amylases qui ont un rôle contre les microorganismes oraux²⁸. La répartition de ces protéines ainsi que leur nature dépend du donneur (sexe, age, alimentation...).

I. 1. 1. Les Protéines Riches en Proline (PRP)

Les Protéines riches en Proline (PRPs) constituent environ 2/3 des protéines totales sécrétées par la glande parotide et par le conduit submandibulaire^{19,20}. L'expression de ces protéines suit un schéma constitutif, c'est à dire qu'elles ne sont produites en masse qu'en présence d'agents activateurs, ce qui explique leur relative faible abondance dans la cavité

orale, en terme de concentration. Malgré cela de hauts niveaux sont présents chez les primates, contrairement à d'autres espèces tels que les chevaux, chiens et chats qui ne les synthétisent pas. Dans la grande famille des muridés, les PRPs ne sont présentes de manière naturelle qu'en infime quantité, mais la présence d'agents activateurs tels qu'un type de catécholamine ou les tanins, induit une forte augmentation de leur concentration.

I. 1. 1. a. La notion de multifamille

A ce jour, plus d'une vingtaine de PRPs ont été identifiées dans la salive humaine. Elles sont classées en trois groupes selon leur nature chimique et leurs charges caractéristiques qui leur confèrent des propriétés basiques ou acides, plus ou moins phosphorylées ou encore en fonction de leur aptitude à fixer covalamment un groupement carbohydrate, on les appelle dans ce cas les glycosylées²⁹. Leur appartenance à une multifamille vient d'un précurseur nucléotidique commun. En effet toutes ces protéines ne sont codées que par seulement 6 gènes dotés d'un fort polymorphisme³⁰. Deux gènes codent pour les acides, les autres pour les basiques et les glycosylées³¹. Pour chaque gène, un précurseur est traduit puis il est clivé par des convertases par protéolyse post-transcriptionnelle ce qui explique le haut degré de similarité entre les différentes protéines³².

I. 1. 1. b. Rôles des PRPs

Tous leurs rôles ne sont pas encore totalement identifiés. Cependant, comme toutes protéines, les rôles des PRPs au sein de l'organisme sont dictés par leurs groupements fonctionnels. Elles sont impliquées dans de nombreuses réactions enzymatiques et dans les processus de reconnaissance biologique chez l'homme. Elles sont aussi connues pour avoir de multiples fonctions de défense au sein l'organisme.

Les PRPs acides possèdent des sites de liaison spécifiques de l'ion calcium (avec un nombre de sites variable en fonction des protéines³³) et pourraient agir comme un réservoir à calcium permettant le maintien de l'homéostasie calcique dans la bouche²⁹. Différentes actions vont en ce sens : elles lient le calcium et inhibent la croissance des cristaux d'hydroxyapatite à la surface des dents³⁴⁻³⁶, phénomènes reliés au processus de minéralisation dentaire et jouent un rôle dans la stabilisation des sels de phosphate de calcium dans la salive³⁷. Les protéines basiques glycosylées agissent comme des lubrifiants³⁸ et fixent les bactéries buccales³⁹ de sorte qu'une augmentation du taux de PRPs basiques entraîne une

diminution de la formation de caries⁴⁰. Plus récemment, leur rôle antiviral a été démontré, par liaison aux virus et plus particulièrement à une protéine du manteau du HIV-1^{41,42}. En ce qui concerne les PRPs basiques, aucun rôle spécifique n'a encore été identifié. En revanche, ces protéines basiques ont une très forte capacité à fixer les polyphénols. En effet, toutes les PRPs possèdent cette propriété mais les basiques de part leur structure fibrillaire sont davantage aptes à neutraliser les effets néfastes des tanins alimentaires, par complexation et précipitation⁴³⁻⁴⁵. Cette propriété constitue un rôle protecteur contre le potentiel toxique et carcinogénique des tanins alimentaires^{34,46,47}. En effet, les PRPs liant les tanins préviendraient du cancer de l'œsophage⁴⁸.

I. 1. 2. Les PRP basiques

Les PRPs basiques représentent ~23% des protéines totale de la salive¹⁹, sécrétées majoritairement par la glande parotide⁴⁹ mais aussi par les glandes Ebner⁵⁰. La majorité des études faites sur les PRPs utilisent un donneur humain unique mais très peu de différences en composition ont été mises en évidence d'un individu à l'autre par comparaison entre les différentes études.

I. 1. 2. a. Diversités des PRPs basiques

L'ensemble des PRPs présente une forte homologie de séquence du fait du peu de nombre de gènes codant pour ces protéines. En effet six gènes seulement codent pour les PRPs basiques, acides et glycosylées. De plus, le nombre important des PRPs (pas encore identifié à ce jour) provient de dégradations protéolytiques post traductionnelles des produits du gène. La présence d'une enzyme kallikreine-like dans la salive parotide humaine, enzyme de dégradation protéolytique spécifique de ces protéines⁵¹, pourrait expliquer cette multiplicité.

Elles peuvent être séparées en deux groupes polymorphiques distincts selon leur poids moléculaire dont Ps1 et Ps2 pour le groupe de haut poids moléculaire et de IB1 à IB9 pour le second groupe de faible poids moléculaire avec peu ou pas de carbohydrates.

I. 1. 2. b. Structure primaire des PRP basiques

L'ensemble de ces protéines est principalement constitué par trois résidus : la proline ($40\% \pm 2$), la glycine ($21\% \pm 2$) et la glutamine ($17\% \pm 3$). Ces trois acides aminés représentent donc 70 à 85% des résidus totaux dans les protéines. Certaines micro-séquences sont inscrites en tandem dans la chaîne peptidique tel qu'un nonapeptide PPGKPQGPP (représenté en rouge dans la figure I-1) ainsi qu'un pentapeptide PQGPP (souligné dans la figure I-1). La figure I-1 (a) offre une vision claire de la superposition en séquence des protéines basiques dites « classiques ». Entre IB7, IB8c et IB9 l'identité compte pour 95% des résidus. Saitoh *et al.*^{53,54} montrent également une forte homologie entre IB5 et IB4 pour plus de 70%, de même Shibata *et al.*⁵⁵ remarquent des ressemblances frappantes entre IB9, IB8c et IB5 (72% entre IB9 et IB5).

IB1 et II-2 sont des phosphoprotéines basiques qui contiennent 35 résidus supplémentaires en partie amino-terminale (*N*-ter). Leur phosphorylation potentielle sur la serine 8 leur procure des propriétés moins basiques que les « classiques ». Cependant la protéine IB7 s'inscrit parfaitement dans la partie carboxy-terminale (*C*-ter) de ces deux protéines (figure I-1 (b)). IB8a et IB6 sont quant à elles potentiellement glycosylables, propriété émanant des résidus en partie *C*-ter. En effet, Kauffman *et al.* démontrent que IB9 constitue pour partie IB6 à laquelle est rajoutée en partie *C*-ter une extension d'une trentaine de résidus³⁰. Malgré cette différence fonctionnelle, l'alignement de séquences avec IB7 sur la partie *N*-ter est remarquable (figure I-1 (c)).

Il est intéressant de noter qu'une grande homologie de séquence existe aussi entre les protéines acides et basiques.

I. 1. 3. Choix d'une protéine et d'un peptide modèles

Les PRPs sont les protéines les plus représentées en nombre dans la salive humaine. Au sein de cette multifamille, les protéines basiques ont le plus d'affinité pour les polyphénols. Parmi elles, une protéine voit sa séquence reprise dans la majorité des autres, la protéine IB7⁵², une des plus petites des PRPs. Elle possède 59 résidus répartis principalement en 3 acides aminés : 25 prolines (42%), 13 glycines (22%) et 9 glutamines (15%) composant ainsi la chaîne peptidique pour 79%. Elle est superposable par trois reprises en segments de 12 résidus : PPGKPQGPPPQG comme le montre la figure I-2. Certaines microséquences sont présentes de nombreuses fois comme le pentapeptide PQGPP répété 5 fois.

SPPGKPQGPPPQGGNQPQGPP
PPGKPQGPPPQGGNKPQGPP
PPGKPQGPPPQGDNKSR ⁵⁹

Figure I-2 : Séquence primaire de la protéine IB7 de la salive humaine composée de 59 résidus. En rouge est représenté un motif répété trois fois et un second motif très ubiquitaire est souligné.

La protéine IB7 apparaît comme un bon modèle représentatif des PRPs. Nous nous sommes également intéressés à un motif répété à trois reprises dans la protéine IB7 et très ubiquitaire dans l'ensemble de la population des PRPs, la partie *N*-ter de la protéine IB7 appelée IB7₁₄, car constituée des 14 premiers résidus de la chaîne peptidique ($N_{\text{ter}}\text{SPPGKPQGPPPQG}^{14}\dots$).

I. 2. LES TANINS

Les composés phénoliques ont un rôle primordial en œnologie, notamment ils sont responsables des différences de couleur et de saveur entre vins blancs et rouges. Ils interviennent dans la matière colorante du vin de part les anthocyanes libres ou combinées aux tanins. Les tanins appartiennent à cette grande famille. Ils sont indispensables à la réalisation d'un vin de qualité, ils participent à des propriétés organoleptiques telles que le « corps », la « charpente » ou encore le « volume » du vin⁵⁶. La saveur du vin résulte en partie des propriétés physico-chimiques des tanins et de leurs caractéristiques structurales. Ces propriétés étant nécessaires à la compréhension du comportement organoleptique des tanins.

Une première définition physico-chimique des tanins est donnée par Bate-Smith et Swain^{4,57} : « composés phénoliques solubles dans l'eau, possédant une masse moléculaire comprise entre 500 et 3000 Da, qui en plus de la réactivité usuelle des composés phénoliques, ont la capacité de faire précipiter les alcaloïdes, la gélatine et les autres protéines ». Il y a alors formation de colloïdes dont la dimension est comprise entre 10 nm et 1 µm environ.

Ces métabolites secondaires sont localisés dans les feuilles, l'écorce et les fruits de nombreuses plantes. Ils font ainsi partie intégrante de notre alimentation (vin, thé, divers fruits...).

I. 2. 1. Les tanins du vin et du raisin

Les tanins du vin et du raisin sont responsables des propriétés dites « positives » telles que le corps d'un vin, mais également de propriétés « négatives » comme la rusticité ou l'agressivité. Ils sont par nature astringents, c'est-à-dire capables de donner des combinaisons stables avec les protéines.

Les tanins appartiennent à la classe abondante des polyphénols caractérisés par leurs multiples groupements phénoliques. Les tanins sont séparés en deux grandes catégories selon leur composition chimique : les tanins condensés qui sont des procyanidines, ou oligomères de flavan-3-ol principalement composés d'unités monomériques de type épicatechine en majorité et catéchine (**1**) ou des prodelphinidines composés d'unités monomériques de type (épi)gallocatéchine, et les tanins hydrolysables, esters de l'acide gallique et ses dérivés souvent estérifiés avec le glucose dont le pentagalloyl glucose est le précurseur (**2**)⁵⁸. Certaines espèces végétales synthétisent également des tanins mixtes constitués d'unités de type tanin hydrolysable et de type tanin condensé reliées par une liaison carbone-carbone.

La présence de ces deux types de tanins dans le vin a des origines différentes. En effet, les tanins condensés proviennent directement des parties solides du raisin, pépins⁵⁹ et pellicule^{60,61}, alors qu'en quantité moindre, les tanins hydrolysables sont issus du bois et cédés au vin lors de la macération dès lors que les vins sont élevés en fût de chêne.

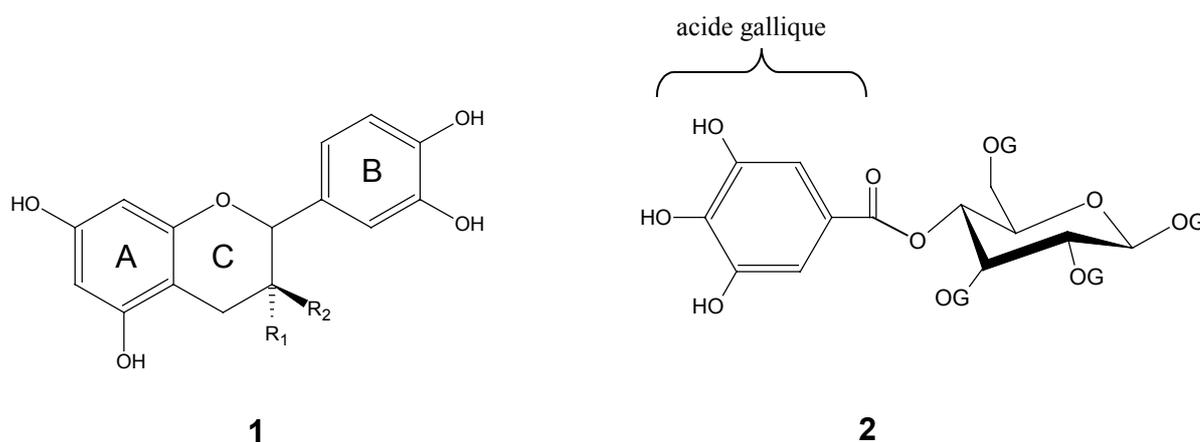


Figure I-3 : Formules chimiques développées du monomère flavan-3-ol, **1**, et du penta galloyl glucose, le plus représenté des gallotanins, **2**. $R_1 = \text{OH}$, $R_2 = \text{H}$: épicatechine ; $R_1 = \text{H}$, $R_2 = \text{OH}$: catéchine ; A : cycle phloroglucinol ; B : cycle catéchol ; C : cycle pyrane ; G = acide gallique.

I. 2. 2. Les tanins condensés

Les tanins les plus rencontrés dans le vin appartiennent à la famille des tanins condensés. Leur teneur globale dans les vins rouges a été estimée entre 1 et 4 g.L⁻¹, dépendant du cépage et des conditions climatiques et de la vinification. Les tanins condensés font partie de la famille des flavonoïdes. Ils possèdent comme unité structurale fondamentale le noyau flavane. Ils sont également appelés proanthocyanidines, nom provenant de leur capacité à libérer des anthocyanidines de couleur rouge par rupture de la liaison interflavane en milieu acide et en conditions d'oxydation. Les tanins condensés ont une masse comprise entre 500 et 20000 Da en fonction de leur degré de polymérisation et de différents tissus de plantes^{11,62}

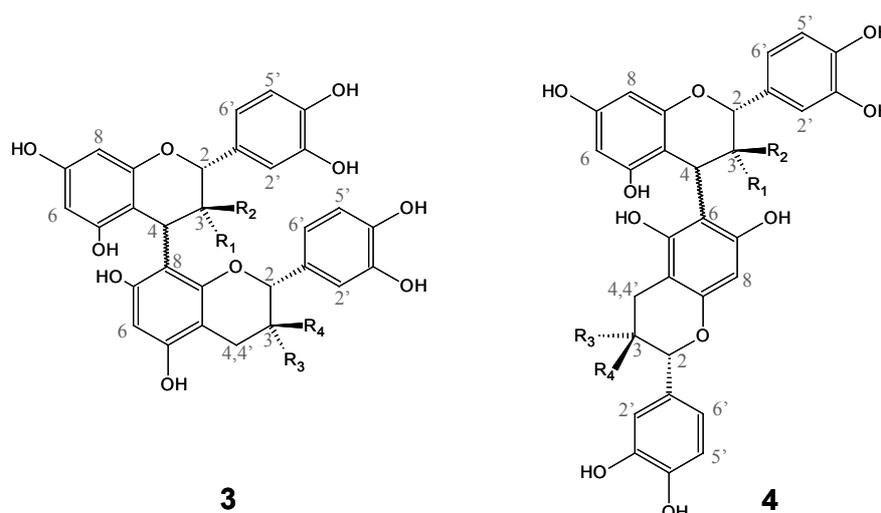
I. 2. 2. a. Nature des proanthocyanidines et degré de polymérisation

La grande diversité structurale de ces flavan-3-ols résulte du nombre de groupes hydroxyles, de leurs positions sur les noyaux aromatiques A et B, de la stéréochimie des carbones asymétriques du cycle pyrane C ainsi que du nombre et du type de liaisons dites « interflavaniques » entre les unités monomériques. Nous ne parlerons ici que des dérivés de

flavan-3-ols ayant la catéchine ou l'épicatéchine comme unités de construction, mais certains peuvent être estérifiés notamment par l'acide gallique.

Les propriétés des proanthocyanidines dépendent de leur structure et particulièrement de leur degré de polymérisation. Les tanins ont ainsi la propriété d'auto-association jusqu'à former des complexes plus ou moins importants. Il existe deux classes de dimères de flavan-3-ol en fonction de la nature de la liaison reliant les unités monomériques. Les plus abondants appartiennent à la série B (figure I-4), pour lesquels la liaison se fait par des liaisons C₄-C₈ (**3**) majoritaires ou C₄-C₆ (**4**). Il existe aussi des polymères de type A (très peu représenté dans le vin et le raisin), ayant une liaison éther entre les carbones C₅ ou C₇ de l'unité terminale et le carbone C₂ de l'unité supérieure en plus de la liaison interflavane existant dans la série B.

Les seuls trimères trouvés dans le vin possèdent deux liaisons de type B. Il existe également des polymères dans le vin, provenant des raisins dont le degré moyen de polymérisation avoisine 10 dans les pépins de raisin et peut aller jusqu'à 30 dans la pellicule.



	~~~~~		~~~~~
B1 : R1=OH ; R2=H ; R3=H ; R4=OH	▲	B5 : R1=OH ; R2=H ; R3=OH ; R4=H	▲
B2 : R1=OH ; R2=H ; R3=OH ; R4=H	▲	B6 : R1=H ; R2=OH ; R3=H ; R4=OH	
B3 : R1=H ; R2=OH ; R3=H ; R4=OH		B7 : R1=OH ; R2=H ; R3=H ; R4=OH	▲
B4 : R1=H ; R2=OH ; R3=OH ; R4=H		B8 : R1=H ; R2=OH ; R3=OH ; R4=H	

### Nomenclature

B1 : épicatechine-(4 $\beta$ -8)-catéchine  
 B2 : épicatechine-(4 $\beta$ -8)-épicatechine  
 B3 : catéchine-(4 $\beta$ -8)-catéchine  
 B4 : catéchine-(4 $\beta$ -8)-épicatechine

B5 : épicatechine-(4 $\beta$ -6)-épicatechine  
 B6 : catéchine-(4 $\beta$ -6)-catéchine  
 B7 : épicatechine-(4 $\beta$ -6)-catéchine  
 B8 : catéchine-(4 $\beta$ -6)-épicatechine

Figure I-4 : Structure et nomenclature des procyanidines dimères de flavan-3-ols de la série B : dimères liés en 4 – 8, **3**, et en 4 – 6, **4**.

## I. 2. 3. Choix du tanin modèle

Les tanins condensés sont les plus abondants dans le vin et ils proviennent directement des raisins. Parmi eux, les oligomères sont les plus représentés en nombre. Etant donné qu'ils sont quantitativement les plus importants, les dimères et trimères pourront représenter la population de tanins issus des pépins du raisin. La synthèse par voie chimique de ces composés a été optimisée pour le dimère B3 par Karine Barathieu⁶³ (au Laboratoire de Chimie Organique et Organométallique), dont nous pouvons avoir des quantités de produits

intéressantes. Le dimère B3 (catéchine-(4 $\beta$ -8)-catéchine) (figure I-5) a donc été choisi comme modèle de tanin dans le raisin et le vin.

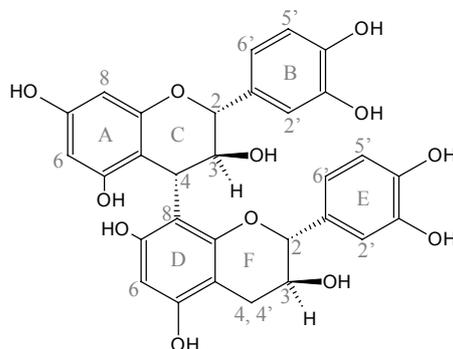


Figure I-5 : Structure semi-développée du dimère B3 (catéchine-(4 $\beta$ -8)-catéchine).

## I. 3. L'INTERACTION TANINS-PROTEINES

### I. 3. 1. Généralités

Les tanins sont des polyphénols qui agissent comme barrière protectrice de certaines plantes vis-à-vis de la prédation animale^{1,16,64,65} : en effet, Bate-Smith explique que, d'un point de vue biologique, l'importance des tanins dans les plantes provient de leur efficacité à écœurer les prédateurs, animaux ou microbes. La propriété commune à tous les tanins est d'être astringents en précipitant certaines protéines de la salive. La communauté scientifique leur confère également des propriétés médicinales à la suite de leur ingestion par l'homme, d'où la théorie du « French paradox ». En effet, leurs multiples groupements phénoliques auraient des effets antiradicalaire⁶⁶, antioxydant⁶⁷, anticancérigène, antibactérien, vitaminique... et préviendraient donc de maladies cardiovasculaires^{58,68}, de certains cancers, d'infections bactériennes... Les effets néfastes des tanins sur la santé sont souvent oubliés, mais ils perturberaient l'absorption minérale⁶⁹, provoquent l'inhibition de certaines enzymes digestives⁷⁰ et participent au développement de certains cancers (comme celui de l'œsophage)^{48,71,72}. Toutes ces propriétés dépendraient de leur structure même et surtout de leur degré de polymérisation.

D'un point de vue du prédateur, l'interaction tanins-protéines joue un rôle protecteur contre le potentiel toxique et carcinogénique des tanins alimentaires^{46,47}. A titre d'exemple, citons la comparaison de l'homme avec le hamster⁷³ : en effet, cet animal ne synthétise pas de PRPs, donc lorsqu'il ingère des polyphénols il voit sa croissance ralentir, contrairement à l'homme sur qui la présence de ce composé dans la nourriture n'a aucun effet phénotypique. Pourtant si l'homme synthétise en continu les PRPs, cette synthèse est fortement stimulée par la présence de polyphénols suggérant que l'induction de la synthèse des PRPs constitue une défense contre la présence des polyphénols dans l'alimentation^{74,75}.

### I. 3. 2. Principes de la complexation

Un modèle de complexation des tanins avec les protéines a été proposé par Haslam⁷⁶ puis raffiné par Saucier et Ribereau-Gayon^{77,78} dont le schéma est présenté figure 1-6 :

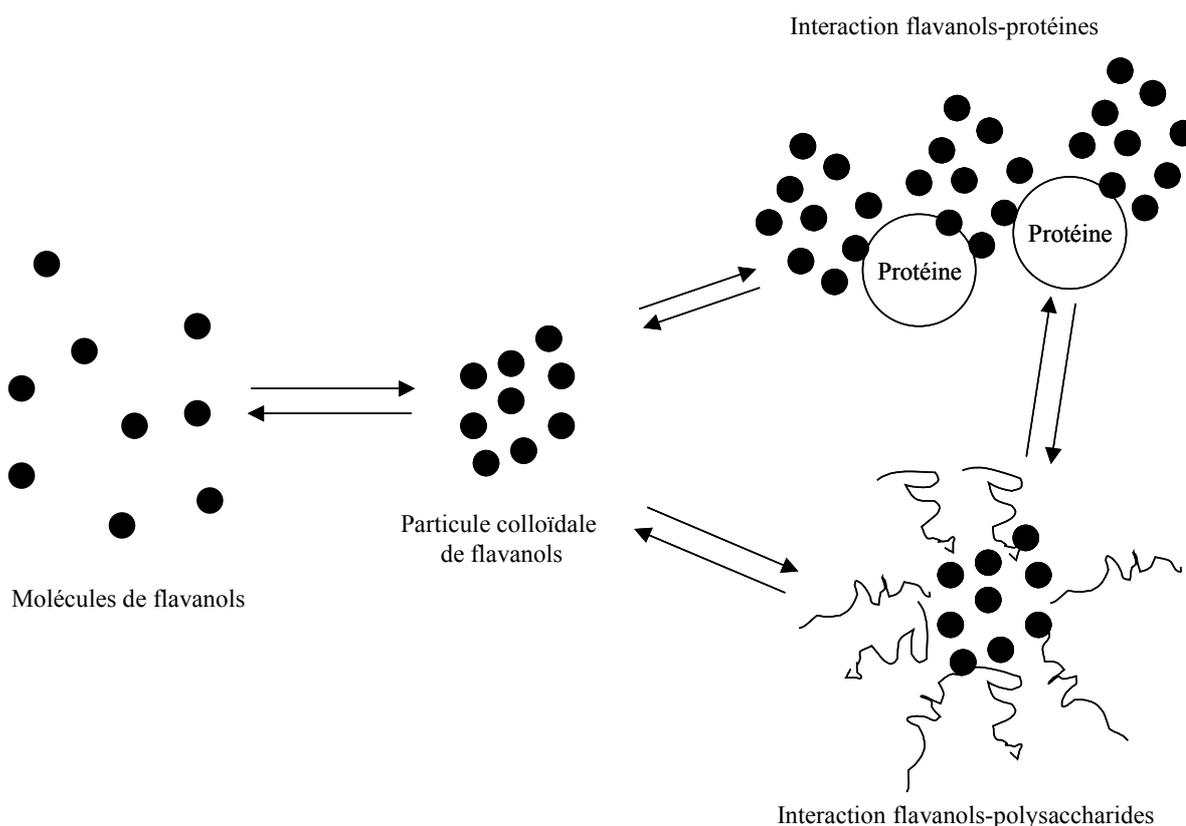


Figure I-6 : Modèle de précipitation des protéines par les polyphénols^{77,78}.

Ce processus de complexation s'effectuerait en deux étapes : en premier lieu, l'interaction est telle qu'un ou plusieurs tanins fixe(nt) la protéine formant une mono couche, diminuant ainsi leur caractère hydrophile. Ensuite, l'augmentation de la concentration en protéines permettrait aux polyphénols de jouer le rôle d'agents pontants entre les différentes molécules, grâce à leur propriété multidentate⁷⁹. La couche superficielle hydrophobe se reformerait alors et entraînerait la précipitation des protéines. Cette précipitation pouvant être réversible par un ajout de protéine⁸⁰, chacune n'étant apte à fixer qu'un nombre fixe de polyphénols⁸¹ ou bien par ajout de polysaccharides avec lesquels les tanins se complexent ce qui provoque une diminution de la liaison avec les protéines par effet compétitif⁸⁰. Cette complexation pouvant être favorisée par de nombreux facteurs (nature des protagonistes, concentration, solvant, pH, force ionique...) d'où l'intérêt de se rapprocher au mieux des conditions naturelles.

### I. 3. 3. Capacité d'interaction avec les polyphénols

Toutes les PRPs ont la capacité à complexer les polyphénols. La différence d'affinité des protéines pour les tanins tient à leur nature chimique. Les PRPs acides et glycosylées possèdent une extension *N*- ou *C*-ter par rapport aux basiques. Ces extensions leur confèrent leurs propriétés « acides » ou « glycosylables » et induisent des conformations globulaires (figure I-7-B), ne se prêtant pas à la fixation de tanins, contrairement aux PRPs basiques, fibrillaires (figure I-7-A). Les protéines acides ou glycosylées auxquelles sont clivées les séquences leur conférant cette propriété (partie *N*-ter et partie *C*-ter respectivement) retrouvent la même affinité pour les tanins à l'image des PRPs basiques⁴⁷.

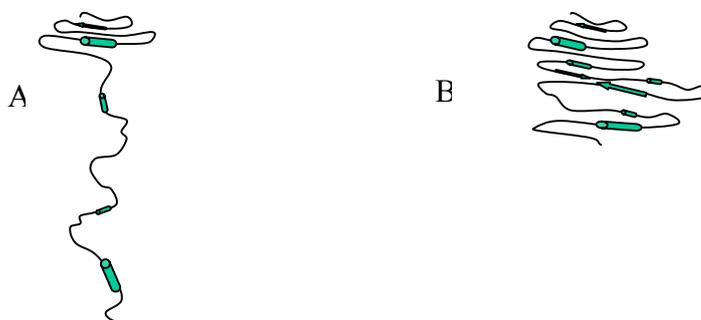


Figure I-7 : Deux assemblages possibles de protéines : en A : modèle de protéine fibrillaire, en B : modèle de protéine globulaire (d'après Cid *et al.*⁸²).

### I. 3. 4. Nature du site de liaison

La structure tridimensionnelle des protéines est primordiale dans l'interaction tanins-protéines. Elles doivent adopter une structure ouverte, plus exposée au solvant, pour faciliter l'approche des tanins. Il n'existe pas de séquences spécifiques d'interaction, seule leur particularité à posséder en tandem des résidus proline, glycine et glutamine suffirait à provoquer une forte affinité pour les polyphénols. Ces trois résidus suffisent à engendrer une conformation ouverte et étendue à la protéine. Les nombreux résidus prolines joueraient un rôle central dans cette interaction : la conformation du squelette proline, la restriction de la conformation du résidu précédent la proline et la nature de l'amine et du carbonyle font de ce résidu un site préférentiel pour l'interaction avec les polyphénols⁸⁰.

Les principaux mécanismes d'interaction entre les tanins et les protéines sont au nombre de trois, représentées sur la figure I-8 : les interactions hydrophobes (forces attractives de Van der Waals), les liaisons hydrogène et enfin les interactions de type ioniques absentes au pH utilisé, acide.

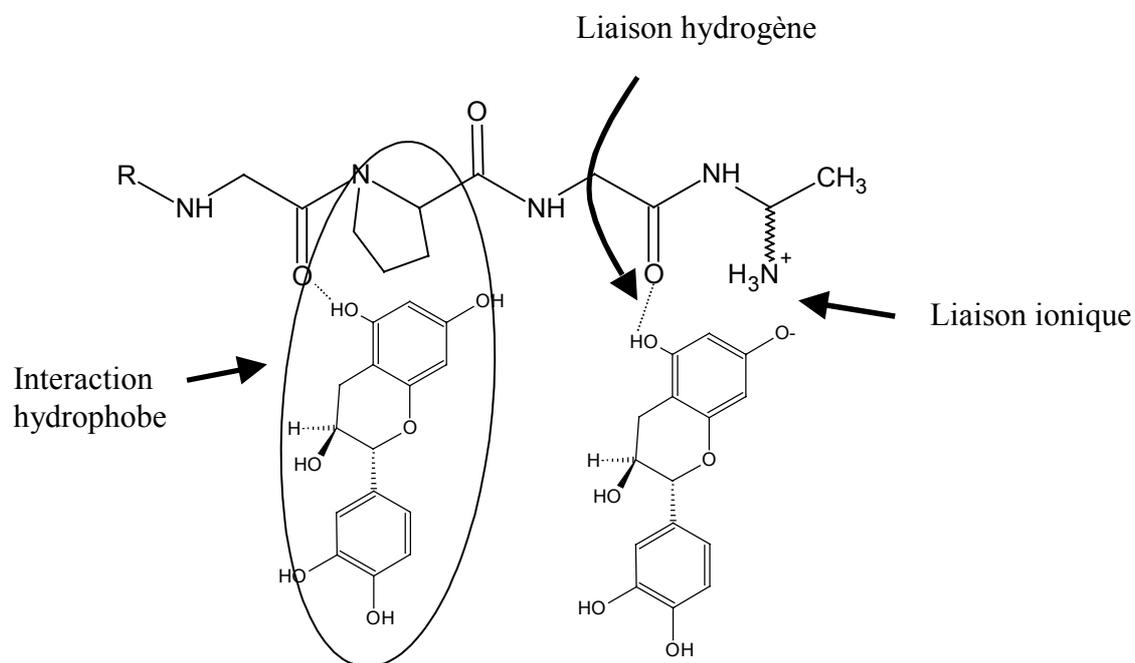


Figure I-8 : Interaction des protéines avec les polyphénols⁸³.

A ce jour, la force majeure d'interaction n'a pas pu être définie et liaisons hydrogène ou interactions hydrophobes se battent la vedette. Ces deux forces d'interaction auraient un rôle dans l'association polyphénols-protéines et l'ensemble de la littérature révèle le flou existant. Tous sont pourtant d'accord avec la nécessité de la présence de prolines dans la séquence peptidique : d'un côté car elle induit le caractère hydrophobe du squelette carboné et de l'autre car la fonction carbonyle des prolines des PRPs, liée à une amide tertiaire, est reconnue comme plus accepteur de liaisons hydrogène que les autres résidus possédant une amide secondaire⁸⁰. Ces liaisons hydrogène seraient fournies pas les fonctions hydroxyles des polyphénols^{1,84}. L'empilement hydrophobe du phénol sur la proline serait le mode d'interaction majeur³. Le cycle pyrrolidine de la proline offre un site ouvert et rigide d'interaction avec la partie aromatique du phénol⁸⁵. Le troisième type d'interaction, ionique, est probablement moins importante mais peut s'ajouter à l'ensemble des interactions précédentes.

## **I. 4. OBJECTIFS ET PLAN DE LA THESE**

L'astringence, vue comme une sensation tactile, est habituellement étudiée de manière tactile ou gustative. Ce travail concerne une première approche quant à l'étude de l'astringence d'un point de vue physico-chimique : l'adaptation conformationnelle du peptide IB7₁₄ au sein du complexe IB7₁₄-B3 ainsi qu'une étude approfondie concernant la nature de l'interaction pourrait nous amener à quantifier la notion d'astringence.

La compréhension de l'interaction tanin-protéine passe par leur étude structurale. En effet, malgré l'importance de la relation structure-activité des protéines, aucune étude structurale à l'aide de ces deux techniques n'a été entreprise sur de telles protéines à ce jour. L'analyse du complexe formé lors de la mise en présence des deux composés peut prendre différents chemins : en effet, l'adaptation de la conformation tridimensionnelle du peptide en présence de tanins, tout comme la nature des liaisons et la nature des sites de fixation entre tanin et protéine ou encore l'aspect thermodynamique du complexe (stœchiométrie et constante de dissociation) sont des approches complémentaires quant à la compréhension de l'interaction Tanin-Protéine. Le complexe est ainsi étudié par le biais de la spectrométrie de masse, la RMN ainsi que la dynamique moléculaire.

La multiplicité des techniques utilisées nous amène à considérer un premier chapitre davantage théorique afin de bien comprendre le but de chacune d'entre elles dans notre étude, le chapitre II. Ensuite, la synthèse et la purification des polypeptides IB7₁₄ et IB7 ainsi que du dimère tannique B3, font l'objet du chapitre III. L'étude structurale du peptide IB7₁₄ et de la protéine IB7 fait l'objet du chapitre IV. Ceci a été réalisé par Dichroïsme Circulaire (DC), RMN et dynamique moléculaire. La mise en présence de dimères B3 avec le peptide ou la protéine entraîne la formation d'un complexe. Le chapitre V relate de cette interaction qui a été mise en évidence par DC, spectrométrie de masse, RMN (1D et 2D) ainsi que par dynamique moléculaire. Une conclusion générale et quelques perspectives termineront ce mémoire.

*Introduction :*

*Interaction Tanin-Protéine : la clé du « goût » ?*

## **Chapitre II METHODES PHYSICO- CHIMIQUES**



L'analyse structurale des peptides (IB7 et IB7₁₄) et de la procyanidine B3, ainsi que l'étude de l'interaction de ces composés au sein d'un complexe non covalent ont été entreprises par différentes méthodes. Le dichroïsme circulaire (DC) nous apporte des renseignements de structure secondaire, puis la spectroscopie de résonance magnétique nucléaire (RMN) multidimensionnelle suivi de la dynamique moléculaire (DM) nous permettent d'accéder à la conformation tertiaire des peptides. L'interaction a pu être mise en évidence par spectrométrie de masse (SM), RMN et DM également. Nous allons rappeler brièvement les principes de ces différentes techniques.

## II. 1. LE DICHROÏSME CIRCULAIRE

Le dichroïsme circulaire est une technique largement développée pour étudier la structure secondaire des protéines, car les spectres sont obtenus rapidement et sur de faibles quantités de produits (de 8 à 80 µg pour le peptide IB7₁₄ et de 35 à 350 µg pour IB7). Cette spectroscopie de différence est sensible à la quantité et à l'activité optique des molécules chirales, qui, en interaction avec une lumière polarisée, modifient leur absorbance. En effet, un composé chiral aura un spectre de DC caractéristique contrairement à un mélange racémique. Dans le cas des protéines, seul l'un des composés optiquement actif est présent donc chaque protéine possède une empreinte spectrale. Les spectres dichroïques sont très sensibles aux repliements de la chaîne peptidique, et présentent alors des courbes différentes selon la structure secondaire et l'arrangement spatial qu'adoptent les protéines⁸⁶. Ainsi, le pourcentage de chaque type de structure (hélice  $\alpha$ , feuillet  $\beta$ , hélice de type II ou pelote statistique . . .) peut être obtenu en déconvoluant le spectre DC d'une protéine grâce à des spectres de peptides possédant des structures secondaires définies.

### II. 1. 1. Principe

Selon la loi de Beer-Lambert, si  $I_0$  est l'intensité de la lumière entrant dans la cellule et  $I$  l'intensité sortant, l'absorbance  $A$  est définie par la relation suivante :

$$A = \log \left( \frac{I_0}{I} \right) = \epsilon c l$$

$c$  étant la concentration dans l'échantillon (en molaire),  $l$  la longueur de la cellule de mesure (en cm) et  $\epsilon$  est le coefficient d'extinction molaire (ou coefficient d'absorption molaire, en  $M^{-1}.cm^{-1}$ ). Le dichroïsme circulaire est l'absorption différentielle par un chromophore des deux composantes circulaires gauche et droite du champ électrique d'une lumière polarisée plane. On mesure la différence d'absorption de la lumière entre les deux coefficients d'extinction molaire, définie de la manière suivante :

$$\Delta\epsilon = \left( \frac{\epsilon_{\text{D}}}{\epsilon_{\text{M}}} - \frac{\epsilon_{\text{M}}}{\epsilon_{\text{D}}} \right) \frac{1}{cl} \Delta A$$

Après la traversée de l'échantillon, la résultante de ces composantes décrit une ellipse dans le plan perpendiculaire à la direction de propagation. L'ellipticité, appelée  $\psi$ , est définie par  $\tan \psi = (a/b)$ ,  $a$  et  $b$  étant respectivement les demi petit et demi grand axes de l'ellipse⁸⁷. L'ellipticité molaire est enfin décrite par la relation suivante :

$$[\theta] = \frac{100 \Psi}{cl} = 3300 \Delta\epsilon \quad [1]$$

L'ellipticité molaire  $[\theta]$  est exprimée en  $\text{deg}.dmol^{-1}.cm^2$ ,  $\psi$  en degrés.

Dans le cas des protéines, le chromophore dominant l'absorption dans l'UV lointain (180 à 240 nm) est la liaison peptidique. Chaque liaison peptidique étant un chromophore, l'ellipticité molaire mesurée doit être pondérée par le nombre de liaisons peptidique dans la protéine étudiée, noté  $n$  :

$$[\theta] = \frac{3300 \Delta\epsilon}{n} \quad [2]$$

L'ellipticité molaire  $[\theta]$  est notée de la même manière que pour l'équation [1] en accord avec la littérature, mais elle devient l'ellipticité molaire par résidu.

## II. 1. 2. Courbes standard et méthode de déconvolution

### II. 1. 2. a. Méthode de déconvolution

Il existe de nombreuses méthodes pour déconvoluer les spectres de DC. En général, la déconvolution des spectres expérimentaux utilise une bibliothèque de structures connues d'après des expériences préalables de RMN ou de diffraction aux rayons X⁸⁸⁻⁹⁰. Il est donc

d'autant plus fiable que le nombre de structures dans la base de données est important. La fiabilité des structures standards a été assurée par l'utilisation de peptides synthétisés au laboratoire. De plus, les déconvolutions effectuées au moyen des banques préexistantes utilisent des protéines standards dont les familles de structures ont été déterminées par rayons X ou RMN. Or, certaines structures comme l'hélice de type II ou encore les protéines membranaires n'apparaissent pas dans les banques données, d'où l'intérêt de confectionner nous-mêmes une banque de données à l'aide de protéines connues ayant des structures secondaires que l'on retrouve dans les protéines étudiées comme les PRPs.

L'ellipticité est caractéristique de chaque structure, or nous allons considérer pour nos protéine et peptide courts, qu'il n'existe que les quatre conformations dont les dichrographes caractéristiques sont présentés figure II-1. L'ellipticité globale est alors déterminée par la relation suivante⁹¹ :

$$[\theta] = [\theta]_{\alpha}f_{\alpha} + [\theta]_{\beta}f_{\beta} + [\theta]_{II}f_{II} + [\theta]_p f_p$$

où  $[\theta]_{\alpha}$ ,  $[\theta]_{\beta}$ ,  $[\theta]_{II}$ ,  $[\theta]_p$ , sont les ellipticités molaires par résidu correspondant à chaque structure standard, et  $f_{\alpha}$ ,  $f_{\beta}$ ,  $f_{II}$ ,  $f_p$  les fractions de résidus appartenant à chaque type de structure ( $\alpha$  pour l'hélice  $\alpha$ ,  $\beta$  pour l'hélice  $\beta$ , II pour l'hélice de type II et p pour la pelote statistique). La proportion de chaque structure au sein de la protéine suit la relation :  $f_{\alpha} + f_{\beta} + f_{II} + f_p = 1$ . Un programme de déconvolution codé en FORTRAN a été mis au point au laboratoire (Dufourc, non publié). Ce programme détermine toutes les combinaisons possibles et détermine la meilleure solution par la méthode des moindres carrés.

### **II. 1. 2. b. Obtention des courbes standard**

Afin d'accéder aux structures secondaires des protéines, la réponse dichroïque des différentes structures secondaires classiques (le feuillet  $\beta$ , l'hélice  $\alpha$ , l'hélice de type II ou encore la pelote statistique - désordonnée) de protéines connues doit être enregistrée. Les spectres dichroïques de protéines de structures définies ont été enregistrés en utilisant les standards suivants :

- la poly(L-lysine) à température ambiante et pH neutre (7,6) prend une forme intégralement en pelote statistique, représentant une structure aléatoire. Une poly(L-lysine) de 20 résidus a été synthétisée par SPPS selon la stratégie Fastmoc dans les mêmes conditions de synthèse que IB7₁₄ (cf § III. 1. 2).

- la poly(L-lysine) est également capable d'adopter les conformations hélice  $\alpha$  ou feuillet  $\beta$  (pH = 11,4 à température ambiante et à 60°C respectivement)⁹². Cependant, en raison du manque de précision obtenu avec la poly(L-lysine) pour l'hélice  $\alpha$  et le feuillet  $\beta$ , nous avons utilisé des standards dont les motifs structuraux ont été déterminés par spectroscopie InfraRouge^{93,94}. Chaque peptide de référence présente une structure en hélice  $\alpha$  (hélice droite avec 3,6 résidus par tour, c'est le motif structural le plus répandu dans les protéines) ou bien en feuillet  $\beta$  (forme particulière d'une hélice à deux résidus par tour parallèle, antiparallèle ou mixte) pour l'ensemble de leurs résidus. Un peptide poly(L-valine) a également été synthétisé par SPPS dans le but de mimer parfaitement l'hélice  $\pi$  (hélice droite « aplatie » avec 5 résidus par tour).

- l'hélice de type II est le composant structural majeur des conformations appelées « non-ordonnées » dans la littérature⁹⁵. Cependant, elle correspond à une hélice allongée, typique de la poly(L-proline), qui est une hélice gauche avec 3 résidus par tour. Nous utilisons la poly(L-proline) de Sigma, qui a une masse proche de celle de la protéine IB7.

Les spectres dichroïques des quatre structures secondaires standard sont montrés figure II-1. Chaque motif structural possède un spectre dichroïque caractéristique. En effet, le spectre de l'hélice  $\alpha$  possède trois points singuliers : un maximum d'intensité très important à 192 nm, puis deux minima d'intensités voisines à 208 et 220 nm. Le feuillet  $\beta$  conduit à un spectre de moins grande intensité que l'hélice  $\alpha$  avec un maximum d'absorption à 198 nm et un minimum à 217 nm. La structure en pelote statistique a ses maximum et minimum inversés par rapport au feuillet  $\beta$ : un minimum à 196 nm et un maximum de très faible intensité à 220 nm. L'allure du spectre représentant l'hélice de type II est proche du spectre de la pelote statistique, mis à part pour l'intensité et pour le déplacement de son minimum à 206 nm.

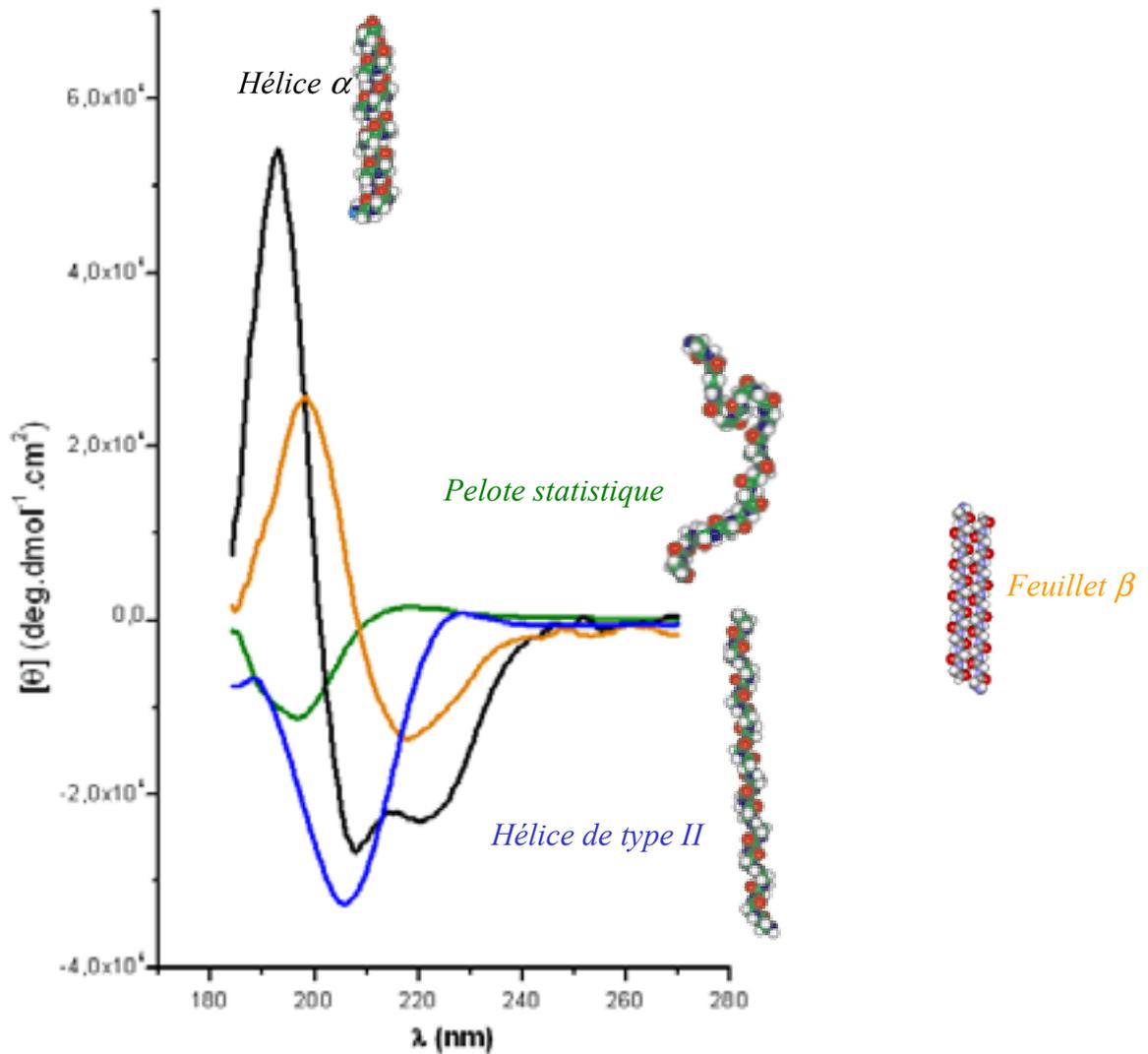


Figure II-1 : Spectres dichroïques de peptides caractéristiques des principales structures secondaires de protéines : l'hélice  $\alpha$  (noir), le feuillet  $\beta$  (orange), l'hélice de type II (bleu), ainsi que la pelote statistique (vert).  $[\theta]$  est l'ellipticité molaire par résidu, définie dans l'équation [1].

## II. 2. LA SPECTROMETRIE DE MASSE (SM)

La spectrométrie de masse s'est révélée longtemps inefficace pour l'étude des composés faiblement volatiles, ce qui est le cas pour des molécules polaires et/ou de masses élevées comme les protéines. En effet, lorsque l'on chauffe ces composés afin de les introduire dans une source de spectromètre de masse, ils se pyrolysent. Ce phénomène rend alors leur étude difficilement réalisable par cette technique. Afin de surmonter ces problèmes, de nouveaux procédés ont été développés que l'on peut regrouper sous le titre « d'ionisation douce », notamment la désorption-ionisation laser assistée par matrice (MALDI, Matrix Assisted Laser Desorption Ionisation) et l'électronébulisation (ESI, ElectroSpray Ionisation). Ces deux techniques couplées à un analyseur à temps de vol (TOF, Time Of Flight) nous ont permis de contrôler la pureté à l'issue de la synthèse du peptide et de la protéine. De plus, l'ESI associé à un analyseur de type trappe à ion permet l'étude de complexes non covalents.

### II. 2. 1. Le MALDI

La technique d'analyse MALDI a connu un essor considérable au cours de ces dernières années pour l'analyse de molécules complexes, de biopolymères et de polymères synthétiques. Son succès repose sur des déterminations précises de masses moléculaires, des études structurales avec une sensibilité dépassant souvent une picomole allié à une gamme de masse très étendue (de 500 à >200000 Da).

Le MALDI doit son origine à Karas et Hillenkamp⁹⁶. Le principe de cette méthode consiste à diluer l'échantillon dans une matrice, ici l'acide  $\alpha$ -cyano-4-hydroxy-cinnamique, composée de petites molécules organiques absorbant fortement à la longueur d'onde du laser. Le mélange est déposé sur une cible métallique et « cristallise » après évaporation du solvant. La cible est ensuite introduite dans la source du spectromètre de masse, sous vide. L'échantillon est ensuite irradié par un faisceau laser impulsionnel. La relaxation de l'énergie absorbée conduit à une volatilisation très rapide du solide. Dans ces conditions les molécules intactes du composé analysé sont désorbées et ionisées. Les ions formés sont généralement du type  $(M+H)^+$ .

Les ions expulsés de la source MALDI sont accélérés et parcourent une distance  $l$  dans le tube à temps de vol (TOF, Time Of Flight) de l'analyseur en un temps  $t_m$  qui est

fonction de leur rapport masse sur charge. Les masses moléculaires des ions sont alors facilement analysées par la relation :

$$m/z = 2V\Delta t_m^2/l^2$$

où  $m$  et  $z$  sont la masse et la charge de l'ion;  $V$  est le potentiel accélérateur;  $l$  représente la longueur du tube analyseur et  $\Delta t_m$  est l'intervalle de temps entre l'impulsion laser et l'impact de l'ion sur le détecteur. L'équation ci-dessus montre que deux ions de masses différentes et de même charge auront des vitesses différentes. Ils auront donc des temps de vol différents dans l'analyseur, de telle sorte que l'ion le plus léger provoquera un impact sur le détecteur avant l'ion le plus lourd. Ainsi les protéines peuvent être triées de manière linéaire en fonction de leur masse.

Cette technique offre un grand nombre d'avantages pour l'étude de peptides et de protéines. Tout d'abord, la sensibilité avoisine la picomole et l'échantillon peut aussi être dissout dans une large gamme de solvant, détergents ou tampons sans perturber la mesure. Un autre avantage important tient du fait que la pureté du produit n'est pas nécessaire. Enfin, la lecture du spectre est facile étant donné que l'on lit une masse associée à une, deux, trois ou  $n$  charge.

A chaque étape de purification suivant les synthèses, une analyse de masse a été effectuée par la technique de MALDI-TOF en mode reflectron (Bruker Reflex III).

## II. 2. 2. L'ESI

L'ESI doit son développement à Fenn et Yamashita⁹⁷. Cette technique produit des molécules ionisées par électronébulisation. L'échantillon est introduit dans la source *via* un capillaire à l'aide d'un gaz de nébulisation chaud (azote). L'application d'un champ électrique entre ce capillaire et le cône d'entrée entraîne une accumulation de charges en surface du liquide, telle que les répulsions électrostatiques des charges entre elles sont supérieures à la tension superficielle qui garantit la cohésion. Le liquide se transforme alors en fines gouttelettes chargées. Lorsque le solvant s'évapore, grâce au chauffage de la source et à l'introduction du gaz de désolvatation, les gouttelettes deviennent plus petites mais gardent leur charge initiale. A cause des répulsions électrostatiques, la cohésion d'une goutte n'est

plus assurée lorsqu'elle atteint une taille limite par rapport à sa charge, la limite de stabilité de Rayleigh. A ce stade, les gouttelettes explosent en plus petites unités, de même polarité. Finalement, le processus aboutit à la formation d'ions multichargés qui, pris dans un flux gazeux, sont introduits dans la partie vide du spectromètre et analysés. L'électrospray peut aussi être couplé à un analyseur à temps de vol.

Un spectre de masse complexe est obtenu, montrant une distribution des ions multichargés. Un traitement mathématique permet de définir la masse réelle de la macromolécule. Elle peut être calculée en considérant les valeurs  $m/z$  de deux pics distincts d'une charge, de masse  $m_1$  et  $m_2$  respectivement :

$$m_1 = (M + n_1)/n_1 \text{ et } m_2 = (M + n_2)/n_2$$

Avec  $n_2 = n_1 - 1$  et  $m_1 < m_2$

Soit :  $n_1 = (m_1 - 1)/(m_2 - m_1)$  et  $M = (n_1 - 1)(m_2 - 1)$

Les nombres de protons  $n_1$  et  $n_2$  ainsi que la masse moléculaire  $M$  de la molécule sont alors identifiés.

L'ESI est très utilisé pour l'étude de petites molécules polaires et permet également l'accès à des protéines de haut poids moléculaire (jusqu'à 100 kDa). Cette technique a toutefois ses limites car lors de l'analyse de mélanges complexes chaque composé donne un ensemble de pics qui se superposent. Au-delà d'un mélange d'une dizaine de composés, il est impossible de résoudre le spectre. La sensibilité de l'ESI est la même que celle du MALDI (de l'ordre de la picomole), en revanche, elle ne supporte pas la présence de quantités importantes d'impuretés (sels, tampons . . .) et donc nécessite un dessalage préalable des échantillons⁹⁸.

### II. 2. 3. L'ESI couplé à la trappe à ions

La visualisation de complexes par l'ESI couplé à un analyseur de type trappe à ions consiste en l'isolement d'un composé de masse précise, puis en sa fragmentation en différentes unités non liées de manière covalentes. La distinction de chacune des masses de chaque unité, ainsi séparées, nous permet de connaître les membres appartenant au complexe originel.

### **Technique de la fragmentation des complexes non covalents : la trappe ionique**

La production des ions a lieu par l'utilisation de la technique nanospray, développée à partir de l'ESI, qui permet de réduire considérablement les quantités de matériel nécessaire. La compatibilité de cette méthode d'ionisation avec le piège ionique quadripolaire, ou « trappe à ions » permet par exemple, de sélectionner des ions d'une masse donnée au sein de l'échantillon. Une fois le nuage d'ions focalisé au centre du piège ionique, les trajectoires des ions sont rendues instables par une augmentation progressive de l'amplitude radiofréquence appliquée au niveau des électrodes. A chaque valeur de la tension, un ensemble d'ions de même valeur  $m/z$  sera éjecté. Ces ions éjectés sont ensuite détectés, en fonction de la masse des molécules, pouvant résulter d'un complexe non covalent composé de plusieurs unités indépendantes. Ensuite, une certaine quantité d'énergie fournie au système provoque une collision entre les différents ions présents, chacun représentant une unité d'un complexe. Ces ions sont alors libérés du complexe par pallier, ce qui peut nous amener à distinguer tous les complexes intermédiaires. Par exemple, si une masse ( $M_{C1}$ ) représente un complexe non covalent avec trois macromolécules, une première fragmentation génère des masses d'une unité seule ( $M_i$ ) et d'un complexe secondaire ( $M_{C2}$ ) ; une seconde fragmentation va libérer tous les intervenants du complexe ( $2M_i$  et  $M_j$ ) comme présenté schématiquement figure II-2 :

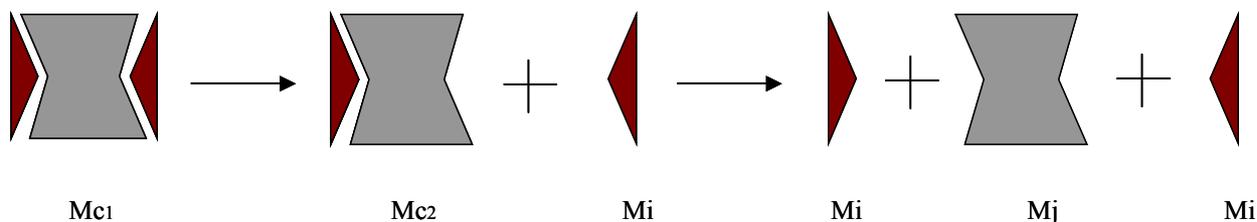


Figure II-2 : Représentation schématique de la méthode de fragmentation permettant de décomposer un complexe non covalent.

## II. 3. LA SPECTROSCOPIE DE RESONANCE MAGNETIQUE NUCLEAIRE

La spectrométrie de résonance magnétique nucléaire s'est largement développée pour l'étude des protéines : elle permet d'obtenir des informations sur leurs structures (primaires, secondaires et tertiaires), ainsi que sur leur dynamique et les interactions entre protéines et ligands. La RMN a été utilisée non seulement dans le but d'obtenir des informations structurales sur le peptide mais également lors de l'étude physico-chimique du complexe peptide-tanin formé, par mesure de la variation de déplacement chimique, de coefficient de diffusion, etc. L'intérêt de travailler dans des conditions les plus proches du milieu naturel nous impose de solubiliser nos échantillons dans l'eau. L'élimination du signal de l'eau ainsi que l'inconvénient relatif à la formation d'agrégats colloïdaux (complexes tanins-protéines) seront expliqués par la suite. Ces différentes approches vont être rappelées après un bref résumé du principe de la RMN, issu de la littérature⁹⁹⁻¹⁰¹.

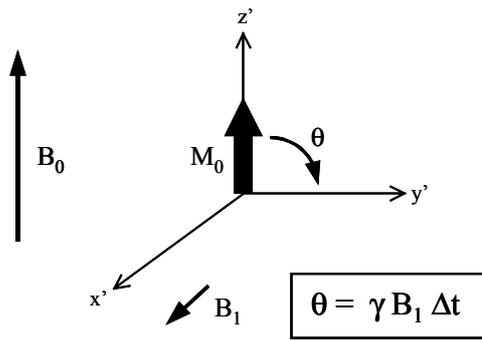
### II. 3. 1. Brefs rappels du principe de la RMN

Les noyaux de certains atomes possèdent un moment cinétique, et un moment magnétique, non nuls ce qui leur confère un spin nucléaire non nul (¹H, ¹³C, ³¹P par exemple). A l'équilibre, lorsque ces noyaux sont placés dans un champ magnétique principal, B₀, les spins précessent à la fréquence de Larmor, que l'on note ν₀. La valeur de ν₀ est liée à B₀ par la relation de Larmor :

$$\nu_0 = \frac{\gamma B_0}{2\pi}$$

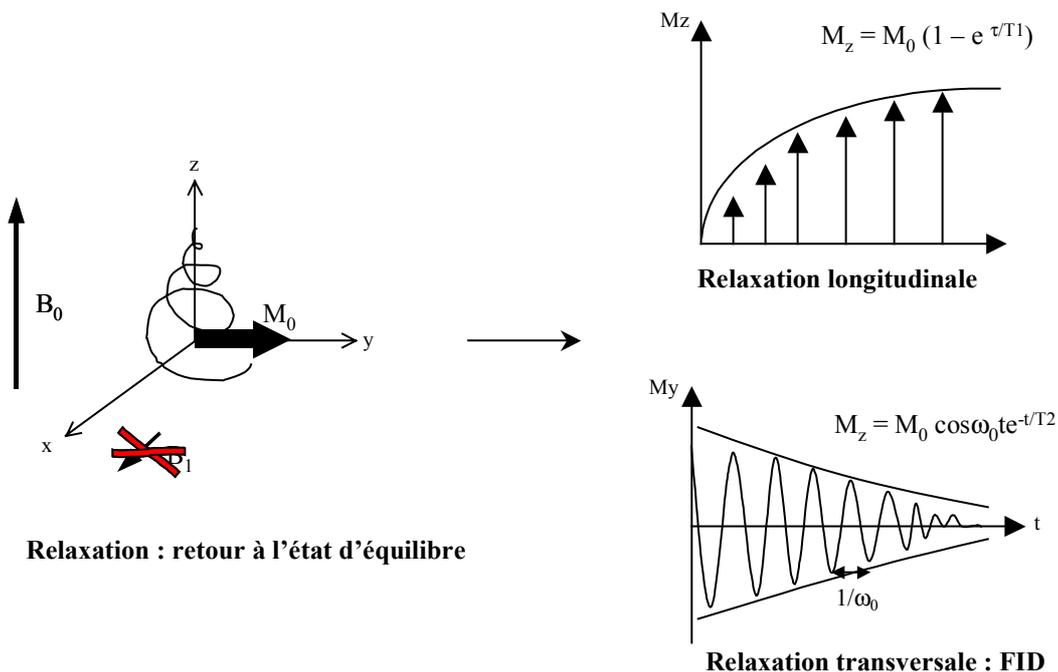
γ étant le rapport gyromagnétique, caractéristique de chaque noyau.

Si on applique sur l'échantillon, perpendiculairement au champ principal, un deuxième champ radiofréquence, B₁ ( $\vec{B}_1 = 2B_1 \cos \omega_0 \Delta t$ ), oscillant à une fréquence ν₀, le champ B₁ interagit avec les spins et perturbe leur distribution entre les niveaux d'énergie due au champ fixe B₀.



$B_1$  est appliqué pendant une durée  $\tau = \pi/(\gamma B_1)$  le long de l'axe  $x'$  (du référentiel tournant), de sorte que le vecteur résultant  $M_0$  bascule le long de l'axe  $y'$  (précession autour de  $B_1$ ). Une projection sur l'axe  $y'$  est obtenue pour un angle de basculement  $\theta = 90^\circ$  (schéma ci-contre).

Lorsque le champ radio-fréquence  $B_1$  cesse, le système retourne progressivement à l'état d'équilibre initial, phénomène appelé la relaxation. Ce retour à l'état d'équilibre peut se décomposer en une relaxation longitudinale (ou spin-réseau) selon l'axe  $z$  et une relaxation transversale (ou spin-spin) par rapport au plan  $xy$  de détection. La figure II-3 traduit le signal du système au cours de la relaxation décomposée selon l'axe  $y$  et selon l'axe  $z$ . L'évolution par rapport à  $y$  constitue le signal de précession libre des spins ou FID (Free Induction Decay), évoluant sous la forme d'une cosinusoïde amortie en fonction du temps.



**Figure II-3 :** Retour à l'équilibre après une impulsion de champ radiofréquence  $B_1$ . L'aimantation précesse au cours du temps à la fréquence de Larmor des signaux ( $\omega_0 = \gamma B_0$ ) et décroît dans le plan  $xy$  et croît le long de l'axe  $z$ . Le signal de précession libre, détectable le long de l'axe  $y$ , est une cosinusoïde d'argument  $(\omega_0 t)$  multipliée par une exponentielle décroissante caractérisant la relaxation transversale  $T_2$ . Le retour à l'équilibre sur l'axe  $z$  suit une exponentielle caractérisant la relaxation longitudinale  $T_1$ .

Le passage du domaine temporel au domaine fréquentiel est assuré par un outil mathématique, la transformée de Fourier et conduit à l'obtention d'un spectre composé de raies dont la position en fréquence est associée à l'environnement électronique des noyaux et la largeur des raies à la dynamique des molécules.

Le déplacement chimique  $\delta_i$  caractérise la résonance d'un noyau  $i$  et il est déterminé par sa position dans le spectre par rapport à une référence. Ecrit sous la forme  $\delta_i = [(v_i - v_{\text{réf}})/v_{\text{réf}}].10^6$ , il est indépendant de  $B_0$  et s'exprime en ppm. Afin de calibrer la valeur du déplacement chimique d'un spectre à l'autre nous utilisons l'acide propionique triméthylsilyl (TSP) comme référence interne (à 0 ppm).

### II. 3. 2. Détermination de la structure de Protéines

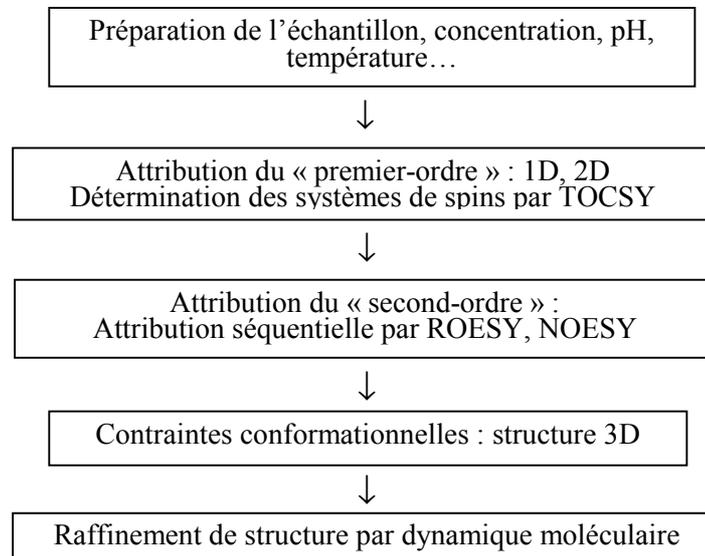
Dans une protéine le nombre de protons est très important ce qui entraîne un recouvrement partiel de nombreux pics. En outre, les PRPs sont constituées à 70% en trois résidus ce qui augmente le chevauchement des signaux. L'interprétation du spectre 1D est alors difficile (figure II-5 (B) et (C)). Cette difficulté est contournée en utilisant la RMN multidimensionnelle qui permet d'éclater les raies de résonance dans un espace à 2, 3 voire 4 dimensions. La RMN bidimensionnelle (2D) homonucléaire ( $^1\text{H} - ^1\text{H}$ ) nous a suffi quant à l'attribution des protons du peptide IB7₁₄. La RMN 2D hétéronucléaire ( $^{13}\text{C} - ^1\text{H}$ ) a quant à elle permis l'attribution ainsi que la reconnaissance de la régio-chimie de la liaison interflavane du dimère B3.

#### II. 3. 2. a. Stratégie d'attribution

La détermination structurale de protéines par RMN se déroule en plusieurs étapes, dont la stratégie a été définie par Wüthrich¹⁰² : un échantillon de protéine est solubilisé à une concentration suffisante (plusieurs mM) dans des conditions de pH et de température fixées. La première étape consiste en l'attribution de tous les protons de la protéine par acide aminé, chacun considéré comme un système de spin. La deuxième étape permet de relier les acides aminés en fonction de leur emplacement dans la chaîne peptidique et donc d'accéder à la structure primaire. Ensuite, un autre jeu de séquences multidimensionnelles procure des contraintes conformationnelles d'angles et de distances entre les atomes, qui sont enfin

utilisées par la modélisation moléculaire pour remonter à la structure secondaire et tertiaire de la protéine.

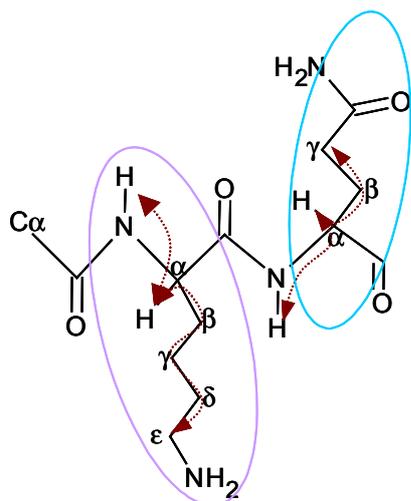
La stratégie d'attribution développée par Wüthrich est schématisée ci-dessous :



### ***II. 3. 2. b. Première étape : Identification des systèmes de spin***

L'attribution des protons de la protéine commence par l'utilisation de séquences de type COSY et TOCSY. La séquence COSY¹⁰³ met en œuvre des couplages scalaires entre les protons appartenant à un même système de spin (c'est-à-dire à un même résidu dans le cadre des protéines). Seuls les protons géminés ou vicinaux donnent naissance à des pics de corrélation (couplage  2J  ou  3J ). En revanche, la séquence TOCSY¹⁰⁴ met en évidence les couplages scalaires de type COSY et relayés existant entre tous les protons appartenant à un même système de spin. Ceci par l'introduction d'un temps de mélange isotrope permettant le transfert de l'aimantation entre les spins couplés.

Le schéma ci-dessous représente la répartition des protons par système de spin pour un dipeptide constitué d'une lysine (●) et d'une glutamine (●).

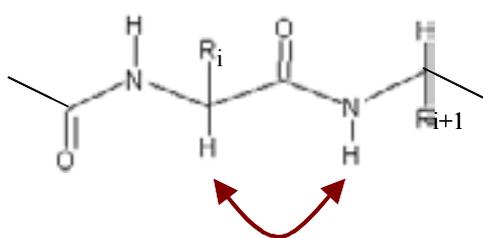


Le schéma ci-contre représente des acides aminés par système de spin. Les groupes de protons sont corrélés entre eux par couplage J et identifiables lors des expériences de type TOCSY.

Cette étape permet de cartographier tous les systèmes de spin en attribuant tous les protons par résidu.

### II. 3. 2. c. Deuxième étape : Structure Primaire

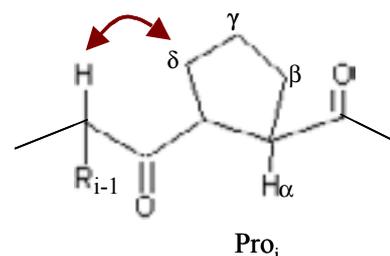
Après l'identification de chaque système de spin, il faut leur attribuer une position exacte dans la séquence primaire en utilisant la méthode d'attribution séquentielle développée par Wüthrich¹⁰². La méthode repose sur l'existence systématique dans la carte NOESY ou ROESY de corrélations (interaction spin-spin directe) entre le proton  $H\alpha$  d'un résidu (i) et le proton HN du résidu suivant (i+1), schématisé ci-dessous : la méthode de détermination de



l'enchaînement en acides aminés par attribution séquentielle est effectuée grâce aux corrélations entre le  $H\alpha_i$  et le  $NH_{i+1}$  obtenues avec des séquences de type NOESY ou ROESY.

Cette méthode est valable pour tous les acides aminés mis à part la proline dont la chaîne latérale est cyclisée sur l'azote tertiaire.

Des informations séquentielles sont fort heureusement accessibles en observant les effets NOEs entre les  $H\delta$  de la proline et le  $H\alpha$  du résidu précédent (i-1).



Les expériences utilisées, NOESY et ROESY, sont basées sur l'interaction entre protons proches dans l'espace, interaction dipôle-dipôle (moins de 6 Å de distance) et s'appuient sur l'effet NOE (Nuclear Overhauser Enhancement)¹⁰⁵. Celui-ci correspond à une variation de l'aimantation (intensité des signaux) d'un spin  $i$  lorsqu'on perturbe l'aimantation d'un spin  $s$  en interaction avec  $i$  au travers des phénomènes de relaxation (couplages dipolaires). L'effet NOE est négatif, nul ou positif en fonction de la masse molaire des peptides ou protéines.

L'expérience ROESY^{106,107} diffère de l'expérience NOESY par l'introduction d'un champ radiofréquence supplémentaire (le « spin lock ») pour verrouiller la composante de l'aimantation transversale le long de l'axe  $y'$  du référentiel tournant. Dans ces conditions l'effet NOE est toujours positif quelle que soit la taille du peptide. L'expérience ROESY a donc été privilégiée lors de l'étude du peptide IB7₁₄ où les expériences NOESY ne donnaient aucun signal.

### ***II. 3. 2. d. Troisième étape : Etablissement d'une liste de contraintes***

A ce stade de l'analyse, il est nécessaire d'obtenir des informations spatiales pour le repliement du squelette ou la conformation des chaînes latérales. Deux paramètres nous apportent des informations structurales : les valeurs de constante de couplage  3J  et les effets NOE. Les valeurs de distances inter proton et d'angle de torsion, déterminées expérimentalement permettent d'établir des listes de contraintes géométriques, utilisées par la suite pour la détermination de structure spatiale à l'aide de la modélisation moléculaire.

#### **Effets NOE :**

Lorsque tous les effets NOE séquentiels ( $H\alpha_i$  et  $NH_{i+1}$ ) sont attribués, il reste les taches de corrélations impliquant des proximités spatiales ( $i \rightarrow i+2, i+3, i+4\dots$ ), indiquant une estimation de la distance des protons dans l'espace en fonction de l'intensité des pics (entre 2 Å et 6 Å). L'examen des effets NOE permet de repérer l'existence de motifs de structure secondaire ainsi que des repliements tridimensionnels de la protéine. En effet, l'intensité des pics corrélés en ROESY est proportionnelle à l'inverse de la distance entre deux protons. De fait, si une distance entre deux protons est connue, que l'on nomme  $r_{\text{réf}}$ , il est alors possible de comparer une distance à déterminer  $r_i$ , selon la relation suivante :

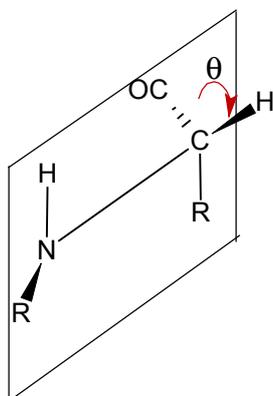
$$r_i = r_{\text{réf}} \left( \frac{S_{\text{réf}}}{S_i} \right)^{1/6}$$

où  $S_{\text{réf}}$  et  $S_i$  sont respectivement l'intensité de la tache de corrélation entre les protons dont la distance est connue et l'intensité d'une tache entre proton dont la distance reste à déterminer.

Les contraintes de distances sont obtenues par l'intensité des pics corrélés en ROESY et déterminées par XEASY¹⁰⁸.

### **Constantes de couplage 3J :**

Les valeurs des constantes de couplage scalaire  3J  sont corrélées aux valeurs des angles dièdres entre atomes voisins donc nous informent sur les contraintes d'angles du squelette protéique¹⁰⁹. Elles sont mesurées soit directement sur le spectre 1D, soit grâce à l'utilisation d'une séquence COSY avec un filtrage double quanta (COSY-DQF)^{110,111}.



La relation existant entre les constantes  3J  et les angles dièdres  $\theta$  est définie selon la relation de Karplus :

$$^3J = A \cos^2 \theta + B \cos \theta + C$$

A, B et C sont des constantes, dont les valeurs sont respectivement 6,4, -1,4 et 1,9 Hz dans le cadre des protéines¹¹².

### ***II. 3. 2. e. Structure tridimensionnelle***

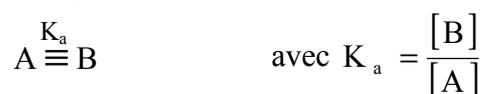
Les renseignements structuraux issus de la RMN, comme les valeurs de déplacements chimiques, et les contraintes d'angles et de distances sont intégrées dans le logiciel DYANA¹¹³. Il nous permet de générer un fichier de contraintes sous forme de

coordonnées spatiales. Ce fichier est ensuite utilisé pour effectuer un raffinement de structure par modélisation moléculaire (cf § II-4).

### II. 3. 3. Dynamique

La RMN permet également d'obtenir des informations dynamiques sur des systèmes en échange.

Soit un échange entre deux espèces A et B :



La figure II-4 montre l'aspect du spectre en fonction de la vitesse de l'échange entre les formes A et B :

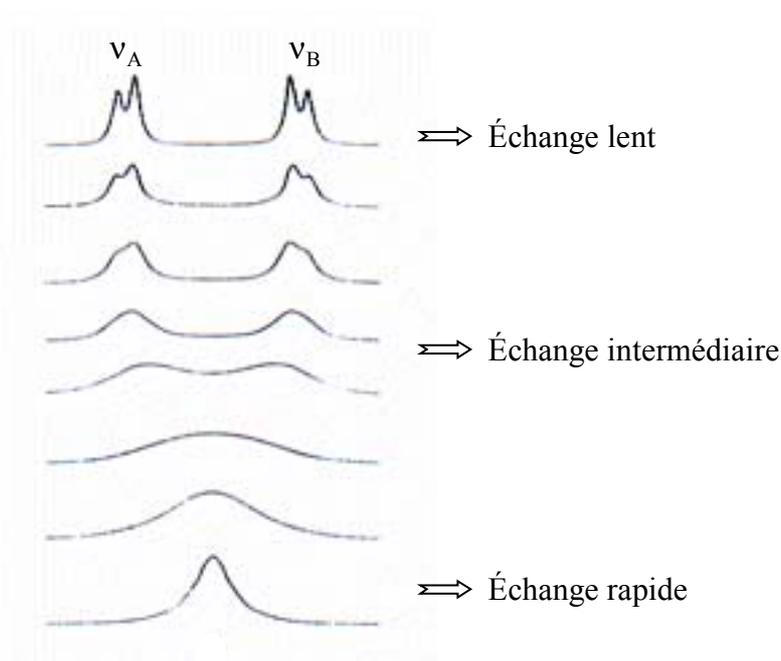


Figure II-4 : Variation de la nature du spectre RMN  $^1\text{H}$  en fonction de la vitesse de l'échange.

Dans le cas d'un échange dit « rapide », un signal unique est observable qui représente un déplacement chimique moyen dont la position dépend de la fraction de A ( $f_A$ ) et de B ( $f_B$ ) dans le mélange, tel que :

$$\delta = f_A \delta_A + f_B \delta_B$$

Ainsi, si l'on suit la variation de la valeur du déplacement chimique ( $\delta$ ) en faisant varier la concentration de l'une des composantes de l'échange (A ou B) on peut remonter à  $K_a$ ¹¹⁴. Cette propriété sera appliquée à la détermination de la constante de dissociation taninprotéine (cf chapitre V).

### II. 3. 4. Coefficient de Diffusion translationnel, D

La diffusion translationnelle d'une molécule dans un solvant suit un mouvement brownien de temps de corrélation  $\tau_c$  caractérisé par un coefficient de diffusion D. Ce dernier est associé à la taille de la molécule comme le montre l'équation de Stokes-Einstein [3] :

$$D = \frac{kT}{6\pi\eta r} \quad [3]$$

où  $k$  est la constante de Boltzmann,  $\eta$  la viscosité de l'eau (en poiseuille), et  $r$  le rayon hydrodynamique de la molécule considérée sphérique.

Le suivi du coefficient de diffusion D du peptide en fonction de la concentration en B3 rajouté dans la solution, nous permettra de rendre compte de la taille du complexe ainsi obtenu mais également de la constante de dissociation, de la même manière que la variation du déplacement chimique.

#### Mesure des coefficients de diffusion

Les mesures de coefficient de diffusion translationnel sont réalisées à l'aide de séquences unidimensionnelles consistant en un écho de spin stimulé en parallèle à l'utilisation de gradients bipolaires (séquence de type `stebpgp1d` - Bruker)^{115,116}.

La technique consiste à faire disparaître le signal au fur et à mesure de l'augmentation des gradients. La puissance et le temps d'impulsion des gradients sont calibrés de telle sorte qu'entre l'application de 2% et 95% de gradient  $z$ , il ne reste qu'environ 5 à 10% du signal.

Il existe une relation directe entre l'intensité du signal résiduel et le gradient appliqué, définie par la relation [4] :

$$\ln(I/I_0) = -[\gamma^2 \times \delta^2 \times G_z^2 \times (\Delta - \delta/3)] D \quad [4]$$

Avec  $I_0$ , l'intensité du pic avec un gradient  $z$  de 2%, et  $I$  l'intensité résultante ;  $\gamma$  étant le rapport gyromagnétique.  $\Delta$  et  $\delta$  sont respectivement le temps de diffusion et la durée d'application des gradients. Lorsqu'on trace  $\ln(I/I_0)$  versus  $G_z^2$ , la pente permet de remonter à  $D$ .

L'étape préliminaire est la calibration du gradient propre au spectromètre. Pour cela différents échantillons dont le coefficient de diffusion est connu (sous certaines conditions de concentration et de température), sont testés. Par exemple, l'eau dans l'eau à 25°C a un  $D \approx 2 \cdot 10^{-9} \text{ m}^2 \cdot \text{s}^{-1}$ . Par un cheminement inverse la valeur du gradient effectif ( $G_e$ ) est de  $G_e = 0,55 \times G_z$  pour la sonde BBI du spectromètre 400 MHz donc il faudra systématiquement corriger la valeur du gradient par 0,55.

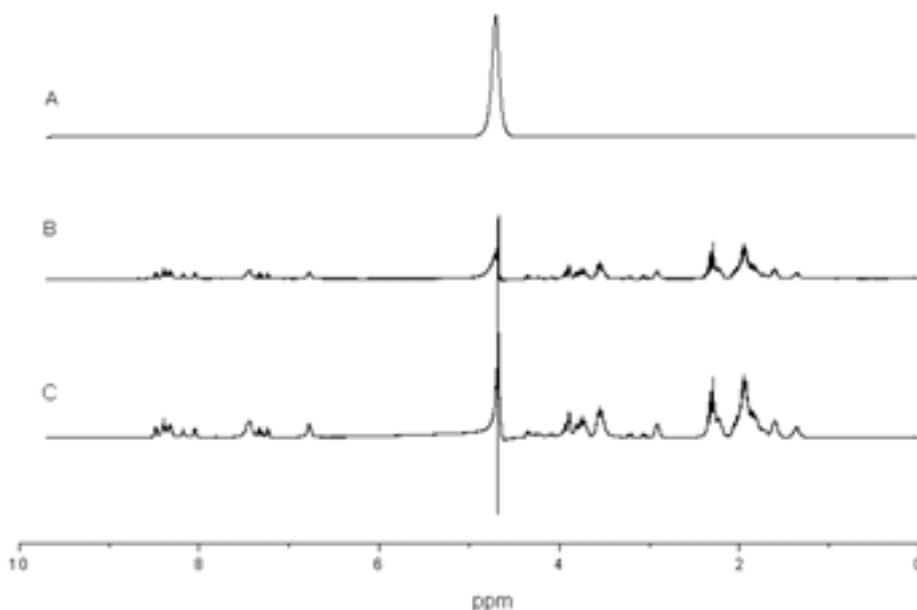
## II. 3. 5. Spectres de protéines dans l'eau et complexées avec les tanins

### II. 3. 5. a. *Elimination de la résonance de l'eau*

L'intérêt de travailler dans de l'eau est primordial pour s'approcher des conditions naturelles. Or la concentration des protons dans l'eau avoisine les 110 M, ce qui est de très loin supérieur aux concentrations de protéine ou de complexe utilisés ( $10^{-3}$  M). De plus, travailler à 100% de  $D_2O$  provoquerait une extinction des protons échangeables. Il est donc indispensable de supprimer la résonance des protons de l'eau pour visualiser ceux de la protéine ou du dimère B3. Deux techniques sont envisageables : la présaturation ou la technique WATERGATE (WATER suppression by Gradient Tailored Excitation).

La présaturation¹¹⁷ est une séquence simple dans laquelle une puissance d'irradiation, centrée sur le signal de l'eau, est insérée en début de séquence (puissance d'irradiation à 48 dB). Le WATERGATE a été introduit par Piotto *et al.*¹¹⁸ en utilisant des gradients de champ.

La figure II-5 représente trois spectres obtenus respectivement avec une simple impulsion (A), avec une présaturation (B) et enfin avec un WATERGATE (C).



**Figure II-5 :** Nécessité de la suppression du signal de l'eau. Spectres 1D du peptide IB7₁₄, avec une séquence d'impulsion classique (A), avec une présaturation sur le signal de l'eau (puissance de présaturation 48 dB) (B), avec un WATERGATE (gradients 3-9-19) (C).

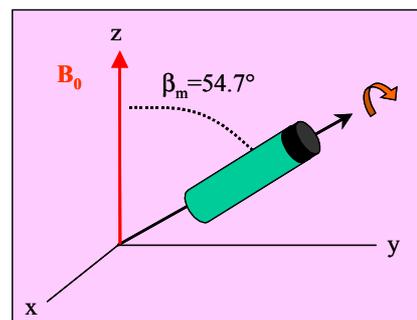
Chacune des techniques a ses avantages et ses inconvénients dans le cas de l'étude de protéines. En effet, lors d'une expérience de présaturation l'élimination du signal de l'eau a lieu avant l'acquisition donc tous les protons de l'eau ou ceux en échange avec l'eau vont être perturbés. Lors de l'expérience WATERGATE, les gradients sont appliqués pendant l'acquisition et donc n'affectent pas les protons labiles. De ce fait, l'élimination du signal de l'eau par présaturation peut diminuer l'intensité des pics protons des groupements NH en échange avec le solvant que l'on irradie. Malheureusement, la technique WATERGATE ne permet pas de visualiser correctement les pics sortant à proximité de la résonance de l'eau. Il faudra jongler avec ces deux méthodes en fonction des groupements que l'on souhaite distinguer.

### **II. 3. 5. b. HR-MAS**

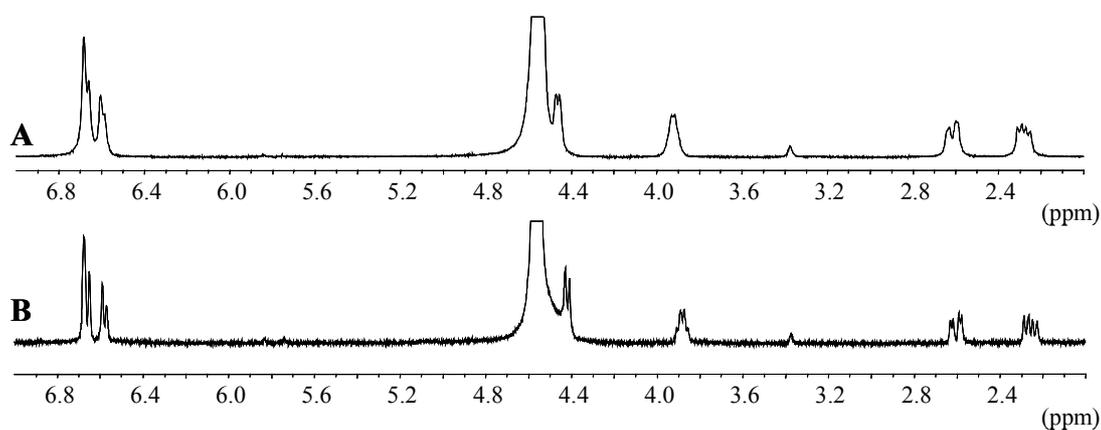
L'utilisation d'une sonde HR-MAS (High Resolution - Magic Angle Spining) a plusieurs intérêts. Les échantillons étudiés ici peuvent se retrouver sous forme colloïdale, d'une part par la nature colloïdale des tanins⁷⁷ et d'autre part parce que les complexes tanins-protéines formés sont susceptibles de former des agrégats. Or les spectres RMN de composés

colloïdaux présentent des raies élargies dues aux interactions de RMN des solides qui ne sont plus annulées par les mouvements rapides des molécules. En effet, les agrégats colloïdaux ont des mouvements beaucoup plus lents que les molécules isolées en solution. Le blindage électronique devenant anisotrope et le couplage spin-spin direct provoquent l'élargissement des raies de résonance. Toutes les interactions ont une loi d'échelle en  $\frac{3\cos^2\theta-1}{2}$ . Ce terme s'annule pour un angle  $\theta$  de  $54,7^\circ$ .

La rotation à l'angle magique ( $54,74^\circ$ ) permet d'éliminer les potentielles interactions dipolaires pouvant apparaître sur des solutions à la frontière entre le liquide et le solide. Les interactions magnétiques internes anisotropes sont alors généralement moyennées à condition que la vitesse de rotation de l'échantillon soit suffisamment rapide. Le spectre obtenu en rotation est donc typique d'un spectre liquide haute résolution.



La figure II-6 montre les spectres de la catéchine après agrégation (en plusieurs jours) en statique (A) et en rotation à l'angle magique (B). Il semble évident que la rotation nous ramène à un spectre résolu.



**Figure II-6 :** Spectres  $^1\text{H}$  de la (+)-catéchine, A : enregistré dans une sonde liquide et B : enregistré dans une sonde HR-MAS¹¹⁹.

En outre, une sonde de ce type avec gradients est disponible sur un spectromètre d'un champ de 11,7 T, à 500 MHz sur le proton, ce qui va tout d'abord nous permettre

d'augmenter la résolution et le rapport Signal/Bruit par rapport au 400 MHz. De plus, cette sonde permet de travailler sur des rotors de 80  $\mu\text{l}$  ce qui nécessite donc moins de matériel (pour une solution à 5 mM, il faudrait 2,6 mg pour 400  $\mu\text{l}$  et 0,52 mg pour 80  $\mu\text{l}$  pour le peptide IB7₁₄, et on passe de 11,3 mg à 2,27 mg pour la protéine IB7).

## II. 4. LA MODELISATION MOLECULAIRE

La modélisation moléculaire utilise des méthodes de calcul théorique (mécanique moléculaire, dynamique moléculaire, mécanique quantique *ab initio* ou semi-empirique) confrontées à des expériences réalisées sur la matière : ces deux approches expérimentales ou purement théoriques deviennent alors complémentaires. La dynamique moléculaire propose une description de la matière à l'échelle atomique ainsi que son suivi au cours du temps. Cette analyse répond à de nombreuses questions sur la structure de protéines et sur la dynamique des systèmes.

### II. 4. 1. Principes de la dynamique moléculaire

La dynamique moléculaire est une méthode d'étude des mouvements et de l'évolution de la configuration spatiale des systèmes moléculaires. La technique repose sur la nature des liaisons et des atomes, avec leurs forces associées. Elle dépend également du milieu dans lequel l'échantillon est placé. Ces deux notions vont être abordées uniquement dans le cadre des expériences réalisées sur l'étude des protéines et des complexes.

#### Le champ de forces

Le mouvement des atomes est inhérent à tout système chimique et les vibrations atomiques dépendent de la température. Au niveau moléculaire, cet effet se traduit par une compression/extension des liaisons ou par la variation des angles de valence ou de torsion. Ainsi une molécule est considérée comme un ensemble d'atomes dont les interactions mutuelles peuvent être décrites par une fonction d'énergie potentielle ou champ de forces. Les atomes représentés par des sphères sont reliés dans le cas d'une liaison covalente par des ressorts. La distribution des électrons, considérée comme implicite, est localisée sur les liaisons du système et représentée par la distribution de charges sur les atomes de chaque

molécule. Ces charges peuvent être calculées de façon à reproduire le moment dipolaire de la molécule ou la surface de potentiel électrostatique obtenue par calcul *ab initio*.

On appelle « champ de forces », le fichier où tous les paramètres de charges, de constantes d'oscillateur ou de rayons de Van der Waals sont définis. On y distingue deux types de termes énergétiques : les termes liés et les termes non liés. Il existe divers champs de forces mais seul AMBER¹²⁰ a été utilisé pour ce travail. Il est reconnu comme le meilleur champ de forces pour l'étude de peptides et de protéines.

Ce champ de forces est défini par trois termes liés pour une molécule et trois termes non liés, selon l'équation suivante :

$$E_T = \overbrace{E_{liaison} + E_{valence} + E_{torsion}}^{\text{Termes liés}} + \underbrace{E_{H-H} + VdW + e-}_{\text{Termes non liés}}$$

Les termes liés sont les énergies de déformation de liaisons, des angles de valence, des angles de torsions et des interactions en dehors du plan (figure II-7). Pour avoir l'expression complète de l'énergie d'une molécule, il faut aussi rajouter les termes d'énergie non liés, dus à l'interaction électrostatique d'une part et à l'interaction de Van der Waals d'autre part. Dans le cadre des protéines, un terme spécifique pour les liaisons hydrogène est indispensable.

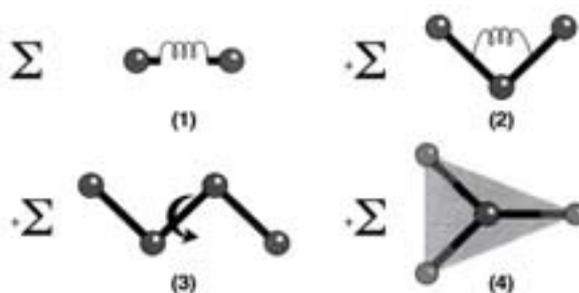


Figure II-7 : Illustration graphique des quatre termes d'énergie liée du champ de force AMBER ; déformation (1) des liaisons, (2) des angles, (3) des angles de torsion et (4) des interactions en dehors du plan. Les atomes sont représentés par des sphères et les liaisons par des ressorts.

### **Le solvant implicite et la périodicité**

Il est important de replacer notre échantillon protéique dans un solvant mimant la réalité. Les calculs ont été faits dans un solvant implicite, l'eau ( $\epsilon = \approx 78,5$ ), où les atomes de l'eau ne sont pas représentés mais remplacés par un potentiel continu, contrairement au solvant explicite qui nécessitait alors beaucoup trop de temps de calculs. Ce solvant implicite ne considère que les énergies de solvation et de cavitation par rapport à la molécule placée artificiellement dans l'eau.

### **Le Cut-Off**

Le nombre de paires d'atomes à considérer augmente avec la taille du système étudié. En effet, si le système étudié comporte 1 000 atomes, la somme des interactions non-liées est  $\sim 500\,000$  termes. Une distance de troncature ou "Cut-Off" est utilisée pour limiter les temps de calcul et réduire ce nombre. Une revue regroupant plusieurs méthodes de Cut-Off a été publiée par Brooks *et al.*¹²¹.

Le Cut-Off est une distance limite au-delà de laquelle les interactions "non-liées" (van der Waals et électrostatiques) sont ignorées. Le Cut-Off peut être défini par paire d'atomes ou par groupe. Avec un Cut-Off défini par paire d'atomes, ne sont calculées que les interactions d'un atome X avec les autres atomes contenus dans la sphère centrée sur X et ayant le Cut-Off pour rayon.

La notion de « cut-off » permet donc de réduire le nombre d'interactions en estimant qu'au-delà d'une certaine distance les interactions sont négligeables.

## **II. 4. 2. Minimisation des structures de protéines**

Lors de la reconstruction d'un système biologique par modélisation moléculaire, les atomes peuvent en un premier temps prendre des positions arbitraires. Tous les angles ne sont pas possibles pour de simples considérations géométriques. Des modèles atomiques, prenant en compte les distances interatomiques connues et les rayons de van der Waals, permettent d'observer que certains angles sont beaucoup plus probables que d'autres selon la nature des acides aminés. Des conformations sont donc plus stables que d'autres demandant moins d'énergie au système. Il est donc impératif de relaxer la géométrie du système afin de diminuer le coût énergétique. Ce mécanisme s'appelle la minimisation.

La minimisation de la structure du peptide a été réalisée en maintenant les contraintes obtenues par RMN sur les angles dièdres et les distances entre les protons. Les minima atteints après minimisation dépendent de la géométrie de départ. Ceci est dû à la complexité des systèmes étudiés et à la multitude de minima de l'hypersurface de l'énergie potentielle qui en résulte (représenté schématiquement figure II-8). Le logiciel DISCOVER permet entre autres deux méthodes de minimisation que nous avons également utilisées : la méthode des gradients par "steepest descent" (SD) (littéralement "à descente le plus rapide")¹²² ou celle des gradients conjugués (CG) (Conjugate Gradient)¹²³. L'utilisation successive des deux techniques permet dans un premier temps, par SD de descendre rapidement vers un minimum local, puis, par CG, d'affiner le calcul. Cette méthode est une méthode de choix pour les gros systèmes, le temps d'intégration est plus long que par SD car la minimisation se fait pour toutes les directions, cependant cette perte est compensée par une meilleure convergence vers un minimum. La minimisation s'arrête lorsqu'un critère de convergence RMS (Root Mean Square) du gradient de l'énergie potentielle  $V$  atteint une valeur inférieure à celle définie par l'utilisateur ou le nombre d'itérations défini par l'utilisateur est atteint.

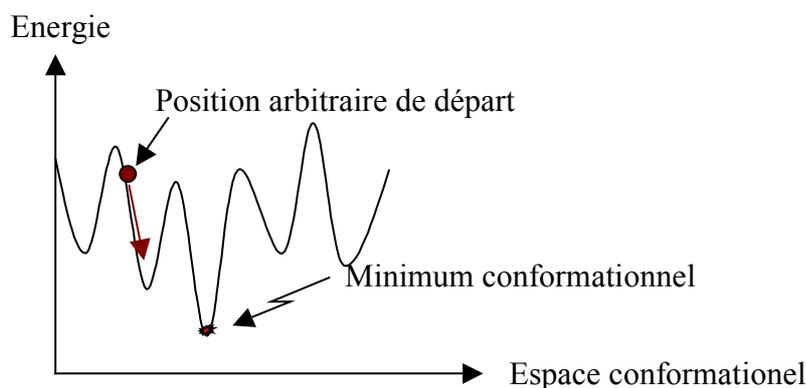


Figure II-8 : Représentation schématique des puits d'énergie.

### II. 4. 3. Dynamique

La simulation en dynamique moléculaire (DM) utilise la seconde loi de Newton pour décrire le mouvement d'une molécule en fonction du temps :

$$F_i = m_i \times a_i$$

Où  $F_i$  est la force,  $m_i$  la masse et  $a_i$  l'accélération de l'atome  $i$ . Comme le montre l'équation de Newton, la vitesse et le sens du mouvement atomique dépendent des forces qui s'exercent entre les atomes. Les vitesses des particules sont réparties selon une distribution de Maxwell-Boltzmann qui centre la distribution des vitesses autour d'une vitesse moyenne choisie en fonction de l'énergie cinétique que l'on attribue au système. Elle est habituellement obtenue en élevant la température d'un système moléculaire placé initialement au zéro absolu et en amenant ainsi progressivement les atomes qui le constituent à l'énergie d'équilibre. La dynamique moléculaire permet de calculer la force exercée sur chaque atome et de déterminer les différentes informations liées à la connaissance de cette donnée (vitesse et position des atomes pendant un intervalle de temps spécifique, de l'ordre de 1 ps à quelques ns suivant les systèmes moléculaires).

La force  $F_i$  qui s'exerce sur un atome  $i$  peut être directement calculée depuis la dérivée de l'énergie potentielle  $V$ . La mécanique moléculaire ne permet de calculer l'énergie que pour des applications n'induisant pas de modifications importantes de la structure électronique (ex : rupture/formation d'une liaison). La chimie quantique donne une valeur de l'énergie permettant d'étudier des processus dynamiques entraînant des changements chimiques mais cette application est d'utilisation limitée compte tenu de la durée des calculs nécessaires à la détermination de la fonction énergie. La dynamique moléculaire permet d'étudier la trajectoire des atomes à partir de la connaissance des forces s'exerçant sur chaque atome et des masses atomiques.

Par la dynamique nous avons accès au minimum d'énergie par une traversée rapide de tous les puits d'énergie (représenté figure II-8). La technique de la dynamique stochastique a été utilisée à ces fins. Elle prend en compte la viscosité et les termes de chocs aléatoires, ceci en élevant la température pour favoriser le passage des barrières d'énergie.

**Chapitre III SYNTHÈSE D'UNE PRP DE LA  
SALIVE HUMAINE ET D'UN TANIN DU RAISIN**



### **III. 1. PRODUCTION DE PROTEINES ET DE TANINS : SYNTHESE OU EXTRACTION DE PRODUITS NATURELS ?**

L'étude structurale de protéines ainsi que la mise en évidence d'un complexe entre tanin et protéine, suivie de l'étude physico-chimique de ce complexe requiert des quantités relativement importantes de matériel. La synthèse des deux protagonistes impliqués a donc été optimisée dans ce but. Nous allons décrire dans ce chapitre les voies empruntées pour l'obtention de produits purs. La protéine IB7 étant une protéine naturelle de la salive humaine, trois voies d'obtention sont envisageables : l'extraction directe à partir de la salive, la synthèse par voie chimique ou bien la surexpression. La salive constitue un mélange de protéines, de carbohydrates et d'ions ayant pour principale fonction de lubrifier les muqueuses¹⁸. La grande diversité des constituants rend difficile l'extraction d'une protéine en particulier. De plus, les protéines peuvent subir des modifications post-traductionnelles, comme par exemple l'hydroxylation des prolines¹²⁴, ce qui ajoute encore de la multiplicité. Dans ce cas, purifier et caractériser une protéine de masse aléatoire n'est pas aisé. Par le biais de la synthèse chimique nous pourrions travailler sur une protéine homogène et de composition primaire connue. Cependant, certains peptides sont connus pour leur difficulté à être synthétisés comme les peptides hydrophobes en raison de leur faible solubilité en milieu aqueux¹²⁵. Certains résidus requièrent également une attention particulière de part leur nature chimique. Les prolines engendrent la formation de coudes qui peuvent gêner l'attachement à l'acide aminé suivant. La glycine, le plus petit des résidus, est capable de s'enfouir entre les différentes chaînes en élongation et donc masquer l'extrémité réactive. Les protéines riches en proline sont, comme leur nom l'indique, constituées de nombreux résidus proline (40 %) et glycine (21 %). Pour ces raisons, la voie de synthèse sur de longues protéines riches en proline n'est pas utilisée. En effet, dès lors que les peptides riches en proline étudiés possèdent plus d'une vingtaine de résidus, l'extraction directe à partir de la salive est préférée à la synthèse par voie chimique : la synthèse a donc été privilégiée pour des tétra-peptides à des peptides d'une vingtaine de résidus^{124,126-130}, alors que l'extraction a majoritairement été effectuée dans le but d'obtenir un ensemble de protéines de la salive humaine et non une protéine isolée^{21,92,131-133}.

Le défi que constitue la synthèse peptidique de PRPs a été entrepris.

En ce qui concerne les tanins, deux voies d'obtention de dimères sont également possibles : l'extraction directe de tanins à partir de pépins de raisin ou la synthèse chimique. Les deux techniques ont été utilisées lors de ce travail. Comme nous le verrons, la synthèse chimique a donné de meilleurs résultats.

Nous allons donc présenter ici la synthèse en phase solide de la protéine, qui nous a permis d'isoler des quantités suffisantes de protéines ainsi que la synthèse par couplage cationique du dimère B3.

## **III. 2. LA PROTEINE IB7 ET LE PEPTIDE IB7₁₄**

### **III. 2. 1. Synthèse du peptide et de la protéine**

La synthèse peptique sur support solide de IB7 et IB7₁₄ (dont les séquences primaires sont reprises ci-dessous), a été entreprise grâce à un synthétiseur de peptides semi-automatique.

**SPPGKPOGPPPQGG¹⁴NQPOGPPPPGKPOGPPPQGGNKPOGPPPPGKPOGPPPQGDNKS⁵⁹**

#### ***III. 2. 1. a. Stratégie de la synthèse : Synthèse Peptidique en Phase Solide (SPPS)***

Merrifield en 1963 introduit la synthèse de peptide en phase solide (SPPS) pour créer une chaîne peptidique linéaire¹³⁴. Depuis, cette technique a été largement développée et reconnue comme la meilleure méthode de synthèse de peptides, permettant l'obtention rapide de peptides plus longs. Il existe deux méthodes de chimie en synthèse sur support solide, les chimies Boc ou Fmoc, la première nécessitant un milieu acide et la seconde un milieu basique, lors des déprotections. La méthode Fmoc, plus douce, a été choisie au détriment de la Boc qui requiert l'utilisation d'acide fluorhydrique hautement acide.

La caractéristique de la SPPS est l'utilisation d'un support solide, qui d'une part, apparaît inerte vis-à-vis des solvants et réactifs de synthèse et d'autre part, permet une solvation optimale du peptide en élongation. Il se présente sous la forme d'une résine constituée de billes de polystyrène réticulé à 1% de divinyl-benzène. Cette résine est préchargée : c'est-à-dire qu'elle porte le résidu aminé par lequel la synthèse doit commencer, soit le carboxy-terminal (*C-ter*) car la synthèse se fait dans le sens *C-ter* vers l'extrémité amino-terminale (*N-ter*).

Tous les acides aminés sont protégés au niveau des groupements amines par un groupement Fmoc ce qui permet de rendre l'extrémité *N-ter* momentanément non réactive. Afin de prévenir les réactions secondaires entre les  $\alpha$ -amines et les chaînes latérales fonctionnelles, d'autres protections masquent les groupements réactifs des chaînes latérales, comme suit :

- le groupement hydroxyle de la serine et la fonction acide carboxylique de l'acide aspartique sont protégés par un t-butyle ;
- l'ammonium de la lysine porte un Boc ;
- les amides primaires des glutamines et asparagines portent un trityle ;
- un 2,2,5,7,8-pentaméthylchroman-6-sulfonyl (PMC) masque le groupement guanidium de l'arginine.

### **III. 2. 1. b. Etapes de la synthèse**

La synthèse d'un peptide est réalisée par fixation séquentielle de chaque acide aminé constituant la chaîne peptidique. Chaque fixation d'un résidu consiste en une série d'étapes résumées dans la figure III-1.

La résine est porteuse du premier résidu aminé, la première étape dans l'assemblage de la chaîne est donc la **déprotection** du groupement Fmoc de l'acide aminé préchargé sur la résine, pour laisser une extrémité  $\alpha$ -amine libre à l'extrémité peptidique accessible. Cette déprotection est assurée par une amine secondaire, la pipéridine.

La seconde étape est l'**activation** de l'acide aminé libre (et encore muni de son Fmoc en *N-ter*), à fixer en deuxième position, par transformation en ester activé. Différentes méthodes d'activation sont alors possibles, comme les milieux HOBt/DCC, l'anhydride acétique symétrique ou le HBTU-FastMoc.

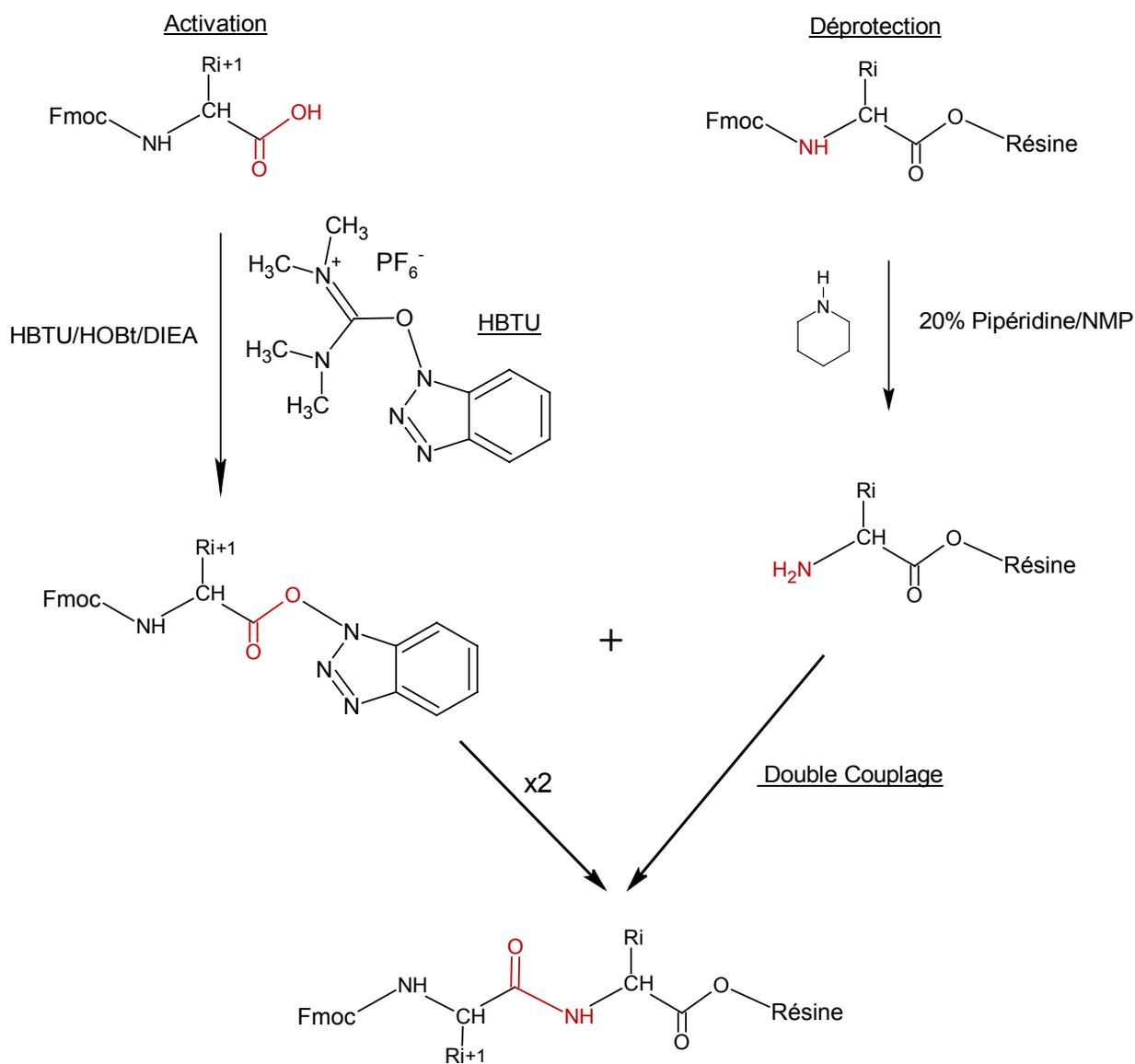


Figure III-1.: Schéma de synthèse sur support solide par la méthode FastMoc. L'activation par HBTU a été introduite d'après Dourtoglou *et al.*¹³⁵, et l'adaptation à la stratégie Fmoc par Fields¹³⁶.

Le succès du couplage d'acides aminés en phase solide dépend non seulement de la réactivité du groupement carboxyle de l'acide aminé protégé mais aussi de l'accessibilité du résidu terminal du peptide en cours d'élongation fixé à la résine¹³⁷. Par exemple, les acides aminés esters HOBt sont connus pour réagir rapidement avec les amines¹³⁸ mais si le temps pour l'activation est insuffisant, la concentration d'ester activé n'est pas suffisante pour faire réagir complètement les sites amines durant le couplage. Les solvants de couplage et d'activation doivent être compatibles avec les solvants de solvatisation du peptide-résine¹³⁹. L'HBTU, combiné à l'HOBt et DIEA, produit une forte activation des Fmoc-acides aminés¹⁴⁰ ceci plus rapidement qu'un activateur de type HOBt/DCC. La méthode FastMoc tient son nom de l'utilisation de HBTU en plus de l'HOBt (Applied Biosystem). Les deux réunis entraînent une réaction cinétiquement favorisée et une meilleure solvatisation de la résine¹⁴¹. Donc l'HBTU est dissout dans une solution de HOBt dans DMF, puis le tout est mélangé à une solution contenant du DIEA dans NMP.

L'acide aminé libéré est alors prêt à recevoir le résidu suivant que l'on a donc préalablement activé par une solution de HOBt/HBTU.

Durant le **couplage**, l'ester activé forme une liaison amide avec le résidu déprotégé en cours d'élongation (figure III-1). La mise en présence des deux groupements réactifs est réalisée en deux passages successifs : un premier acide aminé est présenté à la résine avec un excès de 4 résidus libres pour une chaîne en élongation. Une fois le couplage terminé, la résine est lavée puis incubée à nouveau avec le même acide aminé. Après ce double couplage, les sites n'ayant pas réagi lors du premier passage d'un résidu sont encore susceptibles d'être fixés au cours du second passage. En ce cas, l'objectif à chaque fixation se situe à 98% du couplage. Un des problèmes rencontrés lors du couplage est la formation de liaisons hydrogène intra peptidiques. Elles pourraient alors entraîner la formation d'« agrégats » bloquant ainsi l'accessibilité de l'acide aminé déprotégé pour l'acide aminé libre. Afin de pallier cette éventualité le couplage est réalisé en rajoutant une base (DIEA) dans du NMP (couplage anhydre).

Malgré toutes les précautions prises pour atteindre une réactivité maximale des résidus et donc une fixation optimale, certains acides aminés ( $aa^{n+1}$ ) ne fixent pas les acides aminés ( $aa^n$ ) libres, laissant place à une fonction aminée active pouvant être fixée par le résidu présenté ultérieurement ( $aa^{n-1}$ ). Ce phénomène donnerait naissance à des peptides possédant des délétions en milieu de chaîne ( $aa^{n-1} - aa^{n+1}$ ), comme montré sur la figure III-2. Afin d'éviter ces synthèses tronquées en milieu de chaîne, le blocage de toutes les fonctions amines

qui n'ont pas réagi est réalisé par un traitement à l'anhydride acétique : c'est l'étape de **capping**. Ce produit va bloquer les groupements aminés libres du Fmoc n'ayant fixés aucun résidu, ce qui inhibera la fixation avec un quelconque acide aminé pour éviter de créer des séquences incomplètes en milieu de chaîne.

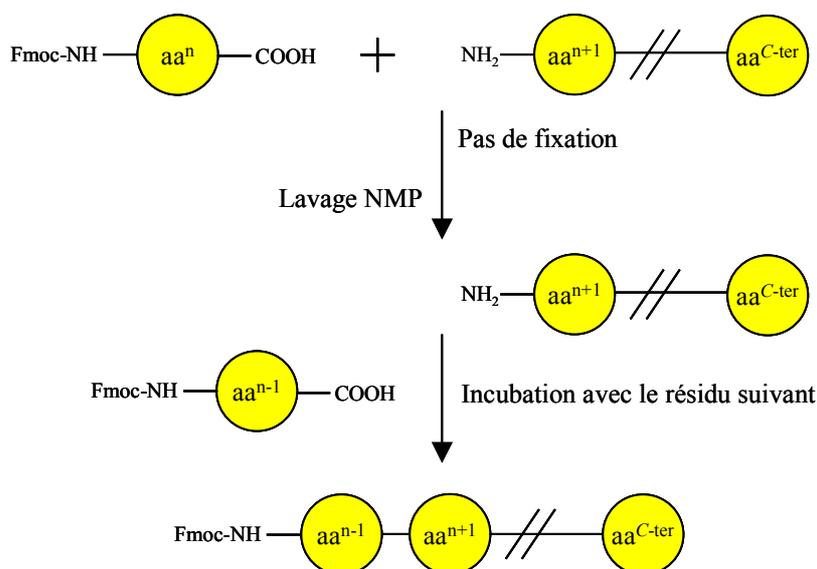
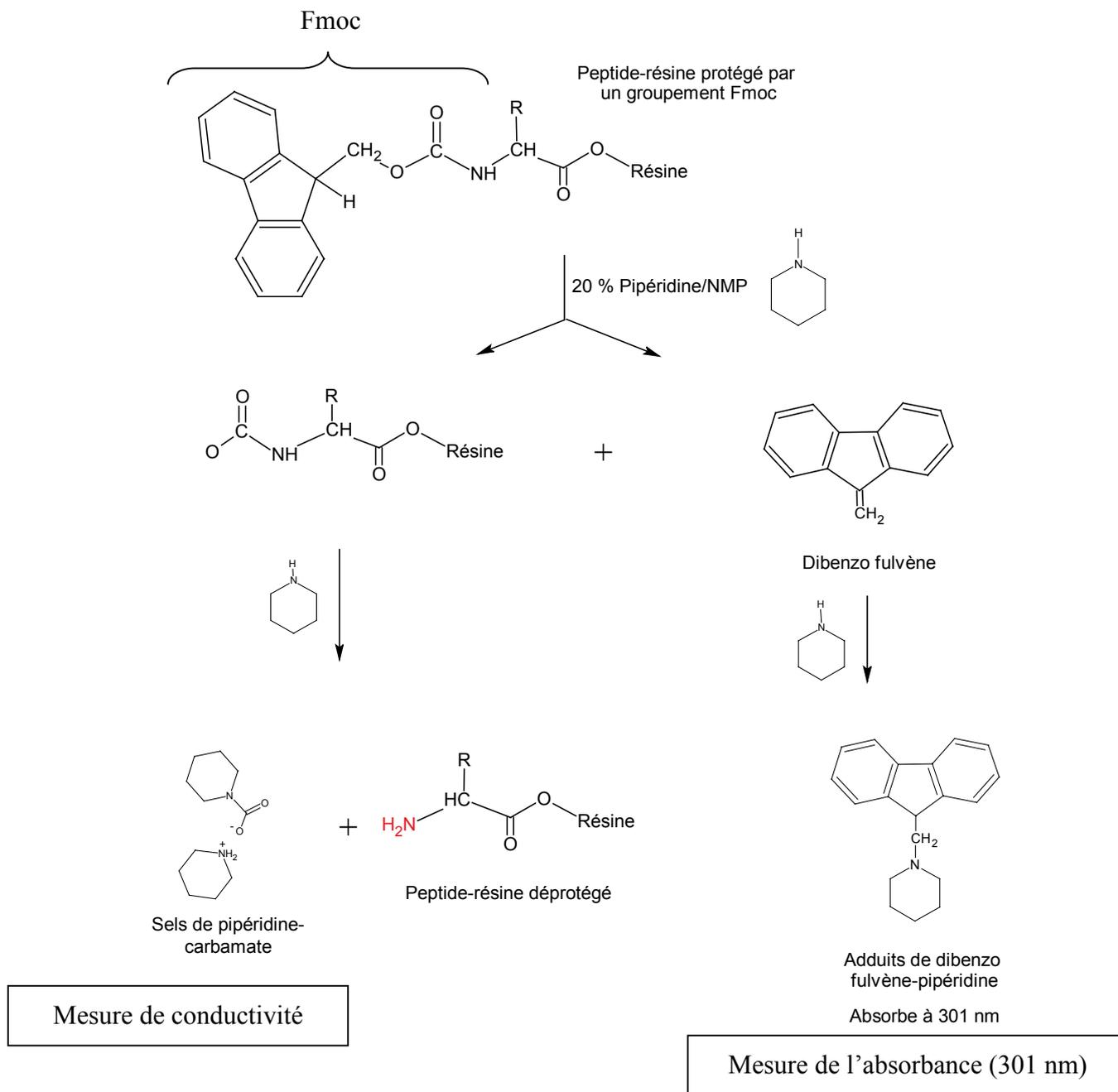


Figure III-2 : Représentation de la formation de peptides tronqués en milieu de chaîne.

Un nouveau cycle peut alors commencer par la déprotection de l'amine sur le peptide en élongation. Ces étapes sont alors répétées jusqu'à l'obtention du peptide complet.

### Suivi de la synthèse par mesure de la conductivité

Au cours de la synthèse, des réarrangements structuraux du peptide peuvent se produire. Or, cette redistribution peut interférer avec la capacité de déprotection et de couplage. Le suivi du succès de l'élongation est réalisé en temps réel après chaque déprotection de manière automatique par la mesure de la conductivité. Le principe repose sur l'étroite corrélation entre la déprotection et le couplage¹⁴². Un faible taux de déprotection indique un couplage faible ou incomplet. Rajoutée en excès l'action de la pipéridine a lieu à deux niveaux successifs. La première génère un dibenzofulvène qui réagit en un second temps avec une molécule de pipéridine pour donner un dibenzofulvène-pipéridine, quantifiable en U.V., et un sel de carbamate détectable par mesure de conductivité (figure III-3). L'efficacité



**Figure III-3 :** Mode d'action de la pipéridine et produits secondaires formés permettant le suivi en temps réels de la synthèse par mesure de la conductivité.

de la déprotection a pu être suivie grâce à un système de mesure automatique de la conductivité installé sur le synthétiseur ABI 433A. Un exemple d'enregistrement schématisé est montré sur la figure III-4. L'action de la pipéridine est effective par cycles (avec un minimum de quatre) à la fin desquels une mesure de conductivité est réalisée. L'étape de déprotection cesse dès lors que le dernier pic observé est égal au précédent (à 10% du signal de départ), ce qui suggère que la conductivité mesurée provient du bruit de fond.

### ***III. 2. 1. c. Isolement et purification***

Une fois la synthèse arrivée à terme après fixation du dernier résidu, le peptide est séparé de son support solide, la résine et les groupements protecteurs des chaînes latérales sont clivés simultanément.

Il reste alors à purifier le peptide parmi les morceaux de séquences tronquées, par chromatographie liquide haute performance. La vérification de la présence de peptide de masse connue est effectuée par spectrométrie de masse de type MALDI.

## **III. 2. 2. Matériels et Méthodes**

### ***III. 2. 2. a. Synthétiseur***

La synthèse peptidique a été réalisée à l'aide d'un synthétiseur de peptide semi-automatique Applied Biosystems Synthesizer 433A (PE Biosystem, Courtaboeuf, France). La méthode de chimie s'est orientée vers la stratégie FastMoc, notée FastMoc-CondMonPrevPeak chez le fabricant. La résine, les Fmoc-acides aminés, protégés au niveau des groupements latéraux, ainsi que HOBt et HBTU proviennent de la compagnie Novabiochem (Läufelfingen, Switzerland) et les solvants de SDS (Peypin, France).

Les solvants et réactifs sont délivrés au synthétiseur sous pression. La composition des bouteilles, numérotées par le fabricant, est détaillée ci-dessous :

- ❶ Bouteille de déprotection : 20% de pipéridine dans NMP,
- ❷ Bouteille de capping : anhydride acétique/DIEA/HOBt (19 ml/9 ml/0,8 g, v/v/v) dans 400 ml de NMP,
- ❸ Bouteille d'activation : HBTU/HOBt (37,9 g/13,6 g) dans 200 ml de DMF,
- ❹ Bouteille d'activation et couplage : 35% de DIEA dans NMP

⑨ et ⑩ Bouteilles de lavage : DCM et NMP respectivement.

### III. 2. 2. b. Synthèse du peptide IB7₁₄ et clivage de la résine

0,1 mmol de résine Wang préchargée avec une glycine sont utilisées. Les acides aminés libres sont présentés en excès de 10 par rapport à la résine.

Le double couplage est un module conditionnel du programme de gestion de l'automate, il dépend de la fixation précédente, dont le bon déroulement est suivi par la mesure en temps réel de la conductivité, comme présenté figure III-4. Le second couplage n'est effectué que lorsque la dernière déprotection observée n'est pas conforme aux restrictions demandées pour la déprotection (§ III.2.1.b suivi de la conductivité).

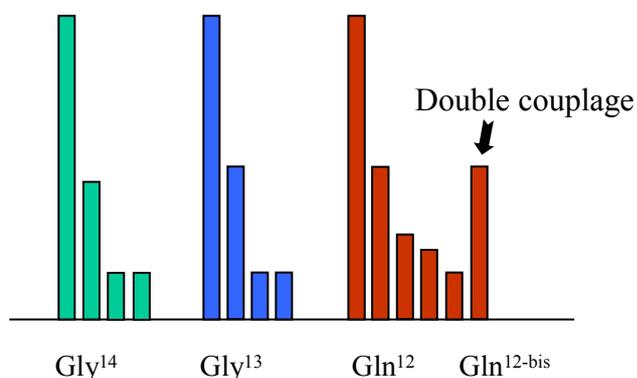


Figure III-4 : Représentation du suivi de la conductivité après action de la pipéridine. Le double couplage est un module conditionnel fonction de la qualité de la déprotection précédente.

A la fin de l'élongation, le peptide fixé à son support solide, est récupéré. Le clivage de la résine ainsi que des groupements protecteurs des chaînes latérales du peptide IB7₁₄ est réalisé simultanément en milieu acide. Le peptide est placé dans un mélange TFA/thioanisole/H₂O Milli Q (95/3/2, v/v/v), 10 ml pour ~0,8 g de peptide-résine, sous agitation non magnétique pour ne pas broyer la résine. Après 120 mn par un passage de 0°C à température ambiante, la résine est filtrée du mélange de peptides libres et des groupements protecteurs sur verre fritté. L'ensemble est lavé avec TFA puis DCM. Ces deux solvants sont ensuite évaporés sous pression réduite et le brut de synthèse est récupéré par une extraction au diéther froid/H₂O (10/10, v/v), répétée deux fois pour chaque phase. Le peptide étant

soluble dans l'eau, seuls les réactants du clivage se retrouvent dans la phase étherée et cela permet une première élimination grossière des groupements protecteurs. La phase aqueuse contenant le peptide est alors lyophilisée.

### ***III. 2. 2. c. Synthèse de IB7 et clivage de la résine***

La synthèse du peptide IB7₁₄, représentant la partie amino-terminale de la protéine IB7, ne nécessite pas de précautions particulières car sa taille le rend relativement facile à synthétiser. Au contraire, pour arriver à son terme avec un rendement correct, la synthèse de la protéine entière IB7 requiert quelques modifications.

La méthode utilisée est la FastMoc à 0,25 mmol d'arginine prégreffé sur la résine, avec un excès d'un facteur 4 en acides aminés libres. La résine utilisée est la résine NovaSyn[®] préchargée avec de l'arginine, que nous avons préférée à la résine Wang, reliée au premier acide aminé par un groupement HMP. En effet, le taux de substitution de la résine NovaSyn[®] est moins élevé (entre 0,1 et 0,3 mmol.g⁻¹ de résine) ce qui la rend plus adaptée à l'élongation de longues chaînes peptidiques. De plus la résine NovaSyn[®] est dotée d'un polyéthylèneglycol qui constitue un bras espaceur empêchant un retour de l' $\alpha$ -amine sur la résine. Ceci permet d'obtenir un meilleur rendement.

Au préalable, nous avons réalisé une plus longue solvataion (doublée) de la résine grâce à deux passages successifs de DCM ainsi que de NMP.

Au niveau du synthétiseur, chaque étape de la synthèse est représentée par un module qui consiste en une série de commandes que nous pouvons modifier, en maintenant un langage Applied Biosystem cohérent. Certains modules de la méthode FastMoc2.5 sont donc optimisés dans ce but :

φ Module B : déprotection : étape 21, fonction 51 : 14 ; soit une augmentation de 30% du temps de délivrance de la pipéridine, étape 22, fonction 56 : 5 ; lavage au NMP

φ Module A : activation : étape 13, fonction 94 : 10 ; étape 14, fonction 98 : 10 ; soit une augmentation de 20% du pourcentage d'anhydride acétique délivré,

φ Module C : capping : étape 3, fonction 52 : 42 ; étape 4, fonction 40 : 5 ; étape 5, fonction 41 : 5 ; soit la quantité d'agents activants (HBTU, HOBt, DIEA) est augmentée ~20% ;

φ Module F : coulage : étape 22, fonction 98 : 35 ; soit le temps de couplage entre le peptide en cours d'élongation et les acides aminés libres a été rallongé de 17% pour chaque résidu.

Les temps d'action de la majorité des produits sont ainsi augmentés le double couplage est rendu obligatoire pour tous les acides aminés.

A la fin de l'élongation, la protéine IB7 est récupérée pour subir un traitement acide dans le but de récupérer un produit libéré de la résine et comportant des chaînes latérales fonctionnelles, sans groupements protecteurs. Il est à noter que le processus de clivage et d'extraction diffère entre le peptide IB7₁₄ et la protéine IB7 entière. IB7 requiert une solution de clivage plus complexe en raison de la substitution du groupement protecteur de l'arginine principalement. Le mélange réactionnel est alors composé de TFA/thioanisol/H₂O/éthanedithiol/cristaux de phénol (84/4/4/2/6, v/v/v/v/g), tel que 20 ml de mixture réagissent avec ~ 2 g de peptide-résine. Cette réaction clivante doit être de plus courte durée possible (ici 150 mn) pour diminuer le contact du peptide néo-synthétisé avec les réactifs de clivage. En effet, une action prolongée de la mixture de clivage pourrait produire des groupements PEG gênant l'étape de purification suivante.

### ***III. 2. 2. d. Purification par Chromatographie Liquide à Haute Performance***

L'analyse ainsi que la purification des peptides et protéines sont effectuées par chromatographie liquide haute performance en phase inverse (RP-HPLC) WATERS dotée d'un détecteur Alliance. La phase stationnaire est une colonne analytique greffée de chaînes aliphatiques de 18 carbones (5 μm, 100 Å, 3,9 x 150 mm). Les solvants d'élution sont en A de l'eau à 0,1 % de TFA et en B de l'acétonitrile à 0,08 % de TFA, délivrés avec un débit de 1 ml/mn. La purification est effectuée sur des colonnes semi-préparative (15 μm, 100 Å, 7,8 x 300 mm) permettant un meilleur débit (2,5 ml/mn) donc le passage d'une plus grande quantité de matériel à chaque injection. La sortie de colonne est reliée à un collecteur de fractions préalablement programmé afin de récolter les fractions un certain temps après le début de l'élution. Ceci est possible car le pic contenant le peptide est élué de la colonne à un temps donné parfaitement reproductible d'une injection à l'autre.

La présence de protéine ou de peptide, est caractérisée par la valeur de la densité optique à 214 nm, spécifique de la liaison peptidique liant les résidus amino acides constituant la chaîne protéique. Le suivi de l'absorption à 254 nm est utilisé comme contrôle. En effet, seuls les groupements aromatiques absorbent à cette longueur d'onde. Or le peptide n'en possède pas, donc la variation de cette valeur va nous informer sur la présence d'impuretés dans la solution, comme certains groupements protecteurs.

### ***III. 2. 2. e. Dessalage***

La SM nous a informés sur la présence de sels dans la solution du peptide pur, par des adduits à +23 (ions sodium) et +39 (ions potassium). Or cette contamination doit impérativement être traitée (par dessalage⁹⁸) afin de réaliser ultérieurement des expériences de RMN. En effet, la présence de sels à forte concentration produit une inhomogénéité du champ  $B_1$  ce qui rend le réglage des impulsions difficiles et provoque une baisse en résolution. Une seconde chromatographie utilisant des micro-colonnes de silice greffée par des chaînes de 18 carbones est donc réalisée (Sep-Pack cartridge). Chaque injection ne contient que 2 mg de produit peptidique pour ne pas saturer la colonne. La colonne est solvatée par du méthanol puis rincée avec de l'eau à 1‰ de TFA (5 volumes de colonne). Après dépôt du peptide, une première élution est réalisée avec de l'eau 1‰ de TFA, qui permet d'enlever les ions directement. Puis le décrochage du peptide de la colonne a lieu lors du passage du second éluant consistant en un mélange acétonitrile/H₂O (3/1) à 1‰ de TFA.

### ***III. 2. 2. f. Vérification de la composition en masse par Spectrométrie de Masse***

Le brut de synthèse obtenu est analysé par spectrométrie de masse en mode MALDI-TOF, capable de définir des masses avec une précision allant jusqu'au dalton. La solution issue de la chromatographie est directement déposée sur une matrice, l'acide  $\alpha$ -cyano-4-hydroxy-cinnamique, avec laquelle elle « co-cristallise ». Le laser va ioniser l'échantillon de telle sorte que chaque molécule comportera un proton (masse et charge) en plus. Nous avons ainsi accès au rapport masse sur charge, ce qui nous permet un suivi fin de la masse +1 de tous les composés.

### III. 2. 3. Résultats et Discussion

#### III. 2. 3. a. Caractérisation du peptide IB7₁₄

Le spectre de masse en MALDI-TOF du brut de synthèse est présenté figure III-5.

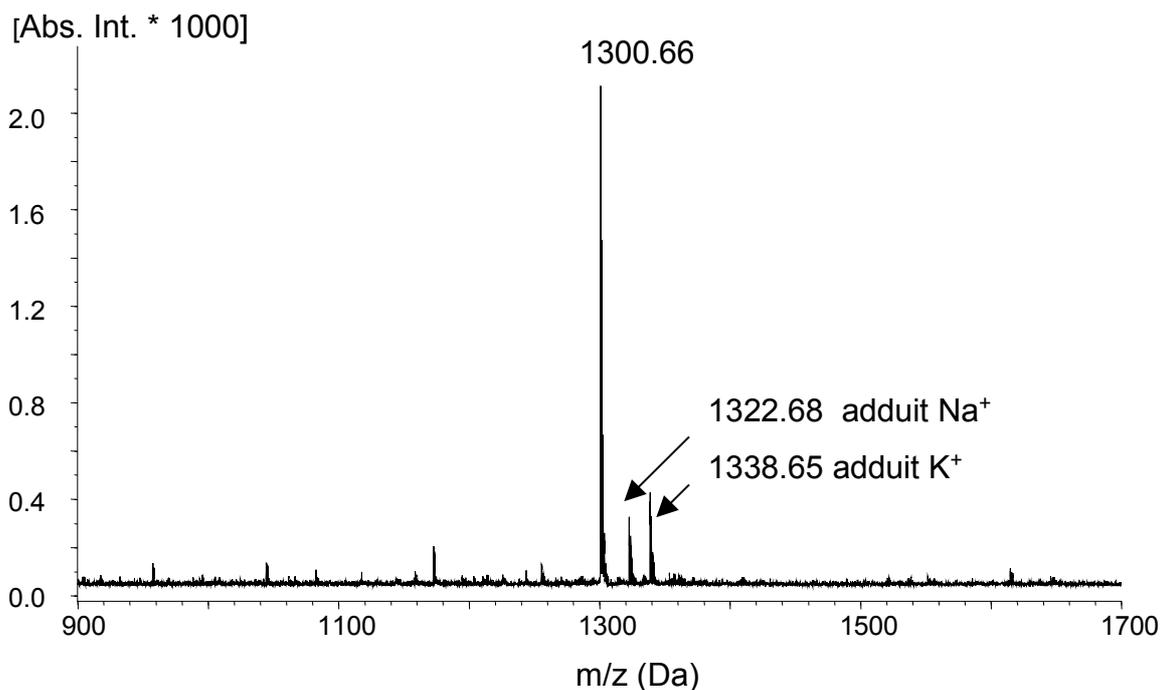


Figure III-5 : Spectre de masse en MALDI-TOF du brut de synthèse du peptide IB7₁₄ après dissolution dans l'eau. Le pic majeur de 1300,66 Da correspond au peptide IB7₁₄ monoprotoné (1+) de masse mono isotopique 1300,66 (masse exacte 1299,66). Les pics secondaires 1322,68 et 1338,65 correspondent respectivement aux masses du peptide lié par du sodium ou du potassium. La présence d'autres pics traduit l'impureté de la solution.

Le pic largement majoritaire d'un rapport masse/charge (m/z) de 1300,66 correspond au peptide IB7₁₄ de masse mono isotopique de 1300,66 soit de masse exacte 1299,66 g.mol⁻¹. Les impuretés présentes dans l'échantillon sont éliminées par RP-HPLC.

La figure III-6 montre les chromatogrammes obtenus avant (en trait plein) et après (en pointillés) la purification. Sur le brut, les pics élués le plus rapidement correspondent aux groupements protecteurs. Le pic étalé reflète un mélange de plusieurs produits peptidiques ayant une composition de nature très proche. Après fractionnement, seul le peptide IB7₁₄ est

présent dans la solution. La pureté de la fraction peptidique apparaît nettement sur le spectre car seul le pic correspondant au temps d'élution du peptide (~11 minutes) est présent.

A l'issue de cette chromatographie, le succès de la purification est confirmé par SM dont le spectre ne contient que la masse du peptide avec ou sans adduits de sels.

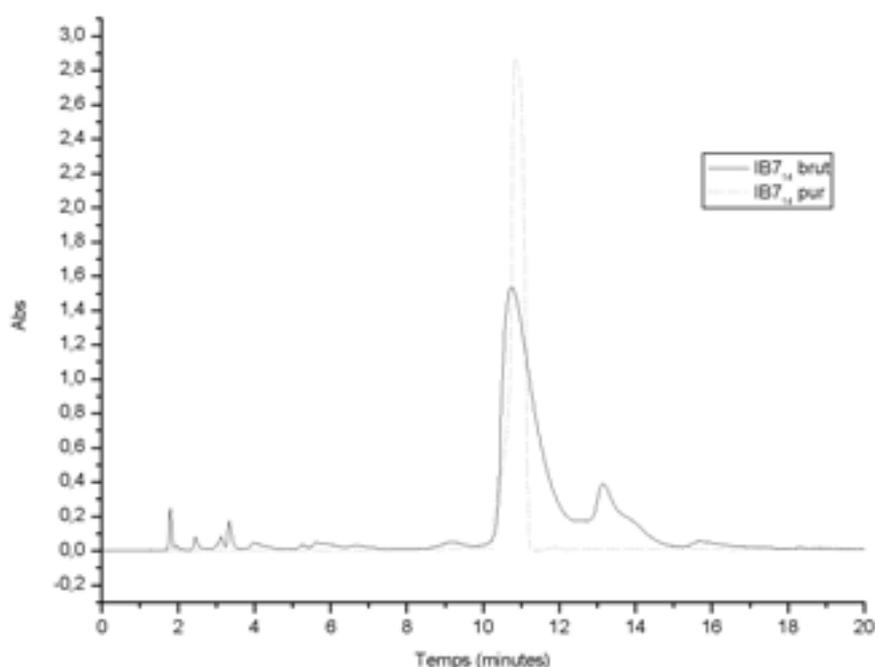


Figure III-6 : Chromatogrammes obtenus après synthèse du peptide IB7₁₄ par RP-HPLC, munie d'une colonne analytique greffée avec des chaînes aliphatiques de 18 carbones, sous un gradient en continu H₂O/Acétonitrile (de 100/0 à 0/100) délivré avec un débit de 1 ml/mn. Le spectre en trait plein représente le brut de synthèse et en pointillé est obtenu après purification.

Le rendement obtenu pour la synthèse du peptide est de 59%, ce qui fait un taux de greffage par acide aminé de  $\sqrt[3]{0,59}$ , soit 96%. Cette valeur est proche du rendement théorique maximal proposé par la technique qui est de 98%. 35 mg de peptide IB7₁₄ pur ont été obtenus.

### III. 2. 3. b. Caractérisation de la protéine IB7

Le spectre de masse en MALDI-TOF du mélange brut obtenu est présenté figure III-7. On observe deux séries de pics majeurs : 5766,8 et 5748,8 correspondant respectivement à la masse  $M+H^+$  et  $M+H^+-18$  d'IB7 ; 2883,7 et 2875,2 correspondant respectivement aux masses  $\frac{M+2H^+}{2}$  et  $\frac{M+2H^+-18}{2}$  d'IB7.

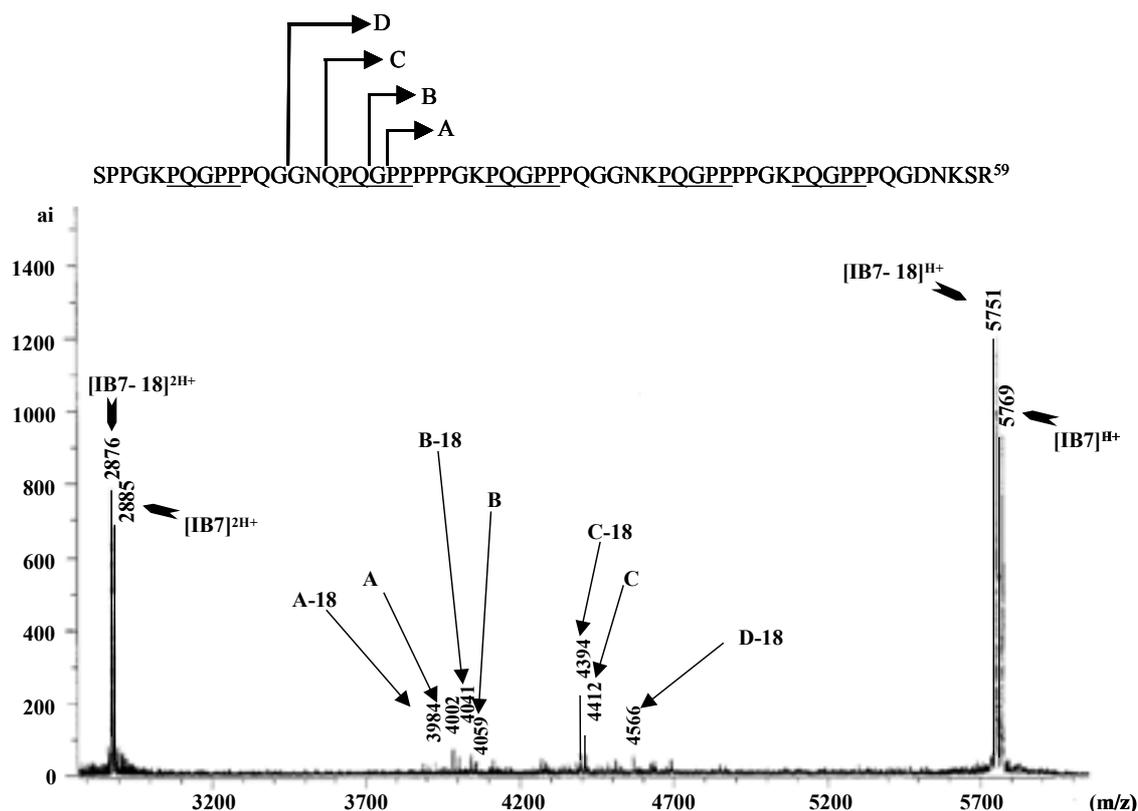
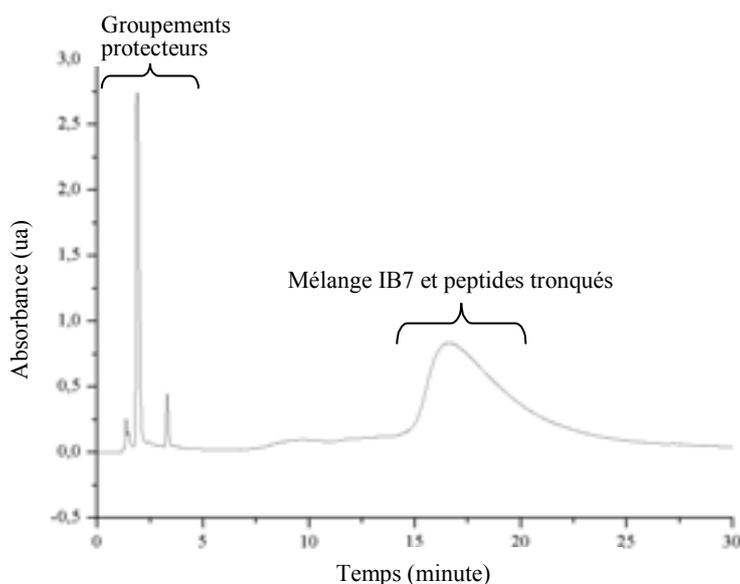


Figure III-7 : Spectre de masse en MALDI-TOF obtenu pour le mélange réactionnel du brut de synthèse après dissolution dans de l'eau. Les pics majeurs correspondent aux masses mono- ( $1^+$ ) et di- ( $2^+$ ) protonées des espèces IB7 et IB7-18. Des pics secondaires représentent des peptides tronqués issus de l'élongation d'après IB7 ou IB7-18. Certains peptides tronqués sont schématisés au-dessus du spectre de masse.

D'autres pics minoritaires sont également observables. Ils correspondent à des peptides tronqués, acétylés lors du capping à l'anhydride acétique. La figure III-7 montre l'attribution de certains d'entre eux, du fragment Arg⁵⁹-Pro²⁰ à Arg⁵⁹-Gln¹⁴ allant respectivement d'un rapport m/z de 4584,3 à 3984,5 Da. L'origine de ces fragments peut être

expliquée par la présence de 5 prolines consécutives en position 22-18. En effet, ces résidus sont capables de se replier pouvant ainsi gêner l'élongation.

La composition en résidus aminés étant la même que celle du peptide IB7₁₄ nous prévoyons un profil d'hydrophobicité quasiment identique, donc la méthode de purification est effectuée selon le protocole RP-HPLC précédemment décrit. Le chromatogramme obtenu ne présente pas une raie fine mais un pic majeur très étalé (figure III-8). Cette allure est due à un mélange de différents peptides dans lesquels on trouve IB7 et un certain nombre de peptides tronqués. Tous ces peptides ont donc la même composition en résidus aminés et présentent donc un même profil d'hydrophobicité. Ils sont ainsi coélus malgré leur différence en taille. Les pics en début de chromatogramme ont été identifiés comme les groupements protecteurs clivés par action du TFA.



**Figure III-8 :** Chromatogramme du brut de synthèse obtenu par HPLC en phase inverse sur une colonne analytique C18. L'élution subit un gradient variant de 18% à 45 % d'éluent B. Les premiers pics correspondent aux groupements protecteurs libérés lors du clivage au TFA et la fraction majeure éluée entre 15 et 20 mn contient un mélange de IB7, IB7-18 ainsi que les peptides tronqués.

La purification a été effectuée par un fractionnement fin du pic élué entre 15 et 20 mn. Trois passages successifs en RP-HPLC ont été nécessaires à l'obtention de la protéine IB7 pure. L'avancement de la purification est suivi par la même technique de spectrométrie de

masse MALDI-TOF que précédemment. La figure III-9 montre un spectre de masse de la fraction purifiée de la protéine IB7.

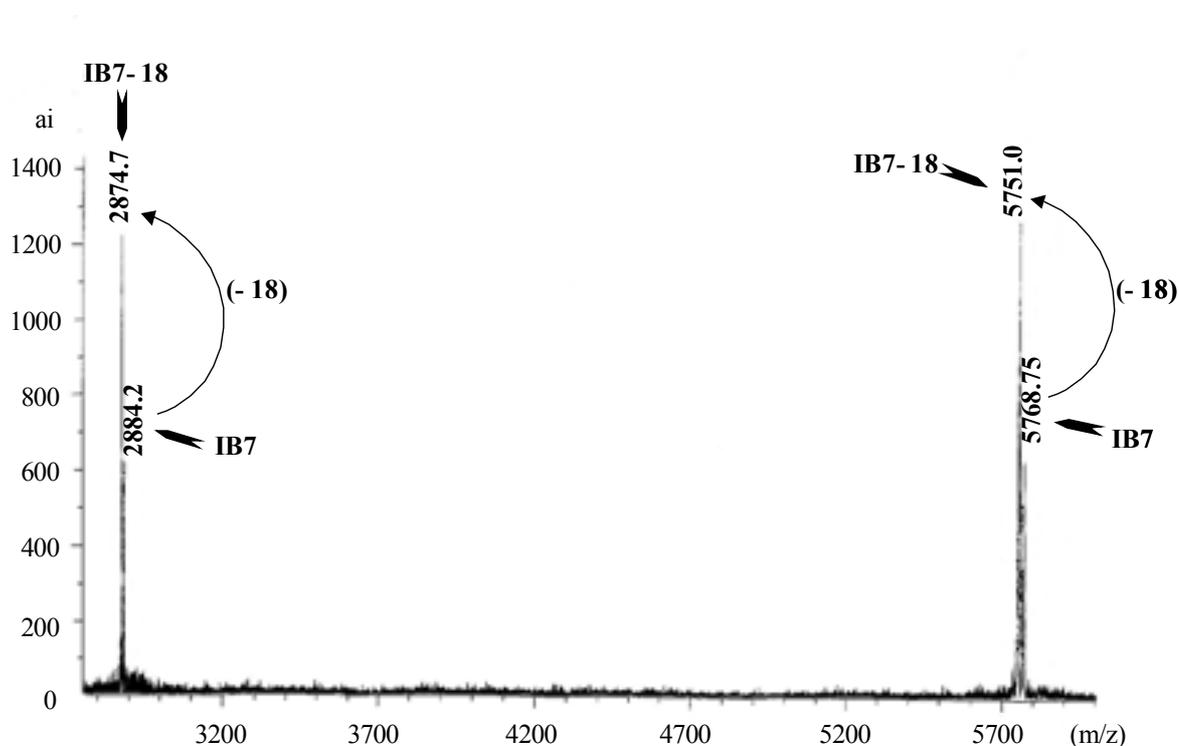


Figure III-9 : Spectre de masse en MALDI-TOF de la fraction purifiée par RP-HPLC. Les deux espèces présentes sont les peptides IB7 et IB7-18 mono- et di-protonées.

Plus aucun pic n'est présent dans la région 4000 à 5000, suggérant que tous les peptides tronqués ont été éliminés. En revanche les deux pics correspondant à la masse de la protéine IB7 M ou M-18, de formes mono ( $1H^+$ ) et di-protonées ( $2H^+$ ) sont toujours présentes. Nous avons donc un mélange de deux protéines, inséparables par HPLC, la protéine IB7 et une protéine dont la masse correspond à la protéine IB7 moins 18 Da.

Un problème majeur rencontré lors de la synthèse chimique de protéines comprenant des résidus acides aspartiques est la formation d'aspartimide^{143,144}. La figure III-10 nous montre une réaction classique basée sur la réactivité de l'aspartate. Plusieurs facteurs peuvent entraîner ce type de réaction comme la nature des groupements protecteurs ou l'utilisation de solvants acides en présence de séquences Asp-Gly ou Asp-Asn. La formation d'aspartimide catalysée en milieu acide a lieu indifféremment sur des résidus aspartates libres ou protégés



### **III. 2. 4. Conclusion de la synthèse**

Après optimisation des conditions de SPPS, les synthèses d'IB7 et IB7₁₄ ont été faites avec des rendements tout à fait acceptables (avec des taux de greffage par résidu de 96 et 97%). Les difficultés inhérentes à la composition en résidus des PRPs (majorité de prolines et glycines) ont ainsi pu être contournées. Dès lors que la synthèse d'IB7 est au point, nous pourrions tenter d'accéder à d'autres types de protéines riches en proline, comme les protéines du collage.

L'obtention de quantités suffisantes d'IB7 et IB7₁₄ permettent d'envisager leur étude structurale ainsi que l'étude des interactions protéines-tanins.

Cependant, la protéine IB7 n'a pas pu être séparée de son homologue IB7-18. En revanche, étant donné que l'acide aspartique est en position 55 sur la protéine IB7, soit à la fin de la séquence, la formation d'aspartimide est supposée n'avoir que peu d'influence sur la structure de la protéine et ne devrait donc pas contrarier les études physico-chimiques réalisées par la suite.

### **III. 3. LES TANINS DIMERIQUES**

Les tanins modèles que nous avons choisi appartiennent à la famille des tanins condensés, tanins présents dans les parties solides des raisins (pépins et pellicule) et transmis au vin au cours de la vinification.

Deux méthodes d'obtention des dimères ont été développées : une méthode d'extraction directe à partir de pépins de raisin (Merlot) et une méthode de synthèse. Ces deux approches nous permettent d'avoir accès à différents dimères présents dans les parties dures des raisins. L'extraction à partir de pépins de raisin nous a permis d'isoler les dimères majoritaires présents dans le raisin et la synthèse par voie cationique, décrite par Barathieu⁶³, nous a permis d'accéder au dimère B3 (catéchine-(4 $\beta$ -8)-catéchine).

### **III. 3. 1. Extraction des pépins de raisin**

L'extraction des composés directement sur la matière vivante (raisin) ou exploitée (vin) est une méthode d'obtention des composés à l'état natif. Cependant la composition phénolique des raisins comme celle du vin est complexe et nécessite un fractionnement délicat effectué en plusieurs étapes : au préalable, les tanins oligomériques sont extraits, puis les dimères sont séparés des autres constituants, enfin chaque dimère peut être purifié séparément.

Ce travail a été effectué dans le laboratoire d'Yves Glories sous la tutelle de Cédric Saucier à la Faculté D'œnologie (Université Bordeaux 2), en nous basant sur les travaux de Victor de Freitas¹⁴⁵ et de Marie Mirabel¹⁴⁶.

#### ***III. 3. 1. a. Matériels et Méthodes de séparation***

##### **Isolement des oligomères et polymères tanniques à partir de pépins de raisin (Merlot)**

50 g de pépins de raisin (issus du cépage Merlot) sont préalablement broyés, lyophilisés puis réduits en poudre. La fraction tannique totale est extraite par deux bains successifs : 1,5 L d'un mélange méthanol/eau (80/20, v/v), puis 1,5 L d'un mélange acétone/eau (75/25, v/v). Une agitation magnétique est réalisée sous azote durant trois heures pour chaque bain. Les solvants sont ensuite évaporés à sec et le produit est repris dans 300 mL d'eau puis lyophilisé.

Sur 100 mg de tanins totaux, une extraction au méthanol/chloroforme (1/3, v/v) permet de séparer les tanins polymériques des tanins oligomériques. En effet, seuls les oligomères restent solubles dans ce mélange alors que les polymères y précipitent, ainsi une simple filtration suffit à les séparer. Le surnageant est ainsi gardé puis évaporé. L'extrait sec (50 mg) est repris dans 2 ml de méthanol.

##### **Extraction des dimères par chromatographie liquide basse pression**

La fraction de tanins oligomériques obtenus précédemment contient différents monomères, dimères, trimères, voire tétramères. Le pré fractionnement est réalisé sur un circuit de chromatographie basse pression dans une colonne de 35 cm de hauteur et de 1,6 cm de diamètre, dont la phase mobile est le méthanol délivré avec un débit de 0,8 ml/mn. La

phase stationnaire est un polymère réticulé qui se présente sous forme de billes poreuses de 5 à 10 mm de diamètre (gel Toyopearl TSK HW40) et qui, hydraté, prend l'aspect d'un gel. Le choix de la porosité se fait en fonction de la taille des molécules à séparer. En effet, la porosité influence directement la migration des molécules de telle sorte que plus la molécule a une taille élevée plus elle parcourt un chemin long au sein des billes. Ainsi, les tanins les plus polymérisés sont retenus plus longtemps dans le gel, au contraire les tanins de faible masse moléculaire comme la catéchine seront élués en premier.

La détection des composés élués est effectuée par spectrophotométrie à 280 nm, spécifique aux groupements phénoliques.

### **La purification par chromatographie liquide haute performance en phase inverse (RP-HPLC)**

20 mg de dimères ont été séparés de la fraction oligomérique. La séparation des différents isomères a été réalisée par chromatographie liquide haute performance en phase inverse (RP-HPLC), cette technique permettant de séparer les molécules en fonction de leur polarité. La phase stationnaire est constituée de billes de silice poreuses sur lesquelles des chaînes aliphatiques de 18 carbones sont greffées, (colonne Beckman ultrasphère ODS C18, 5 mm 100 Å, 15 x 300 mm). L'éluion est réalisée par un solvant constitué par de l'eau à 2,5% d'acide acétique, solvant A, et un solvant B constitué par 20% d'acétonitrile dans 80% de A. La séparation des différents dimères est obtenue pour un gradient allant de 95% à 80% de A avec un débit de 3 ml/mn. Les molécules sont alors éluées de la plus polaire à la moins polaire et leur sortie de colonne est visualisée par absorption UV à 280 nm.

Une chromatographie sur couche mince est réalisée sur chaque fraction pour en déterminer la composition. Les dépôts sont effectués parallèlement à la catéchine utilisée comme valeur contrôle du temps de rétention. L'éluant est un mélange d'acétate d'éthyle/acide formique (48/2, v/v). La migration des spots est suivie soit sous lampe UV soit sous vapeur d'iode. Les différents dimères sont caractérisés grâce à leur vitesse de migration caractéristique¹⁴⁵. Elle est évaluée par la distance de migration du soluté par rapport à la distance de migration du solvant, noté  $R_f$ .

### III. 3. 1. b. Résultats et discussion

La figure III-11 présente le chromatogramme de la fraction de tanins oligomériques. Les molécules de faible masse moléculaire sont éluées en premier, puis viennent les unités monomériques, dimériques et enfin trimériques ainsi que le restant des oligomères.

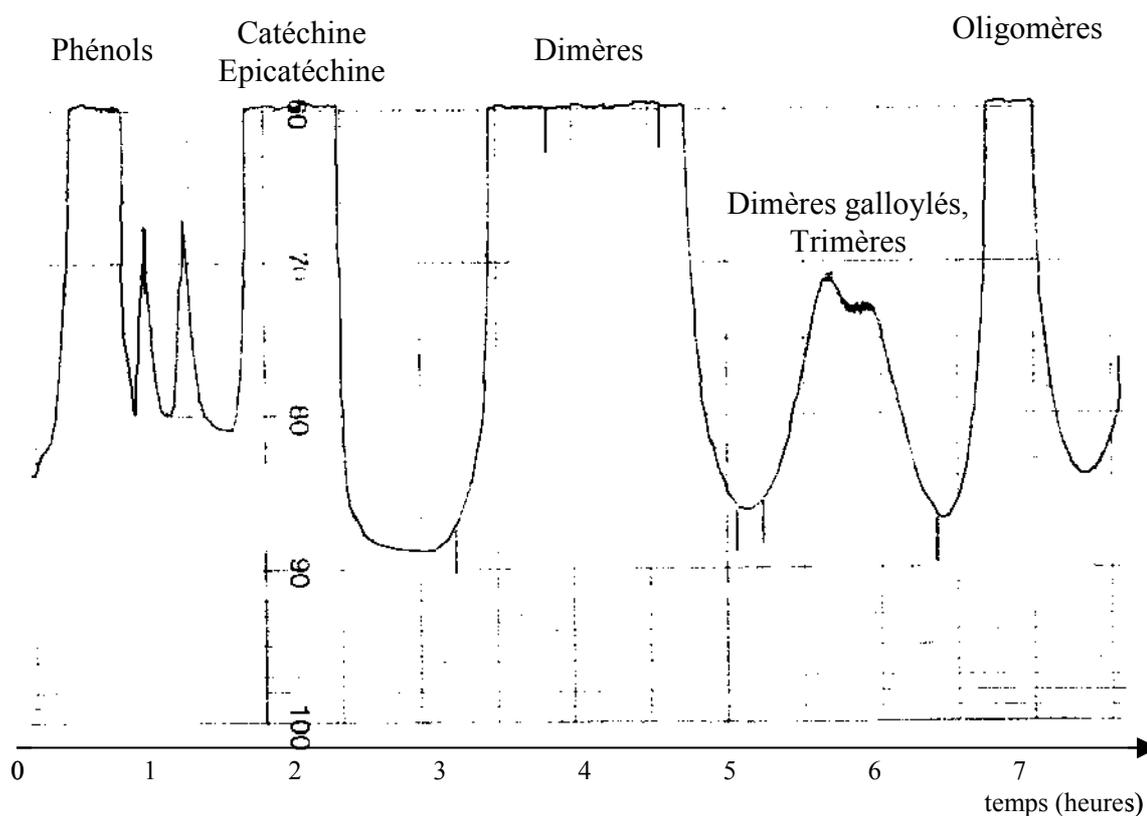


Figure III-11 : Chromatogramme obtenu par détection à 280 nm après fractionnement du brut de tanins oligomériques par chromatographie basse pression sur un gel TSK HW40S. L'éluant est le méthanol délivré avec un débit de 0,8 ml/mn.

La fraction issue de la chromatographie contenant les dimères est donc prélevée de façon à pouvoir purifier les différents isomères. Cette fraction est la plus représentée par rapport à la masse totale des tanins dans les pépins de raisin.

Le profil chromatographique de la fraction dimérique est présenté figure III-12.

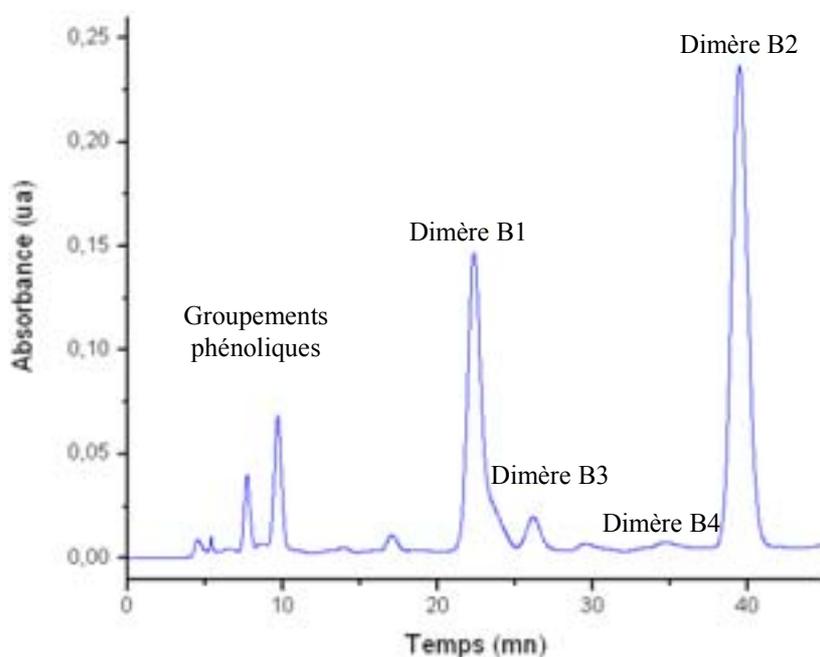


Figure III-12 : Chromatogramme obtenu par RP-HPLC. Les différents dimères B3, B1 et B2 sont caractérisés d'après les temps d'élution décrits par De Freitas¹⁴⁵.

La caractérisation des différents dimères, vus chapitre I, peut être réalisée par la mesure des temps d'élution en HPLC ou bien par la mesure de la distance de migration du soluté par rapport au solvant,  $R_f$ . En effet, chaque dimère a un  $R_f$  qui lui est propre : les dimères B1 et B2 ont un  $R_f$  de 0,56, les dimères B3 et B4 ont un  $R_f = 0,60$ . Il y a peu voire pas de dimères B5, B7 ( $R_f = 0,70$ ) ni de dimères B6 et B8 ( $R_f = 0,83$ ). Le trimère C1 est également présent dans cette fraction avec un  $R_f$  de 0,40¹⁴⁵.

Cette approche nous permet d'isoler plus facilement le dimère B2 que les autres dimères. Non seulement, le dimère B2 est quantitativement le plus important donc non seulement il pourra être représentatif de la population de tanins mais une raison pratique s'ajoute à cela dans la mesure où l'extraction nécessite moins de matériel de départ afin d'obtenir davantage de produit pur. Mais d'un point de vue pratique, il est plus éloigné des autres dimères alors que B1 et B3 risquent d'être co-élus.

Environ 1 mg de dimère B2 a pu ainsi être purifié.

Cette méthode d'extraction présente l'avantage de pouvoir obtenir différents types de dimères, issus du raisin et du vin. Cependant elle reste lourde et les quantités de produits obtenus très insuffisantes pour les études ultérieures d'interaction avec les protéines. Une autre voie a donc été abordée dans le but de disposer de quantité de matériel tannique convenable pour les études physico-chimiques, la synthèse.

### III. 3. 2. Synthèse par couplage cationique

La méthode de synthèse utilisée au laboratoire est le couplage cationique, développé au Laboratoire de Chimie Organique et Organométallique (LCOO-Université Bordeaux 1, dans le groupe d'E. Fouquet par Karine Barathieu⁶³). Cette méthode a été mise au point pour la synthèse du dimère B3 (catéchine-(4 $\beta$ -8)-catéchine). La liaison C4-C8 a été choisie puisque les dimères les plus représentés dans le vin possèdent ce type de liaison (cf chromatogramme figure III-12). D'autre part, ce choix s'est confirmé pour une question de coût : en effet, le gramme de catéchine (1) est vendu 14,5 € alors que le gramme d'épicatéchine est à 68,7 € (Sigma-Aldrich, France).

#### III. 3. 2. a. Etapes de la synthèse

Les différentes étapes de la synthèse par couplage cationique, résumées figure III-13, sont les suivantes :

**a-** la protection des groupements OH de la catéchine (1).

Cette étape est nécessaire à cause de l'instabilité de la catéchine en dehors de la zone de pH 5 à 8¹⁴⁷. Cette protection est assurée par des groupements benzyles, groupements stables en milieu acide et basique et présentant l'avantage d'offrir une déprotection douce en milieu neutre par une réaction d'hydrogénation, compatible avec la fragilité de nos composés. Après protection, on obtient la catéchine tétrabenzylée.

**b-** l'activation de la position 4 du cycle pyrane de la catéchine tétrabenzylée a été développée au LCOO en utilisant la méthode de Tückmantel *et al.*¹⁴⁸. La catéchine tétrabenzylée (2) activée sert d'unité d'extension.

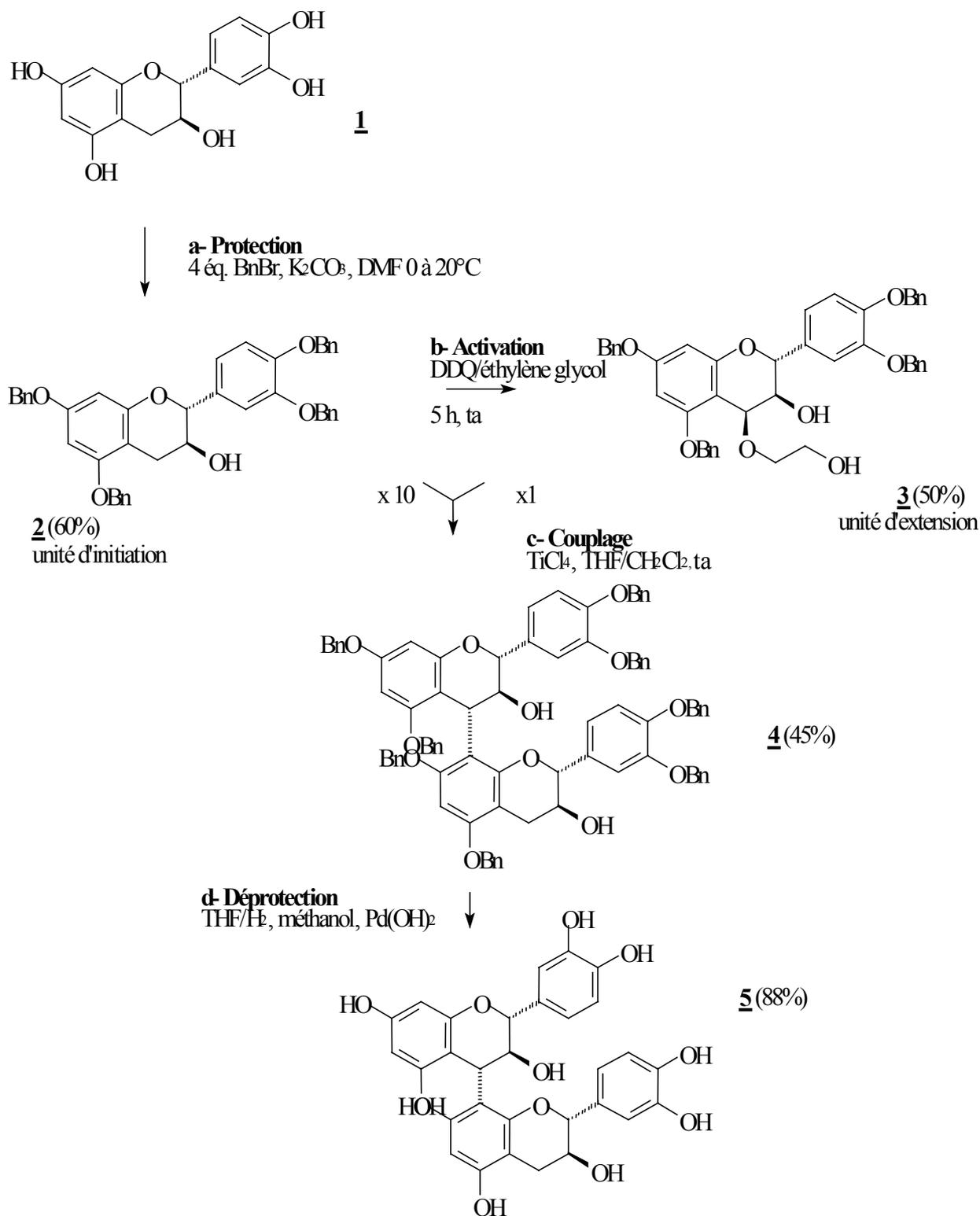


Figure III-13 : Schéma de synthèse du dimère B3 par la voie du couplage cationique⁶³.

c- le couplage va permettre de créer la liaison C₄-C₈ entre la catéchine activée tétrabenzylée (comme unité d'extension, 3) et la catéchine tétrabenzylée (comme unité d'initiation, 2), cette dernière étant en excès (1 unité d'extension pour 10 unités d'initiation). Le dimère catéchine-(4β-8)-catéchine octabenzylé (4) est ainsi formé.

d- la déprotection des groupements benzyles se fait par hydrogénation dans un mélange méthanol/THF sous H₂ avec le catalyseur de Pearlman (Pd(OH)₂) sur charbon, pour obtenir le dimère B3 (5).

### III. 3. 2. b. Matériels et Méthodes

La synthèse est la caractérisation ont été réalisées selon le protocole décrit par K. Barathieu⁶³.

Une purification par RP-HPLC est nécessaire à l'obtention du dimère B3 pur, sur une colonne greffée de chaînes aliphatiques de 18 carbones. La méthode d'élution utilisée est la même que pour l'extraction, avec un gradient allant de 96 à 80% de A en 25 mn.

### III. 3. 2. c. Synthèse du dimère B3

Le dimère B3 a été obtenu avec un rendement de 48%, soit 20 mg. Son identification est réalisée par RMN (figure III-14).

L'attribution totale du B3 a été réalisée et montrée dans le tableau III-1 suivant :

A, B, C		D, E, F	
¹ H	δ (ppm)	¹ H	δ (ppm)
2	4,438	2	4,626
3	4,310	3	3,978
4	4,374	4	2,535
-		4'	2,888
6	6,018	6	6,175
8	5,732	-	
2'	6,854	2'	6,904
5'	6,618	5'	6,836
6'	6,481	6'	6,660

Tableau III-1 : Attribution des protons du B3 par RMN.

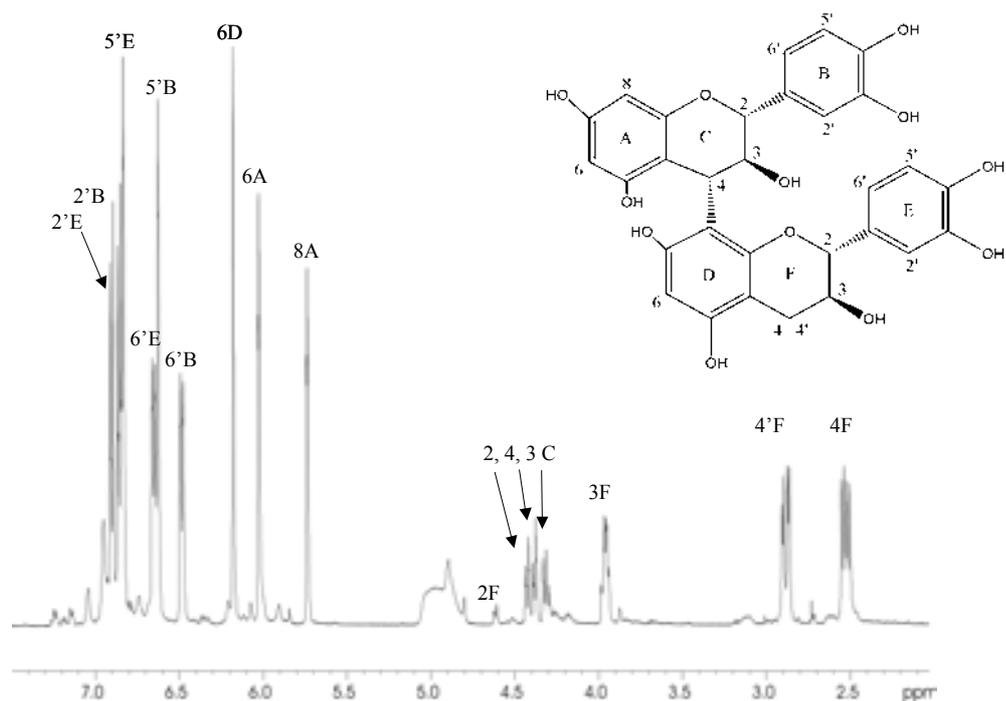


Figure III-14 : Spectre en RMN  $^1\text{H}$  dans  $\text{D}_2\text{O}$  du dimère B3 représenté, obtenu par synthèse cationique et purifié par RP-HPLC.

La régiochimie (liaison C4-C8) a été démontrée par RMN hétéronucléaire HMBC ( $^1\text{H} - ^{13}\text{C}$ ) dont les corrélations principales sont reprises dans la figure III-15 :

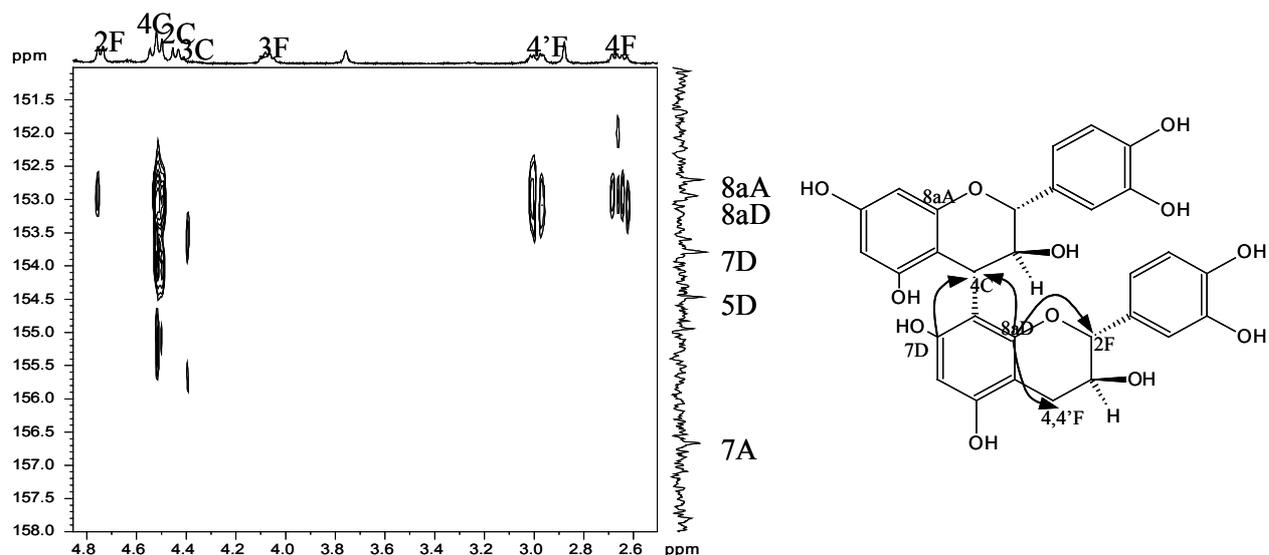


Figure III-15 : HMBC du dimère B3 dans l'éthanol : la région observée démontre les corrélations mettant en évidence la régiochimie du composé.

Les corrélations entre le H₄C et les carbones C₇D et C_{8a}A permettent de connaître la régiochimie 4-8 de B3.

La structure tridimensionnelle adoptée par les tanins a été largement étudiée en dynamique moléculaire où la méthode Monte-Carlo a fait ses preuves. Ainsi nous retrouvons la structure du B3 ayant une forme compacte comme présentée figure III-16 :

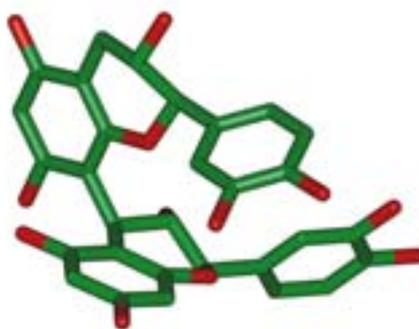


Figure III-16 : Conformation tridimensionnelle adoptée par le dimère B3, obtenue par dynamique moléculaire Monte-Carlo.

### **III. 3. 3. Choix de la méthode retenue**

L'extraction de tanins directement à partir du raisin permet la purification de tous types de dimères ou trimère tel que le C1. Cependant, seulement 1 mg de B2 ont pu être extraits. En revanche, la synthèse a produit 20 mg de dimère B3.

La synthèse par couplage cationique constitue donc une méthode de choix pour l'obtention de dimères. En choisissant l'unité d'extension ainsi que l'unité d'initiation, il est possible d'obtenir sélectivement tous les dimères en liaison C4-C8 (de B1 à B4). D'autre part, cette approche permet également de synthétiser un trimère (C1), et semble généralisable aux autres trimères, voire aux tétramères et pentamères.

### **III. 4. CONCLUSION DES SYNTHÈSES DE PROTÉINES ET DE TANINS**

Les synthèses de peptides et protéines sur support solide ainsi que de tanins par couplage cationique ont ainsi conduit à l'obtention des acteurs de l'astringence dans des quantités exploitables pour l'étude structurale et pour l'étude physico-chimique du complexe. Le développement de ces méthodes et leur optimisation nous permettent désormais de synthétiser rapidement les produits et ce de manière quasi-routinière.

A terme, d'autres types de protéines riches en proline ainsi que d'autres types de tanins pourront être synthétisés dans le but de connaître l'influence de la nature du tanin et de la protéine face à leur capacité d'interagir.



**Chapitre IV DETERMINATION**  
**STRUCTURALE DE**  
**IB7 ET IB7₁₄**



Les différentes techniques pouvant amener à la caractérisation de la structure tridimensionnelle de protéines sont les rayons X, la diffraction des neutrons, la microscopie électronique, la RMN. Cependant, les protéines riches en proline se prêtent peu à ces types d'analyse du fait de leur composition particulière. En effet, elles sont formées à 70-80 % de proline, glycine et glutamine, organisées en une série de séquences répétées de 5 à 9 acides aminés³³. A notre connaissance, aucun cristal de PRP n'a été obtenu, rendant inaccessible l'analyse de la structure 3D par rayons X. De même, aucune structure n'a pu être déterminée par RMN donc il n'existe pas de banques de données concernant ce type de protéines, ne permettant pas une prédiction de structure par homologie de séquence. L'analyse structurale par RMN de protéines riches en proline est également délicate du fait de leur composition : il apparaît un chevauchement important des résonances lié aux séquences répétées avec principalement 3 acides aminés et à la présence d'un fort taux de prolines rendant plus difficile l'attribution séquentielle du fait de leur amine tertiaire.

## IV. 1. INTRODUCTION

### IV. 1. 1. Structures de PRPs : pelote ou hélice de type II ?

A ce jour, les études structurales sur les protéines riches en prolines entières sont principalement basées sur des analyses de structures secondaires obtenues via le dichroïsme circulaire (DC) ou encore par l'analyse de profil d'hydrophilicité / hydrophobicité basée sur le potentiel d'hydrophobicité de chaque acide aminé. Seules les attributions de l'ensemble des protons de deux fragments de 19 et 22 résidus contenant une séquence répétée de la PRP MP5 de souris, solubilisés dans un mélange H₂O/DMSO (80/20, v/v) ont été déterminées par RMN¹²⁷ : ces auteurs n'ont observé que des effets NOEs séquentiels et des constantes de couplages NH-H $\alpha$  oscillant entre 5 et 8 Hz. Ces résultats leur permettent de conclure que ces peptides présentent une conformation en pelote statistique et/ou en hélice de type II. En revanche, les structures étendues comme l'hélice de type II, typique d'enchaînements poly(L-proline) ou la pelote statistique engendrent une perte de corrélation ( $i, i+3$ ), ce qui ne permet pas d'observer directement le pas de l'hélice. Des études de DC en absence ou présence de CaCl₂ les conduisent à penser que leurs peptides adoptent préférentiellement une

conformation aléatoire et ouverte. Conclusions en accord avec d'autres équipes utilisant la même propriété des ions calcium connus pour dénaturer les hélices  $\alpha$ ^{33,149}. En revanche, d'autres équipes observent une structure étalée des PRPs basiques en pelote statistique, entremêlées d'hélices de type II^{55,82}. Cid *et al.*⁸², en se basant sur la nature de la chaîne peptidique (séquence ou profil d'hydrophobicité), concluent que les PRPs ont une conformation aléatoire pouvant également être structurées en hélice de type II typique de la poly(L-proline), ce qui implique que la séquence seule contient l'information nécessaire à l'organisation spatiale d'une chaîne peptidique. En comparant les structures de cristaux obtenus de penta-peptides riches en prolines synthétiques, réalisés par Traub *et al.*^{150,151}, dont l'analyse aux rayons X montre la présence d'hélice de type II, Cid *et al.* suggèrent que les PRPs contenant ce type de motifs doivent également se structurer de la sorte⁸². D'autres études de DC sont en accord avec la présence d'hélice de type II dans les PRPs telles que IB5 et IB9 qui présentent un spectre dichroïque avec un minimum d'intensité comparable à celui obtenu pour la poly(L-proline)^{55,152}.

Malgré la composition comptant pour environ 50% de proline, leur capacité à adopter une conformation en hélice de type PII n'est pas systématiquement reconnue dans la littérature.

## IV. 1. 2. La Proline et l'hélice de type II

La grande particularité de la proline provient de la cyclisation de la chaîne latérale sur la fonction amine ce qui induit une contrainte stérique importante sur l'axe de rotation N-C $\alpha$ . La poly(L-proline) est incapable de prendre une structure secondaire classique dû aux contraintes conformationnelles imposées par ses chaînes latérales pyrrolidine cycliques. Cette nature cyclique lui confère la particularité d'adopter les configurations cis ou trans : dans la configuration cis, l'angle de torsion  $\omega$  étant proche de 0° alors que dans la configuration trans, l'angle  $\omega$  est proche de 180° (figure IV-1). Ces deux configurations différentes de la proline peuvent conduire à deux conformations différentes dans le cas d'une poly(L-proline) : une conformation en hélice de type poly(L-proline) II est adoptée si la proline est de configuration trans et une conformation en hélice de type poly(L-proline) I pour une configuration cis de la proline¹²⁴. Ces deux hélices sont différenciées tant pour leurs propriétés physiques et spectroscopiques que pour leurs structures cristallographiques. L'hélice de type I (PI) est caractérisée par des angles de torsion  $\phi \sim -83^\circ$ ,  $\psi \sim +158^\circ$  et  $\omega$  de  $\sim 0^\circ$ , elle forme une

hélice droite avec 3,3 résidus par tour; l'hélice de type II (PII) est caractérisée par des angles de torsion  $\phi \sim -78^\circ$ ,  $\psi \sim +149^\circ$  et  $\omega$  de  $180^\circ$ , c'est une hélice gauche avec 3 résidus par tour. L'hélice de type PII est beaucoup plus étendue que la PI¹⁵³ ou que l'hélice  $\alpha$  plus classique (pas de l'hélice  $\alpha$  ou PI  $\sim 5 \text{ \AA}$ , pas de l'hélice PII  $\sim 8 \text{ \AA}$ ).

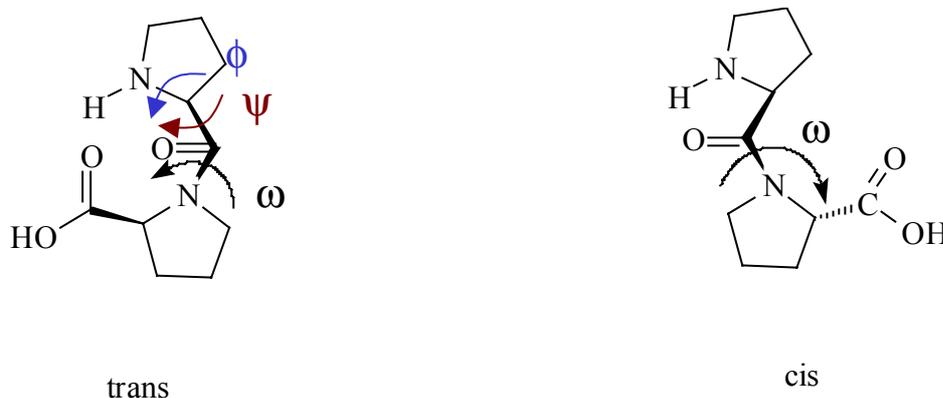


Figure IV-1 : Dipeptide Pro-Pro dans le cadre d'une hélice de PII, en trans, et dans le cadre d'une hélice PI, en cis. Les angles  $\omega$ ,  $\psi$  et  $\phi$  sont représentés.

La forme de l'hélice adoptée dépend de la nature du solvant : la formation d'hélice de type PI est favorisée dans un milieu apolaire, alors que la formation de PII est favorisée dans des milieux polaires¹²⁴. La forme PII est donc la plus représentée dans des conditions biologiques, chez les plantes ou les animaux. En effet, les motifs en hélice de type poly(L-proline) retrouvés dans les PRPs sont tous de type PII. Les angles  $\phi$  et  $\psi$  ont été calculés à plusieurs reprises pour différentes protéines et leurs valeurs fluctuent très peu de  $-75^\circ$  à  $-78^\circ$  et  $+145^\circ$  à  $+149^\circ$  respectivement^{124,154-156}. Les prolines adoptent les conformations cis et trans et influent également sur les conformations cis et trans des résidus précédents tout comme des résidus suivants dans une chaîne peptidique¹⁵⁵. De plus, en condition naturelle, certaines protéines peuvent contenir jusqu'à 45% d'hydroxyprolines, favorisant la conformation trans des prolines¹⁵⁷.

L'hélice de type poly(L-proline) est induite par la nature rigide et cyclique de la proline. L'absence d'un substituant hydrogène sur l'azote de son squelette rend impossible toute stabilisation par liaison hydrogène. La poly(L-proline) forme donc une hélice gauche de 3,3 résidus par tour d'hélice et un pas de 9 Å, comme représentée figure IV-2.

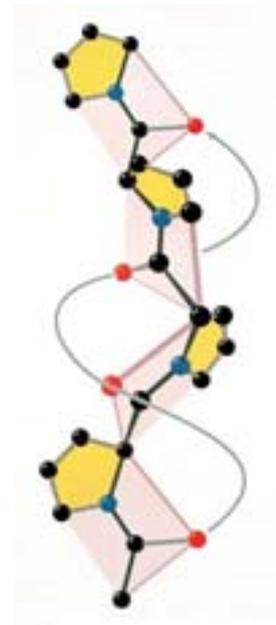


Figure IV-2 : Représentation de l'hélice de type II pour une poly(L-proline).

Si l'on considère un enchaînement de résidus peptidique de type poly-proline, seulement certains couples de valeurs d'angles  $\phi$  et  $\psi$  peuvent exister, car toutes les orientations spatiales ne sont pas admises. Le diagramme de Ramachandran précise les couples de valeurs observées pour les protéines. Le diagramme de Ramachandran d'une poly(L-proline), figure IV-3 montre les valeurs d'angles possibles.

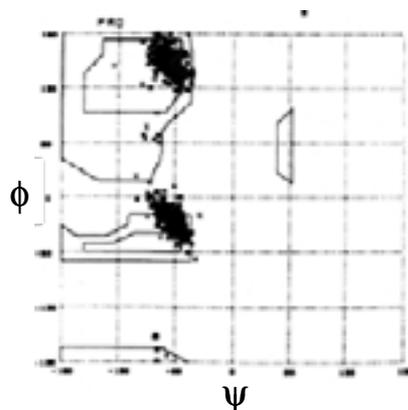


Figure IV-3 : Diagramme de Ramachandran d'un polypeptide poly(L-proline). Les couples de valeurs  $\phi$  et  $\psi$  sont très spécifiques dus à la conformation propre du résidu proline.

### IV. 1. 3. Critères déterminants la formation d'hélice de type II

La structure de type poly(L-proline) trouvée dans les PRPs basiques est due au taux de proline dans la structure primaire complète (~50%) ainsi qu'au nombre de prolines consécutives (~30%). L'hélicité de type PII est favorisée par la présence de 4 ou 5 prolines consécutives^{55,127}. Toutefois, il est notable que certaines protéines sans proline peuvent également se structurer de la sorte¹⁵⁸. La conformation trans des prolines est indispensable à la formation d'hélice PII, or toutes les prolines consécutives sont trans alors que les unitaires, isolées, sont en équilibre entre 70% de trans et 30% de cis.

Ces hélices de type poly(L-proline) ne sont donc pas construites par la formation de liaisons hydrogène intra-chaînes comme les autres structures secondaires. Elles ne sont donc pas identifiables par l'utilisation d'agents susceptibles de rompre des liaisons hydrogène comme le CaCl₂. Elles peuvent être donc facilement confondues avec des structures désordonnées ou en pelote statistique¹⁵⁹.

Nous avons vu précédemment que la formation de l'hélice de type PII requiert la présence d'un solvant polaire (aqueux ou acide). D'autres facteurs sont également à prendre en considération comme le pH et la température. En effet, Shibata *et al.*⁵⁵ montrent qu'à pH acide la conformation PII est favorisée par rapport à un pH alcalin. Dans un même but, Fox *et al.*¹⁶⁰ s'aperçoivent que baisser la température induit la formation de l'hélice de type poly(L-proline). Résultat qu'ils expliquent par le fait qu'en baissant la température, l'énergie libre des molécules d'eau diminue et l'ordre dans la solution augmente, d'où l'induction de la conformation.

### IV. 1. 4. Stratégie de l'étude

Dans ce chapitre, une étude structurale par dichroïsme circulaire de la protéine IB7 et du peptide IB7₁₄, partie amino-terminale de IB7 contenant une séquence répétée, est présentée. La mise au point de cette technique nous a permis de mettre en évidence la présence de structure secondaire de type hélice poly(L-proline) PII dans le peptide et la protéine. Une étude par RMN du peptide IB7₁₄ est également présentée. L'étude RMN nous a permis d'établir une liste d'angles dièdres et de distances qui a pu être utilisée pour modéliser le peptide par dynamique moléculaire et obtenir ainsi sa structure 3D.

## IV. 2. MATERIELS ET METHODES

Les matériaux principaux sont les produits de synthèse du chapitre III. L'éthanol a été acheté chez SDS (Peypin, France). La poly(L-lysine) et la poly(L-proline), le CaCl₂ et le TFE ont été achetées chez Sigma (Saint Quentin Falavier, France).

### IV. 2. 1. Dichroïsme Circulaire (DC)

#### Préparation des échantillons

Les cellules de mesure en quartz utilisées ont un trajet optique 0,1 mm, en rapport avec la concentration de chromophore dans la solution (la cuve de 0,1 mm permet une meilleure précision pour des solutions de 200 µM à 2 mM). Différentes concentrations ont été testées entre 200 µM et 2 mM, valeur au-delà de laquelle nous ne pouvons nous placer pour des raisons de saturation de l'appareil.

La protéine IB7 ou le peptide IB7₁₄ sont dissout soit dans de l'eau soit dans un milieu hydroalcoolique H₂O/Ethanol (88/12, v/v), comme celui du vin, l'idéal étant de disposer d'un solvant le plus proche d'un milieu naturel. La protéine IB7 a également été dissoute dans un solvant structurant, H₂O/TFE (70/30, v/v). En outre, les échantillons sont stabilisés à pH 3,5 et à 298 K, conditions plus proches de la salive pour lesquelles l'interaction est connue pour être quasi-optimale.

Les standards poly(L-lysine), poly(L-proline) et autres sont dissout et enregistrés dans les mêmes conditions.

Chaque essai est réalisé en absence ou en présence de 4 mM de CaCl₂.

#### Conditions d'acquisition et de traitement des données

Les spectres de dichroïsme circulaire ont été enregistrés sur un dichrographe Mark VI (Jobin-Yvon) pour des longueurs d'onde dans l'UV lointain, allant de 190 à 350 nm. Seuls les débuts des spectres seront présentés (190-270 nm) car au-delà de 270 nm, aucun renseignement n'est récolté en raison de la nature des acides aminés présents (seuls les résidus tryptophanes absorbent au-delà de 270 nm). L'accumulation de deux scans consécutifs avec un défilement de 0,5 nm.s⁻¹ et un incrément de 0,5 nm, conduit au spectre.

Les spectres résiduels des solvants sont systématiquement soustraits aux spectres des protéines.

La mesure de l'ellipticité suit la relation [2] décrite §-II. 1, avec  $n = 13$  pour IB7₁₄ et  $n = 58$  pour IB7. La méthode de déconvolution, élaborée au laboratoire a été décrite dans le chapitre II.

## IV. 2. 2. Résonance Magnétique Nucléaire (RMN)

### Préparation des échantillons

La protéine IB7 ou le peptide IB7₁₄ sont dissout soit dans de l'eau soit dans un milieu hydroalcoolique H₂O/Ethanol (88/12, v/v), à pH 3,5 et stabilisés à 298 K.

La concentration en peptide et en protéine est fixée à 5 mM, soit respectivement 0,5 mg pour un volume de 80  $\mu$ L correspondant à un rotor de 4 mm ou 17 mg dans un tube de volume 600  $\mu$ L de diamètre 5 mm. 10% de D₂O sont rajoutés pour le signal du lock.

### Conditions d'acquisition et de traitement des données pour le peptide IB7₁₄

La totalité des expériences pour l'attribution des protons du peptide ont été réalisées sur un DSX Avance Bruker 500 MHz muni d'une sonde HR-MAS avec un lock ²H, équipé de gradient z. L'échantillon est placé dans des rotors de 4 mm contenant 80  $\mu$ l de solution. Les spectres sont enregistrés en rotation à l'angle magique à 5000 Hz. L'impulsion de 90° proton est de 5,9  $\mu$ s à 6 dB. Dans chacune des séquences, une méthode de saturation du signal de l'eau a été intégrée (puissance pour la présaturation de 48 dB ; temps et puissance d'impulsion pour le watergate 9,25  $\mu$ s à 9 dB). Le délai de relaxation est fixé à 3 s. Pour toutes les expériences suivantes l'acquisition est réalisée en mode TPPI (Time-Proportional Phase Incrementation)¹⁶¹. La taille de la fenêtre est de 10 ppm dans les deux dimensions (5000 Hz). Les séquences d'impulsion suivantes ont été utilisées :

COSY-DQF¹⁶² : la séquence utilisée est la cosydfprtp19 (Bruker). Pour chaque incrément  $t_1$ , 256 scans sont accumulés.

TOCSY¹⁶³ : la séquence utilisée est la mlevgptp19 (Bruker). Les expériences utilisées ont un spin lock MLEV17, pour un temps de mélange de 300 ms (après comparaison à différents temps de mélange entre 100 et 500 ms). La durée totale du train d'impulsion encadrant le spin lock est de 2,5 ms. Pour chaque incrément  $t_1$ , 128 scans sont accumulés.

ROESY¹⁰⁷ : la séquence utilisée est la roesyprtp19 (Bruker). La durée totale du spinlock est de 300 ms. Pour chaque incrément  $t_1$ , 128 scans sont accumulés.

Le traitement des données est effectué grâce au logiciel GIFA^{164,165}. Pour toutes les cartes, la matrice initiale ( $t_1$ ,  $t_2$ ) de (512x2048) est étendue à (2048x2048) et double transformée de Fourier et apodisation ( $\sin^2$ , 0,2).

Les contraintes de distance obtenues *via* les corrélations croisées en ROESY, et les contraintes d'angles *via* les constantes de couplage  3J , sont déterminées grâce au logiciel XEASY¹⁰⁸. Les calculs préliminaires de structure sont effectués sur le logiciel DYANA¹¹³.

### **Conditions d'acquisition et de traitement des données pour la protéine IB7**

Les expériences présentées ici ont été réalisées au laboratoire ICSN à Gif sur Yvette chez le Pr. Eric Guittet, sur un spectromètre Avance Bruker 800 MHz muni d'une sonde 5 mm triple résonance inversée (bobine 1H/2H interne et bobine 13C/15N externe) équipée de gradients de champ statique à trois axes avec blindage actif (puissance maximale 50G/cm, température de fonctionnement de -50°C à +80°C). La solution est placée dans des tubes de 5 mm contenant 600  $\mu$ L. L'impulsion de 90° proton est de 9,25  $\mu$ s à -4,2 dB. Dans chacune des séquences, une méthode de présaturation du signal de l'eau a été intégrée (puissance 50 dB). Le délai de relaxation est fixé à 1 s. La taille de la fenêtre est de 9,5 ppm dans les deux dimensions (7575 Hz). Seule l'expérience DIPSI est représentée. L'acquisition est réalisée en mode States-TPPI¹⁶⁶. La séquence utilisée est la dipsi2prst (Bruker). Le temps de mélange choisi est de 70 ms. Pour chaque incrément  $t_1$ , 16 scans sont accumulés. La matrice initiale ( $t_1$ ,  $t_2$ ) de (512x4096) est étendue à (2048x4096) après double transformée de Fourier et apodisation ( $\sin^2$ , 0,2).

## **IV. 2. 3. Modélisation Moléculaire (MM)**

### **Fichier de contraintes issues de la RMN**

Dans le but d'obtenir la structure du peptide, la minimisation ainsi que la DM sont effectuées sous contraintes RMN. Le fichier de contraintes est ainsi constitué par les constantes de couplage  3J  mesurées avec la séquence COSY-DQF et les effets NOE issus des expériences de type ROESY. Les contraintes de distances sont proportionnelles à l'intensité des taches de corrélation NOE : très intense pour:  $2,2 \pm 0,4$  Å, moyenne pour  $3,5 \pm 0,9$  Å et faible pour  $5,0 \pm 0,5$  Å (distances choisies pour assurer trois classes contiguës sans

chevauchement entre elles ni espace libre). La méthode d'utilisation de ces contraintes de distances est appelée « flat bottom up well » ou « puit à fond plat » pour laquelle aucune contrainte n'est mesurée lorsque les atomes s'orientent suivant l'espace autorisé préalablement (par exemple de 1,8 à 2,6 Å pour la région intense). En revanche, dès que cet espace est franchi les contraintes s'additionne.

Les simulations ont été effectuées selon une variante de dynamique moléculaire sur MacroModel, la dynamique stochastique. Aux forces propres du champ de forces sont additionnés des forces de friction et de mouvement brownien, ce qui simule certaines propriétés du solvant¹⁶⁷.

### **Mode de calcul**

Les calculs ont été effectués sur une station SGI Octane R10K avec le logiciel MacroModel¹⁶⁸ version 6,5 (Schrödinger Inc.). Les minima conformationnels ont été localisés avec le champ de force AMBER* modifié (1986 paramètres) tel qu'il est fourni avec le logiciel. Les structures sont minimisées jusqu'à l'obtention d'un critère de convergence RMS inférieur à 0,005 kJ.Å⁻¹.mol⁻¹ via la méthode TNCG (Truncated Newton Conjugate Gradient) (1000 cycles). Dans tous les cas nous avons utilisé la méthode de solvation virtuelle par continuum de solvation GB/SA¹⁶⁹. Le solvant implicite étant l'eau. Dans ce travail, l'option « extended cut-off » a été utilisée, dont les valeurs sont : Van der Waals = 8 Å, électrostatique = 20 Å et liaisons hydrogène = 4 Å.

Le peptide brut a été soumis à trois dynamiques successives de 2 ns chacune à 300, 500 et 1000 K respectivement, ceci permettant la rupture des liaisons aléatoires. Le temps d'intégration est de 1,5 fs dans le premier cas et 1,0 fs dans les deux derniers. A chacune des températures 200 conformations ont été conservées, l'ensemble des 600 solutions a été minimisé sous contraintes et classées par ordre croissant d'énergie requise (TNCG, 1000 étapes). Seulement les 20 structures de plus basse énergie ont été prises en considération.

## IV. 3. RESULTATS :

### IV. 3. 1. Structure secondaire de IB7 et IB7₁₄ par DC

Le dichroïsme circulaire constitue une étape préliminaire pour l'étude de la structure secondaire de protéines. Le spectre obtenu donne des informations sur la proportion de feuillet  $\beta$ , hélice  $\alpha$ , hélice de type II ou pelote statistique (cf § II.1.).

#### *IV. 3. 1. a. Spectres dichroïques de IB7 et IB7₁₄*

Afin de déterminer la structure secondaire de IB7 et IB7₁₄, nous avons enregistré les spectres de la poly(L-proline), modèle d'hélice de type II ainsi que de la poly(L-lysine) qui, à pH 7,4, adopte une structure désordonnée en pelote statistique dans l'eau. Les spectres dichroïques de ces deux polypeptides sont montrés dans la figure IV-2 et présentent des caractéristiques différentes : la pelote statistique est caractérisée par une ellipticité molaire de l'ordre de  $-12000 \text{ deg.dmol}^{-1}.\text{cm}^2$  pour un minimum d'absorption à 196 nm ; alors que l'hélice de type II a un minimum de l'ordre de  $-45000$  à 206 nm.

La figure IV-4 montre également les spectres dichroïques du peptide IB7₁₄ et de la protéine entière IB7 obtenus dans l'eau. Après déconvolution d'après une banque de structures standard connues, on trouve que IB7 et IB7₁₄ sont constitués à 36% de PII et 64% de pelote statistique.

Pelote statistique

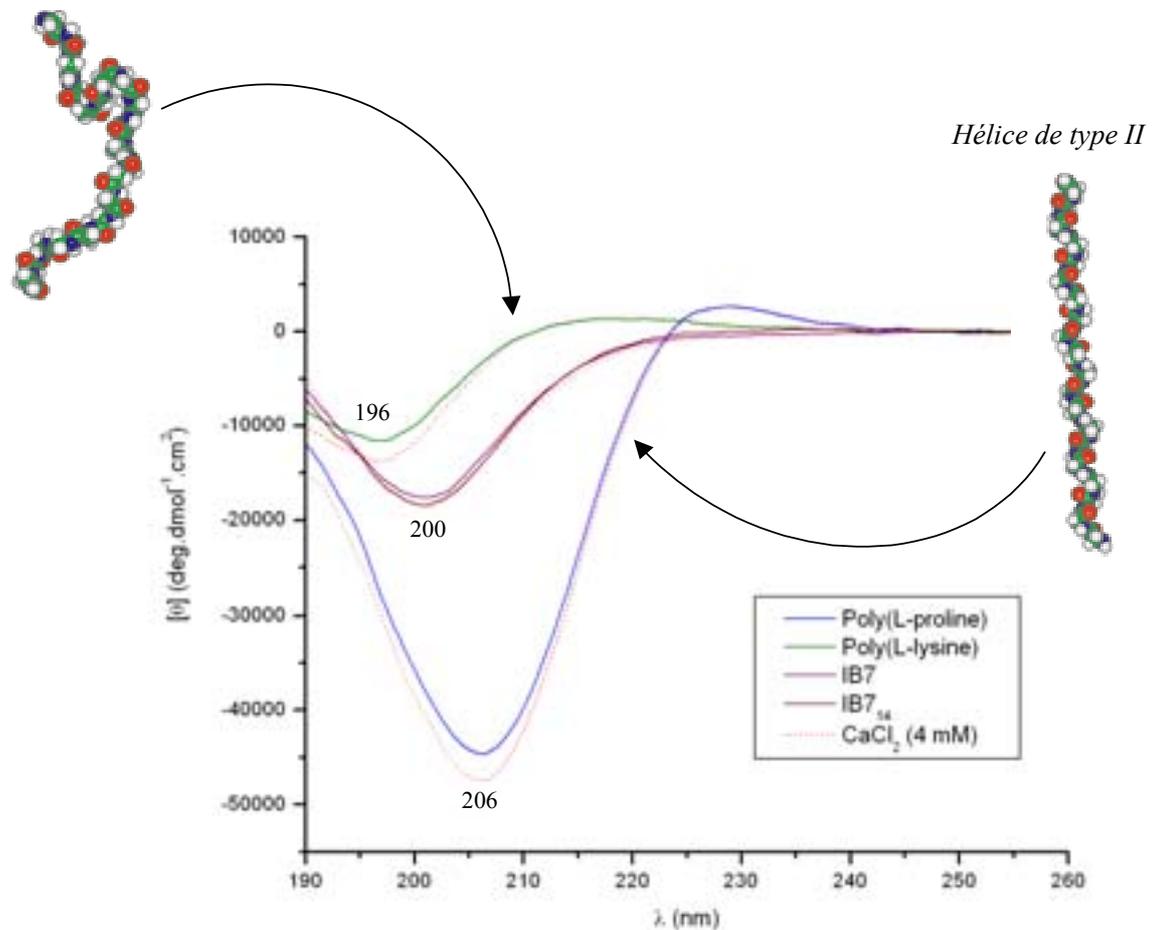


Figure IV-4 : Spectres dichroïques, dans l'eau, de la poly(L-proline) en hélice de type II, de la poly(L-lysine) qui a pH 7,4 adopte une conformation en pelote statistique, ainsi que de la protéine IB7 (à 200  $\mu\text{M}$ ) et du peptide IB7₁₄ (à 200  $\mu\text{M}$ ). Les formes standard sont représentées à côté des spectres.

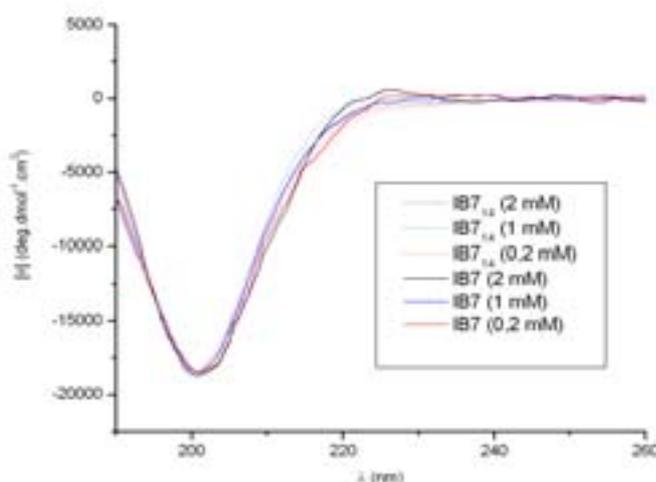
Les mêmes séries d'expériences ont été enregistrées dans un solvant H₂O/EtOH (88/12, v/v) afin de mimer la composition du vin. Les spectres dichroïques présentent alors des profils identiques à ceux obtenus dans l'eau.

Comme prévu, la présence de 4 mM de CaCl₂ dans la solution ne modifie pas l'allure du spectre dichroïque de la poly(L-lysine) en pelote statistique. De même, aucune perturbation n'apparaît sur les spectres d'IB7 et de IB7₁₄ en présence de 4 mM de calcium. Comme l'ont déjà signalé certaines équipes^{127,149}, cette observation induirait que les protéines de type PRP seraient en pelote statistique. Or, le calcium ne perturbe pas la conformation de

la poly(L-proline) en hélice de type II, suggérant que ce type de structure est insensible à l'effet des ions calcium. Aucune conclusion sur la nature des motifs structuraux de ces polypeptides ne peut donc être mise en évidence par l'ajout des ions calcium.

#### ***IV. 3. 1. b. Auto-association des protéines ?***

Afin de mettre en évidence l'influence de la concentration en peptide sur sa propre structure secondaire, nous avons suivi l'évolution des spectres en fonction de différentes concentrations en peptide, de 200  $\mu\text{M}$  à 2 mM. La figure IV-5 rend compte de l'homogénéité des spectres qui ne sont pas modifiés après normalisation par la concentration. Ceci tendrait à prouver qu'il n'y aurait pas de phénomène d'auto-association influençant la structure secondaire des peptides aux concentrations étudiées.



**Figure IV-5 :** Spectres dichroïques IB7 et IB7₁₄ normalisés par la concentration, en solution dans l'eau.

#### ***IV. 3. 1. c. Effet de la présence d'un solvant « structurant » sur la protéine IB7***

La présence de TFE à 30% dans l'eau est connue pour aider à la structuration de protéines¹⁷⁰. Des spectres de DC ont donc été enregistrés dans un mélange TFE/Eau (30/70, v/v) pour la protéine IB7. La figure IV-6 montre une quasi-totale superposition entre les spectres de la protéine IB7 obtenus dans l'eau ou dans l'eau à 30% de TFE. L'utilisation d'un tel solvant ne semble donc pas modifier à la structure secondaire de la protéine IB7.

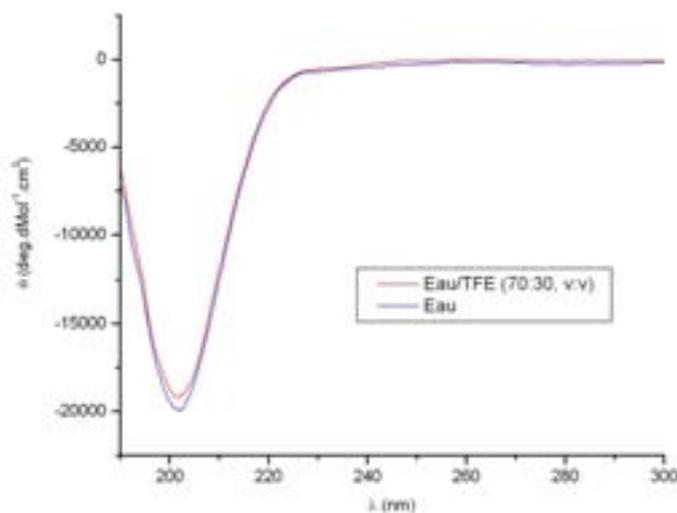


Figure IV-6 : Spectres dichroïques de la protéine IB7 dans de l'eau (bleu) ou dans un mélange eau/TFE (70/30, v/v) (rouge).

#### IV. 3. 1. d. Structure secondaire de IB7 et IB7₁₄

La répartition à environ un tiers (36%) de conformation en hélice de type II et deux tiers (64%) de pelote statistique est en accord avec le nombre de proline dans la séquence du peptide IB7₁₄ et le nombre de prolines consécutives (35% avec deux séries à 2 et 3 prolines consécutives) qui favoriseraient la formation d'hélice de type II. Enfin, les résidus prolines peuvent également influencer la conformation des acides aminés contigus.

Les spectres obtenus pour la protéine entière se superposent parfaitement à ceux du peptide IB7₁₄, représentant un motif répété trois fois dans la protéine. Cette observation laisse penser que la répétition des trois séquences provoque une répartition de conformation secondaire identique à celle du peptide, mais également que ce peptide sera réellement un bon exemple du comportement de la protéine entière.

Dans la littérature, les études concernant la structure secondaire de PRP amènent à des conclusions diverses. Certains auteurs suggèrent une conformation entièrement aléatoire^{33,127} et d'autres considèrent un mélange de motifs PII entremêlés avec de la pelote désordonnée^{55,82}. Le résultat donné comme preuve de l'absence de structuration en hélice PII provient des expériences de dichroïsme en présence de calcium. En effet, l'ion calcium est

connu pour induire la dénaturation des hélices  $\alpha$  en s'insérant au sein de l'hélice, rompant ainsi les liaisons hydrogène qui la structure. Ainsi le profil dichroïque devrait passer d'une hélice  $\alpha$  à une conformation aléatoire ou bien ne rien changer si le peptide était déjà en conformation aléatoire. Cette dénaturation est directement visualisable en DC par le passage d'une hélice à une pelote. Certains auteurs^{127,149} ont observé l'effet du calcium sur leurs peptides riches en proline : l'absence de modification les a conduits à penser que ces peptides étaient déjà en pelote statistique. Pour vérifier cette hypothèse nous avons enregistré les spectres de la poly(L-proline) en absence et en présence de  $\text{CaCl}_2$  : aucune modification dans l'allure du spectre n'a été observée. Cette hélice ne semble pas réagir de la même façon que l'hélice  $\alpha$  vis-à-vis des ions calcium. Nos résultats montrent qu'en présence de calcium, IB7 et IB7₁₄ conservent leur structure partiellement en hélice de type II et en pelote statistique.

### **IV. 3. 2. RMN**

Dans un premier lieu l'analyse par RMN du peptide IB7₁₄ sera présentée, puis l'attribution de la protéine entière IB7 sera évoquée.

#### ***IV. 3. 2. a. Attribution des protons du peptide IB7₁₄ par RMN***

La première étape nécessaire à la caractérisation structurale par RMN est d'identifier l'ensemble des résonances proton de chacun des acides aminés. Les déplacements chimiques ont des valeurs standards pour un acide aminé en solution dans l'eau et, au sein d'une protéine, les valeurs sont sensiblement différentes en fonction des résidus voisins et du repliement de la chaîne peptidique. Il existe des abaques répertoriant pour chaque acide aminé, les résonances des différents protons, en annexe 1. Les expériences TOCSY sont utilisées à ces fins, dont un agrandissement de la région de corrélations NH-H $\alpha$  est présenté figure IV-7. Dans cette région, tous les protons de chaque acide aminé possédant un proton lié à l'azote sont identifiés par système de spin. Les résidus prolines sont identifiés dans la région 2 à 4,5 ppm.

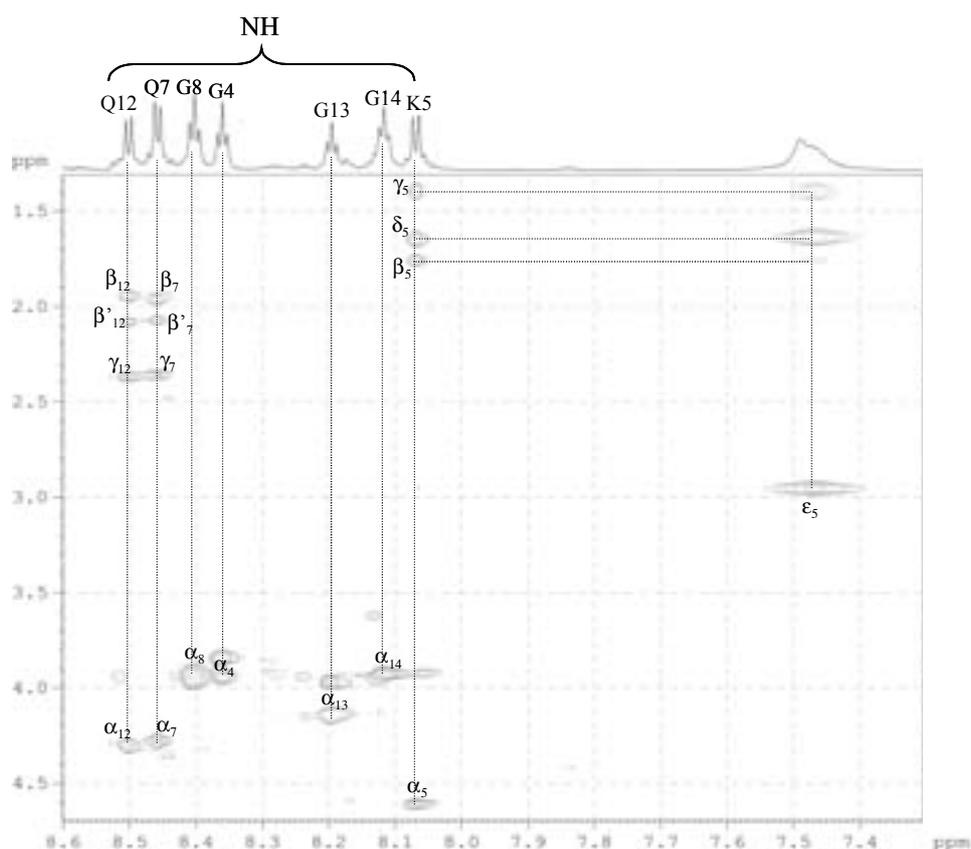


Figure IV-7 : Carte TOCSY représentant la région des NH-H $\alpha$  du peptide IB7₁₄.

De plus, les prolines sont distinguables par la valeur de leur déplacement chimique en fonction des résidus voisins : les déplacements chimiques des protons alpha et delta des prolines sont différents si elles précèdent une autre proline (**Pro-Pro**) ou bien un résidu quelconque (**Pro-X**) comme il a déjà été montré dans la littérature¹²⁷.

La figure IV-8 montre un agrandissement de la région de 2 à 4,5 ppm, typique des protons de résidus proline. Les deux groupes de prolines y sont facilement distinguables.

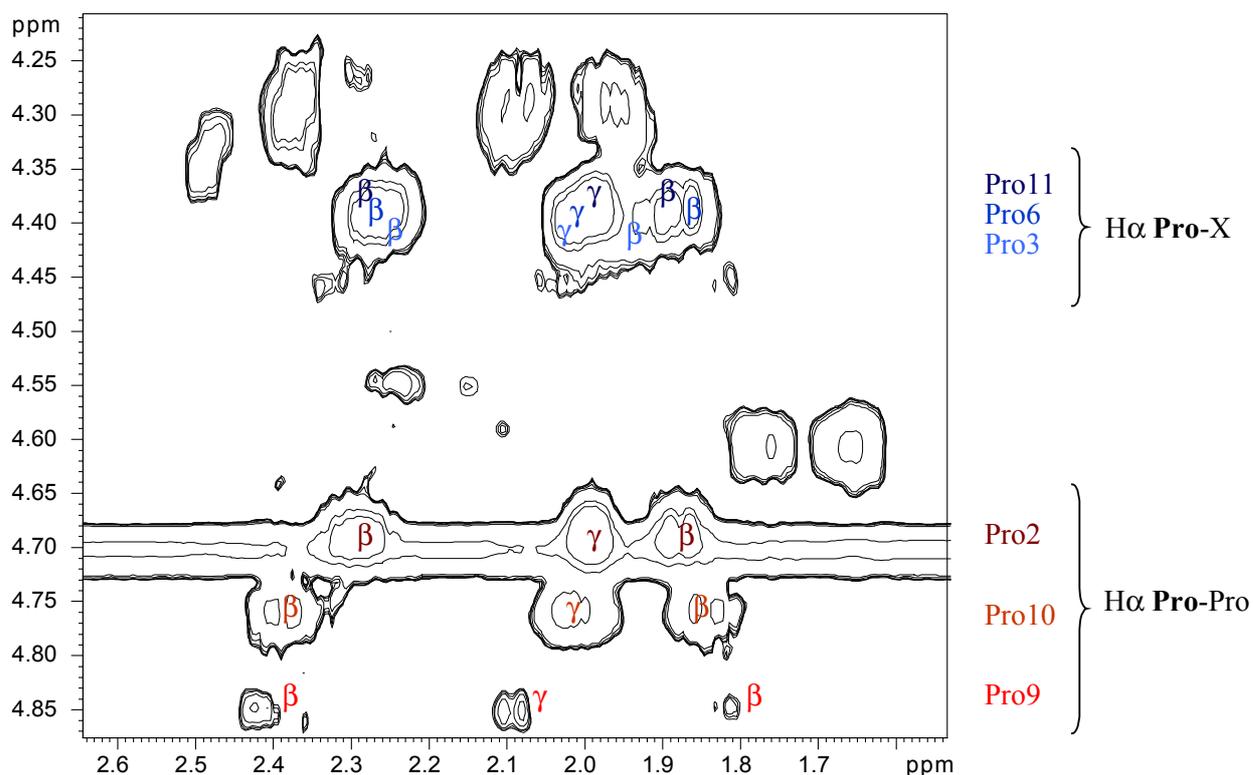
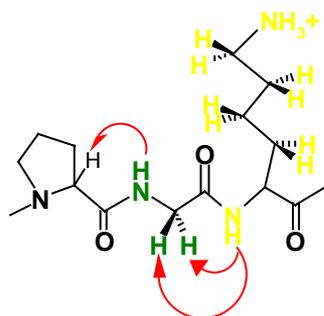


Figure IV-8 : Carte TOCSY représentant la région des prolines du peptide IB7₁₄.

En effet, les H $\alpha$  des prolines 2, 9 et 10 (prolines précédant une proline) sont fortement déblindés, alors que les H $\delta$  des mêmes prolines sont blindés par rapport aux prolines 3, 6 et 11 précédents des résidus autres que proline.

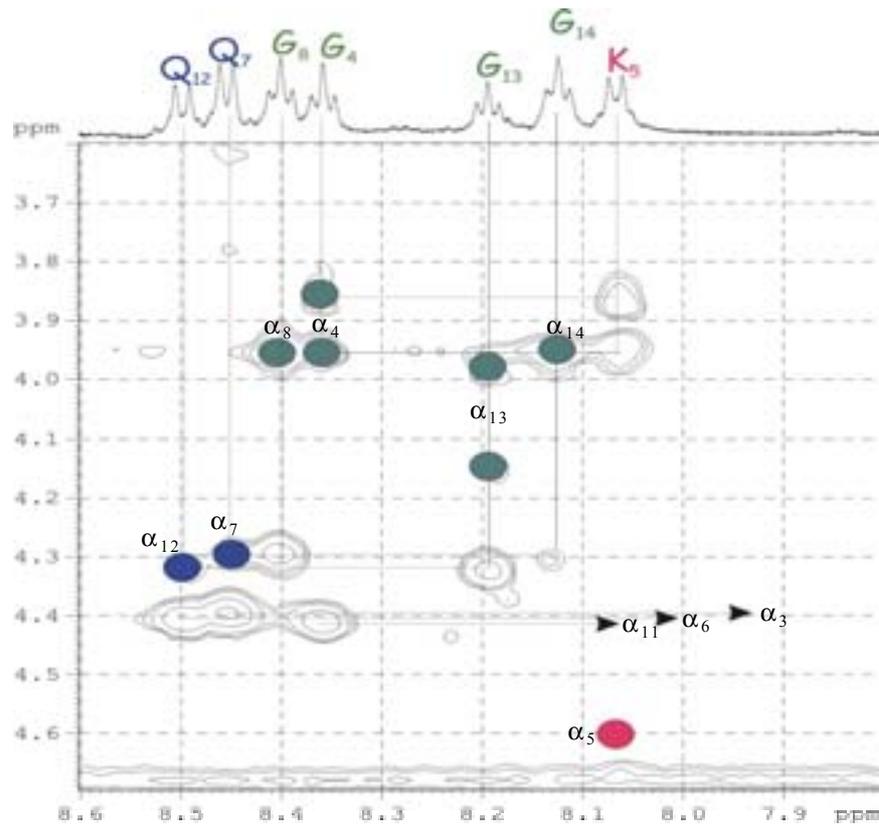
Une fois l'attribution de l'ensemble des protons de chaque acide aminé faite, il reste à savoir dans quel ordre les acides aminés se situent dans la séquence peptidique. La carte ROESY nous permet de réaliser une attribution séquentielle selon la méthode de Wüthrich¹⁰². La présence de corrélations entre les NH et les H $\alpha$  du résidu voisin indique l'enchaînement des résidus sur la séquence, comme représenté sur le schéma ci-dessous :



S1_P2_P3_G4_K5_P6_Q7_G8_P9_P10_P11_Q12_G13_G14

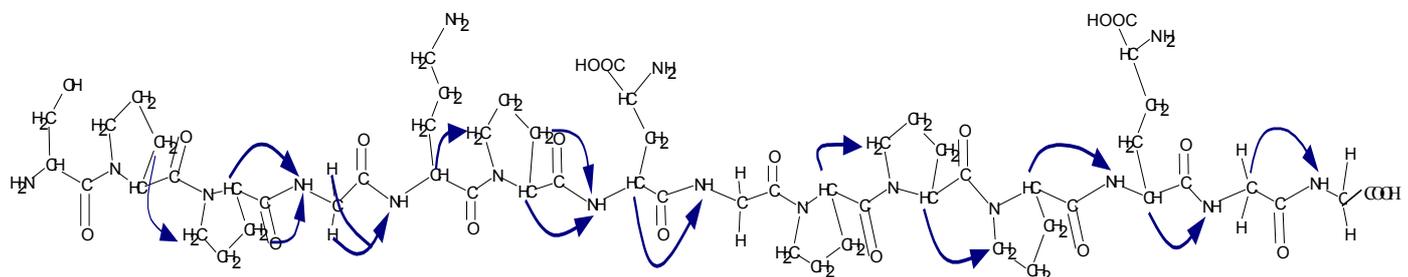
En revanche, l'introduction d'une proline dans la séquence empêche cette attribution par les corrélations  $\text{NH} \rightarrow \text{H}\alpha$ . L'attribution séquentielle au niveau des prolines peut toutefois être réalisée en considérant les corrélations  $\text{H}\delta$  des prolines et  $\text{H}\alpha$  du résidu précédent une proline.

Les corrélations séquentielles entre atomes voisins sont visualisées sur la carte ROESY comme le montre la figure IV-9 :



**Figure IV-9 :** Carte ROESY représentant la région des  $\text{NH-H}\alpha$  du peptide  $\text{IB7}_{14}$ . Les corrélations  $\text{H}\alpha_i \rightarrow \text{NH}_{i+1}$  sont représentées par des pointillés.

Les effets NOE séquentiels observés sont représentés par des flèches sur la figure IV-10 :



**Figure IV-10 :** Représentation de la structure primaire du peptide IB7₁₄, ainsi que les effets NOE séquentiels observés (flèches bleues).

Ainsi de proche en proche nous avons attribué tous les protons, répertoriés dans le tableau IV-1. Nous pouvons noter le manque de NOEs séquentiels entre la glycine 8 et la proline 9 et entre la proline 9 et la proline 10. Nous remontons à l'attribution totale en alignant des fragments de séquences attribués.

	HN	H $\alpha$	H $\beta$	Autres
Serine 1	ND	4,390	3,991/3,842	-
Proline 2	-	4,691	1,874/2,290	H $\gamma$ : 1,994 ; H $\delta$ : 3,574/ ND
Proline 3	-	4,390	1,921/2,251	H $\gamma$ : 2,021 ; H $\delta$ : 3,805/ 3,629
Glycine 4	8,355	3,916/3,860	-	-
Lysine 5	8,067	4,598	1,759	H $\gamma$ : 1,410 ; H $\delta$ : 1,65 ; H $\epsilon$ : 2,963
Proline 6	-	4,385	1,902/2,262	H $\gamma$ : 2,003 ; H $\delta$ : 3,787/3,61
Glutamine 7	8,456	4,275	1,973/2,072	H $\gamma$ : 3,361
Glycine 8	8,400	3,942/3,895	-	-
Proline 9	-	4,839	1,812/2,421	H $\gamma$ : ND ; H $\delta$ : 3,5/ ND
Proline 10	-	4,746	1,845/2,383	H $\gamma$ : 2,014 ; H $\delta$ : 3,565/3,741
Proline 11	-	4,370	1,864/2,271	H $\gamma$ : 1,984 ; H $\delta$ : 3,603/3,76
Glutamine 12	8,493	4,302	1,952/2,076	H $\gamma$ : 2,38
Glycine 13	8,189	4,120/3,988	-	-
Glycine 14	8,123	3,932	-	-

**Tableau IV-1 :** Valeurs des déplacements chimiques  $^1\text{H}$  en ppm du peptide IB7₁₄.

### **IV. 3. 2. b. Essais de l'attribution des protons de la protéine IB7**

La même approche, utilisant la batterie classique (COSY-DQF, TOCSY et NOESY) a été tentée sur la protéine IB7 entière. En revanche les cartes obtenues par RMN montrent un chevauchement important des raies de résonance. L'utilisation d'un spectromètre 800 MHz ou d'un solvant structurant comme le TFE ne nous a pas permis de différencier tous les protons. La carte TOCSY présentée figure IV-11 confirme ces observations.

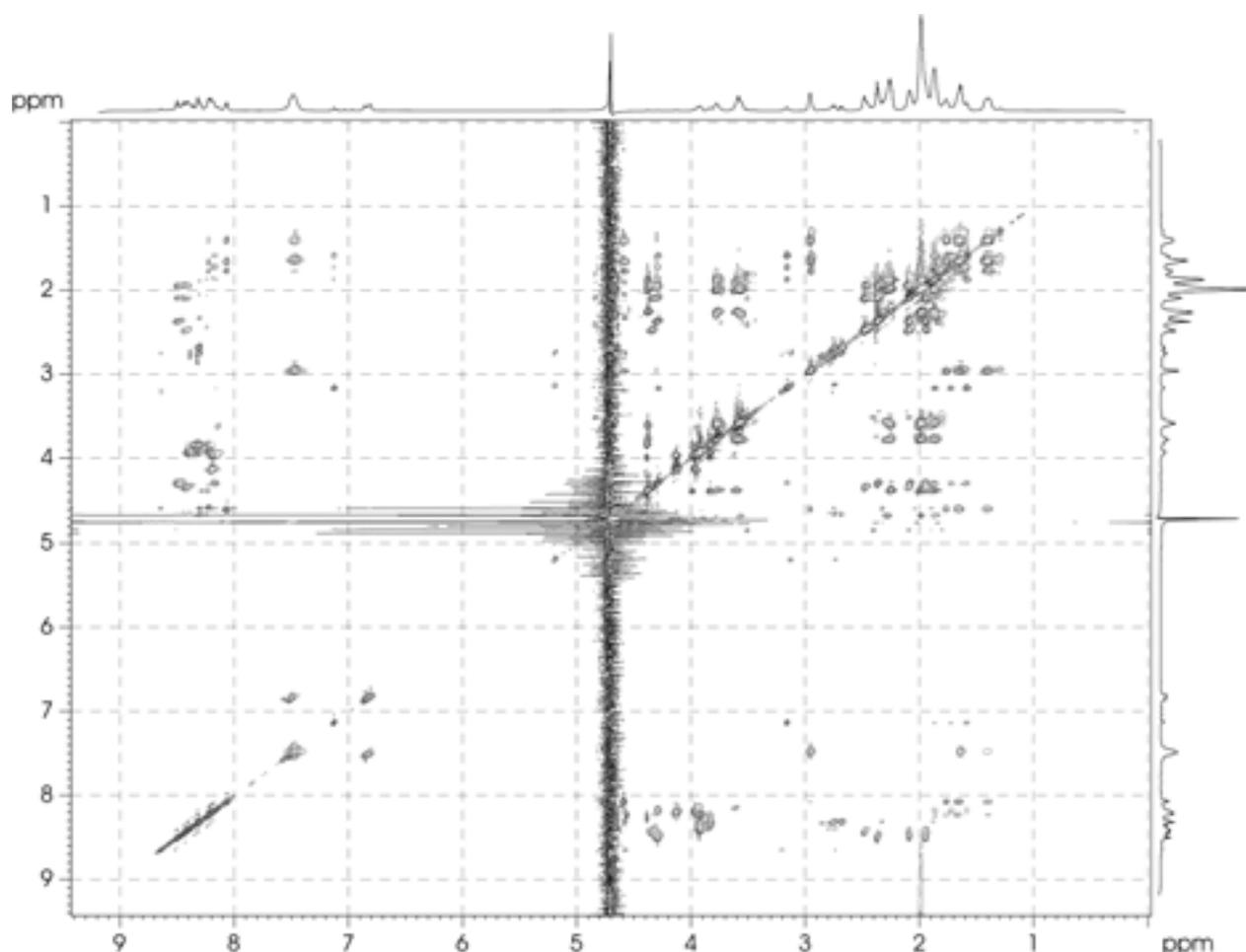


Figure IV-11 : Carte TOCSY de la protéine IB7 (à 5 mM dans l'eau) effectuée sur un spectromètre 800 MHz.

Sur les 59 acides aminés nous devrions dénombrer 34 systèmes de spins dans la région NH-H $\alpha$  (59 – 25 prolines) alors qu'une vingtaine au plus sont identifiables et les 25 prolines ne sont pas toutes visualisables.

#### ***IV. 3. 2. c. Etablissement d'une liste de contraintes pour le peptide IB7₁₄***

Afin d'établir la structure réelle adoptée par le peptide, deux paramètres sont à prendre en considération. Non seulement les effets NOE inter-résidus démontrent la proximité spatiale des noyaux d'une manière séquentielle et non séquentielle, mais également les constantes de couplage  3J  sont en fonction des angles dièdres entre atomes voisins ce qui nous renseigne donc sur les contraintes d'angles du squelette peptidique. Ces deux types de contraintes ( 3J  et NOEs) sont facilement accessibles pour des peptides rigides mais cela devient difficile pour des peptides flexibles.

La carte COSY-DQF, figure IV-12, nous permet d'accéder aux corrélations entre les NH et les H $\alpha$  appartenant au même résidu, mais également aux constantes de couplage  3J  entre ces protons. Sur la carte nous pouvons distinguer toutes les corrélations NH-H $\alpha$  mis à part la sérine N-ter. Celle-ci est identifiée par ses corrélations entre les H $\alpha$  et H $\beta$  spécifiques.

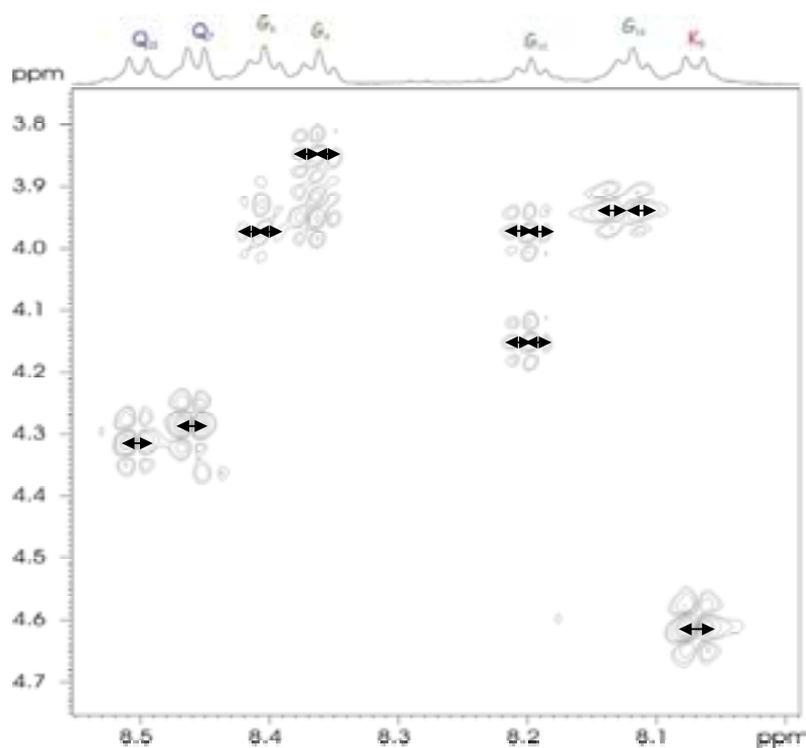


Figure IV-12 : Carte COSY-dqf représentant la région NH-H $\alpha$  du peptide IB7₁₄. La constante de couplage  3J  est directement calculable par la mesure de la distance entre les raies d'un multiplet montrée par les flèches.

41 effets NOEs et 7  3J  constituent le fichier de contraintes. Ce fichier de contraintes est présenté en annexe 2.

### IV. 3. 3. Minimisation et Dynamique Moléculaire sous contraintes RMN : Structure tridimensionnelle du peptide IB7₁₄

Les cartes bidimensionnelles obtenues par RMN nous ont permis de réaliser une cartographie des déplacements chimiques de tous les protons du peptide ainsi que de construire un fichier de contraintes structurales de part les effets NOE et les constantes de couplage. Ces données sont ensuite utilisées pour déterminer la structure de plus basse énergie. A la fin de la minimisation, les 20 conformations de moindre énergie obtenues sont sélectionnées.

La figure IV-13 montre la superposition des vingt structures minimisées retenues.

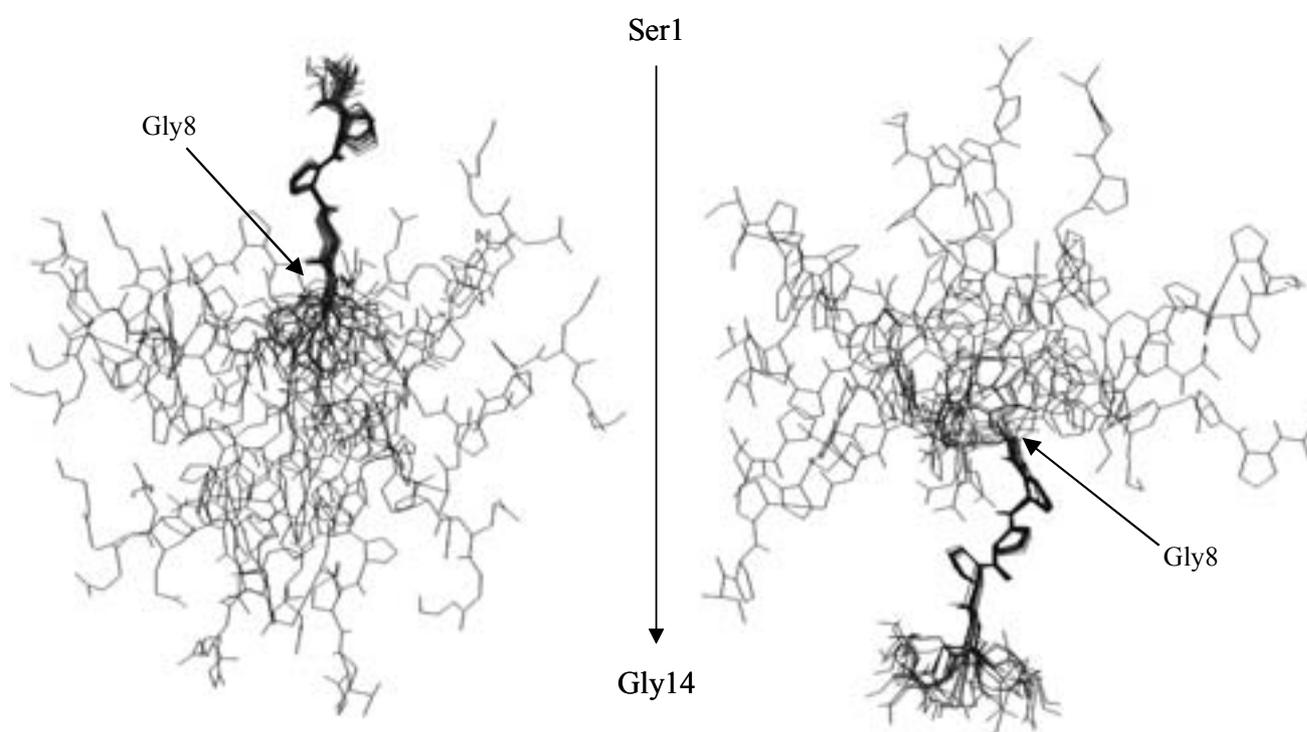
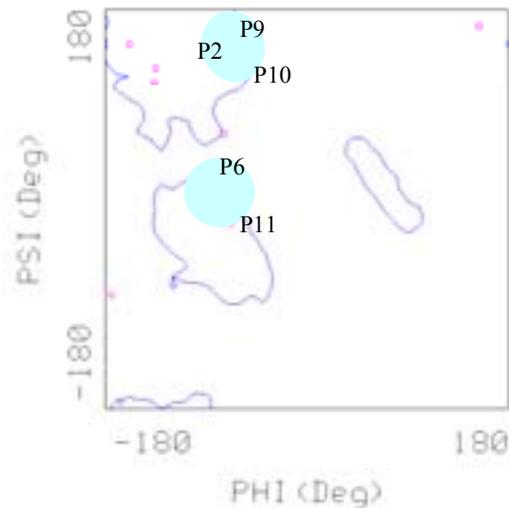


Figure IV-13 : Représentation des vingt conformations de plus basse énergie après minimisation du peptide IB₇₁₄. Deux superpositions de séquences peuvent être réalisées indépendamment l'une de l'autre : d'une part entre Pro3 et Pro6 (gauche) ou bien entre Pro9 et Gly13 (droite).

Le principe de la minimisation requiert de travailler à très hautes températures (500 ou 1000 K) afin de casser les barrières énergétiques. En revanche, malgré l'utilisation de conditions dénaturantes, le peptide conserve certaines zones dans des conformations identiques en hélice de type II dont le pas est particulièrement ouvert. Deux superpositions sont possibles de part et d'autre de la glycine 8, apparaissant comme un axe de libre mouvement. D'une part, le segment Pro2-Pro6 a une structure conservée d'une minimisation à l'autre et indépendamment, le motif Pro9-Gly13 est également superposable sur les vingt structures.

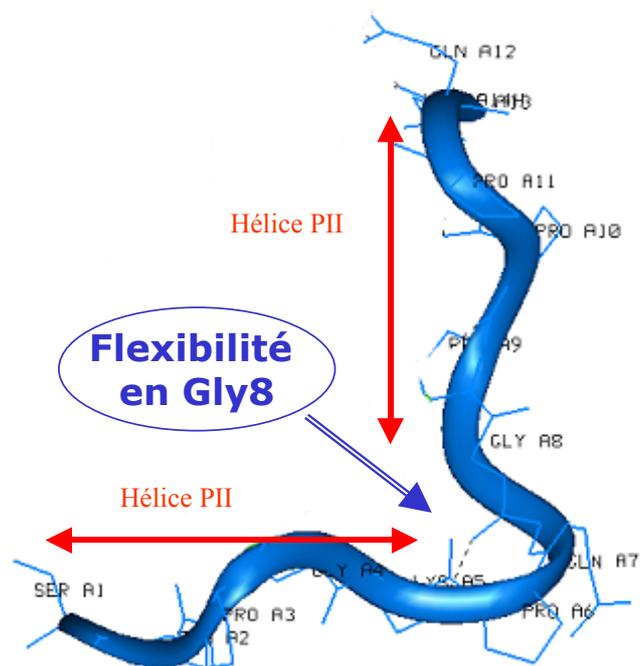
Ces observations confirment la conformation étendue du peptide, avec des micro-séquences en hélice de type II principalement au niveau des prolines consécutives Pro2-Pro3 et Pro9-Pro10-Pro11. Un diagramme de Ramachandran du peptide IB₇₁₄ a été réalisé (figure IV-14), afin de comparer les angles adoptés par le peptide aux valeurs  $\phi$  et  $\psi$  des résidus au

sein d'une structure poly(L-proline) (figure IV-3). Les résidus XXX s'inscrivent parfaitement dans le domaine d'existence d'une hélice de type II.



**Figure IV-14 :** Carte de Ramachandran du peptide IB7₁₄ : elle présente quelques résidus dans la zone typique de l'hélice de type II, colorée en bleu¹⁵⁷.

La figure IV-15 représente une des structures de plus basse énergie du peptide IB7₁₄. La structure est très ouverte et étalée avec des motifs en hélice de type II de part et d'autre de la glycine 8. La distance entre deux résidus placés l'un au-dessus de l'autre dans l'hélice est estimée à 8 Å, qui est effectivement le pas de l'hélice de type II.



**Figure IV-15 :** Structure d'un conformère de plus basse énergie. Deux parties en hélice sont situées de part et d'autre de la glycine 8.

## IV. 4. CONCLUSIONS ET PERSPECTIVES

La structure secondaire du peptide IB7₁₄ et de la protéine entière IB7, obtenue par DC (36% d'hélice de type II et 64% de pelote statistique) a été confirmée par modélisation moléculaire (MM) pour IB7₁₄ en tenant compte des contraintes émanant des expériences de RMN. La structure tertiaire du peptide a pu également être obtenue par RMN et MM. Pour la première fois une analyse structurale par DC, RMN et DM a été effectuée sur un peptide de type PRP. Murray *et al.*¹²⁷ avaient toutefois effectué une étude structurale par DC et RMN. Le peptide IB7₁₄ adopte une conformation étendue due à une répartition de structure secondaire en hélice de type II et en pelote statistique. Les deux portions en hélice PII sont séparées par la glycine 8, qui semble alors jouer le rôle d'une « rotule », d'un point de flexibilité. Le manque d'effets NOE séquentiels autour de la glycine 8 et de la proline 9 démontre cette potentielle flexibilité. L'hélice PII est plus étendue que l'hélice  $\alpha$  : en effet, le pas de l'hélice PII est  $\sim 8\text{\AA}$  alors que le pas de l'hélice  $\alpha$  est de  $\sim 5\text{-}6\text{\AA}$ . La propriété de cette hélice est principalement due à la présence de multiples résidus proline. Hagerman et Butler ont ainsi proposé que la nature étendue et ouverte des protéines était requise pour l'interaction avec les tanins¹.

La fonction primaire des PRPs basiques consiste à précipiter les polyphénols, propriété émanant principalement de leur conformation. Les caractéristiques spécifiques de la pelote statistique et de l'hélice PII sont très proches. Les désaccords perçus dans la littérature au sujet de la conformation adoptée par les PRPs viendraient de cette particularité. Par dichroïsme circulaire l'allure entre ces deux conformations est la même, seules l'amplitude et les minima diffèrent¹²⁴.

L'étude structurale par RMN s'est focalisée sur le peptide IB7₁₄, dont la séquence est reprise trois fois dans la protéine entière. En effet, l'attribution de l'ensemble des protons sur la protéine entière IB7 est rendue très délicate de part sa composition : trois résidus (25 prolines, 13 glycines et 9 glutamines) représentent 70 % de la composition primaire en acides aminés. Chaque acide aminé possède un groupe de protons ayant un déplacement chimique qui lui est propre et qui est fonction de son environnement électronique (dépendant non seulement de la nature de l'acide aminé mais également de la nature de ses voisins). Dans ce cas particulier, les trois acides aminés principaux sont représentés en grand nombre et dès lors qu'ils sont imbriqués dans des motifs répétés, ils ont pour la plupart des environnements

similaires. La différenciation des 59 systèmes de spin est donc rendue difficile, les spectres présentant des raies de résonance chevauchantes. L'utilisation d'un solvant structurant comme le TFE ne nous a pas permis d'accéder à la structure 3D de la protéine IB7. L'utilisation d'un spectromètre 800 MHz ne nous a pourtant pas permis d'obtenir une résolution suffisante pour attribuer les différents acides aminés. De ce fait, l'obtention d'une structure par dynamique moléculaire (DM) sans contrainte RMN ferait l'objet d'un temps de calcul bien trop important pour être envisagé.

L'accès à la structure tridimensionnelle du peptide IB7₁₄ dont la séquence est reprise trois fois dans la protéine, devrait nous donner des informations de contraintes sur la protéine.

Cependant, l'adaptation de techniques de RMN 3D standards aux protéines dont la composition provoque la superposition des raies de résonance se développe de plus en plus. En effet, la troisième dimension de cartes 3D homonucléaires (TOCSY-ROESY ou ROESY-TOCSY) ou hétéronucléaires (HMBC-TOCSY) permettraient de différencier des taches de corrélations chevauchantes en 2D. De même, l'enrichissement isotopique en ¹³C et/ou en ¹⁵N nous permettrait de distinguer préférentiellement certaines positions sur la protéine et donc de pallier le chevauchement des pics¹⁷¹. Le double enrichissement apporte une aide majeure à la caractérisation de protéines de haut poids moléculaire (24 kDa) en RMN¹⁷², bien que la technique soit encore très coûteuse.

Les résultats obtenus par DC nous ont montré que la protéine entière présentait la même répartition de structure secondaire que le peptide IB7₁₄[Simon, 2002 #107]. Nous avons donc choisi de prendre le peptide comme modèle représentatif des PRPs pour étudier la structure 3D de l'interaction Protéine-Tanin. Par contre, les mesures thermodynamiques ( $K_d$  et stoechiométrie) ont été réalisées sur la protéine entière IB7 et sur le peptide IB7₁₄.



**Chapitre V INTERACTIONS TANINS-  
PROTEINES**



## V. 1. INTRODUCTION

Dans la salive, certaines protéines, les PRPs, ont un rôle de lubrifiant. Ces mêmes protéines sont connues pour interagir avec les polyphénols issus de l'alimentation, formant ainsi des complexes insolubles, notamment avec les tanins condensés⁴⁷. Cette interaction réduirait ainsi le potentiel lubrifiant de la salive ce qui est perçu comme une astringence. L'astringence ou sécheresse est entrevue comme une sensation tactile au niveau de la bouche émanant de l'interaction polyphénols-protéines de la salive^{7,21}. Le mécanisme d'interaction est mal connu. Il proviendrait d'interactions hydrophobes et de liaisons hydrogène, en dépendant de nombreux facteurs tant au niveau du polyphénol que de la protéine ainsi que des conditions du milieu. Les critères définissant cette interaction, comme la nature des liaisons entre les deux partenaires (tanins et protéines), la spécificité des sites de liaison sur la protéine ainsi que la stœchiométrie des complexes, vont être résumés à partir de la littérature.

### V. 1. 1. Nature de l'interaction

#### V. 1. 1. a. *Nature et non-spécificité des sites de liaison*

D'après la littérature, il semblerait qu'il n'y ait pas de séquence spécifique d'interaction sur la protéine mais un site privilégié en fonction d'une suite d'acides aminés entraînant une conformation propice à l'interaction. Le seul fait d'avoir des séquences riches en résidus proline, glycine et glutamine suffirait à provoquer la complexation avec les tanins. Ainsi, par le suivi des déplacements chimiques des protons de deux peptides de 19 et 22 résidus n'ayant pas un enchaînement identique, Murray *et al.* constatent une affinité similaire des deux peptides vis-à-vis des polyphénols⁸⁵. La particularité des sites tient majoritairement à la présence de résidus prolins. La proline serait un bon site de fixation car elle a une structure ouverte et rigide. Par ailleurs, la fonction amide tertiaire de la proline est enthalpiquement plus favorable pour former un zwitterion que les amides secondaires. Kay *et al.*¹⁷³ rappellent que les prolins sont un motif de liaison hors pair car l'amide tertiaire induit une augmentation de la basicité du carbonyle plus apte à accepter des liaisons hydrogène

174,175

Le premier résidu d'un enchaînement X-proline est un site préférentiel d'interaction^{80,127}. En effet, Baxter *et al.*¹²⁸ montrent par RMN que le cycle galloyle se lie sur la proline ou sur le résidu la précédent (X-Pro). En outre, Hatano et Hemingway¹⁷⁶ étudient l'interaction du tetrapeptide Gly-Pro-Gly-Gly avec un monomère et un dimère de catéchine et montrent qu'il n'y a pas de sélectivité d'interaction sur les résidus prolines, mais qu'elles définissent les contraintes spatiales permettant l'approche des tanins. Il est donc fortement présumé que le fort taux de prolines requis n'est dû qu'à la conformation ouverte et étendue qu'elles engendrent. Glycine et glutamine joueraient également un rôle dans l'interaction entre tanins et protéines⁴⁵.

### ***V. 1. 1. b. Types de liaison tanin-protéine***

La nature de l'interaction demeure quant à elle mal connue. Liaisons hydrogène et interactions hydrophobes sont les forces majeures de l'interaction en conditions physiologiques¹⁷⁷. Selon certains auteurs, les tanins et les protéines interagiraient majoritairement *via* des interactions hydrophobes entre le groupement latéral (pyrrolidine) de la proline et les cycles aromatiques des polyphénols. Dans ce cas, les liaisons hydrogène, apparaissant comme secondaires, serviraient à stabiliser le complexe^{80,85,178}. Spencer *et al.* suggèrent que la liaison de nature hydrophobe est formée par certains résidus aromatiques de la chaîne hydrocarbonnée des protéines³. Cependant les PRPs et en particulier la protéine IB7, ne possèdent pas de résidus aromatiques dans leur séquence. D'après Baxter *et al.*, le stacking hydrophobe du phénol sur la proline est le mode d'interaction majeur sur un segment PRP de 19 résidus, dans 10% de DMSO¹²⁸. En revanche, d'autres études suggèrent que l'interaction provient de la formation de liaisons hydrogène entre l'hydroxyle phénolique et le carbonyle du peptide^{1,179}. Peleg *et al.* suggèrent également que l'augmentation de l'astringence avec le degré de polymérisation des tanins serait due à la capacité des plus gros polymères à former davantage de liaisons hydrogène entre les groupements hydroxyles des composés phénoliques et le peptide¹⁸⁰. D'une manière générale, on parle de phénomène de surface⁵⁸ : les tanins seraient adsorbés à la surface des protéines en formant ainsi une monocouche hydrophobe.

### ***V. 1. 1. c. Affinité : aspects thermodynamique des complexes***

L'existence de plusieurs sites de liaisons des tanins sur une même protéine est largement envisageable étant donné qu'il n'y a pas de spécificité du site de reconnaissance.

En effet, Hagerman et Butler¹ montrent que le nombre de tanins fixés augmente avec la taille de la protéine. Ce résultat est en accord avec une multiplicité des sites.

Toutes les mesures de constante de dissociation ( $K_d$ ) du complexe tanin-protéine ont été faites en considérant qu'il existe plusieurs sites de liaison sur une protéine et que ces sites sont indépendants entre eux. En effet, la proximité des cycles aromatiques des tanins induit une variation de déplacement chimique de certains protons du peptide dont il est possible de retirer le  $K_d$  et le nombre de sites⁸⁵. Les déplacements chimiques les plus influencés par la présence de tanins sont majoritairement les deux premiers résidus de motifs de type **X-Pro-Pro** (notamment avec X = prolines, glycine ou arginine). Ces variations suggèrent que les tanins se fixent au niveau de ces résidus. L'interaction a également été suivie par la mesure de variation de turbidité de solutions contenant tanins et protéines^{133,181}, ou encore par des essais compétitifs au moyen de protéines marquées¹.

Concernant le tanin, l'affinité PRP-tanin observée augmente avec le nombre de groupements galloylés que possède le tanin : Penta Galloyle Glucose > TétraGG > TriGG¹³⁰. De même, l'épicatéchine interagit davantage avec les PRPs lorsqu'elle est galloylée¹³³. D'une manière générale, la régiospécificité (liaison interflavane 4-6 ou 4-8) ainsi que le degré de polymérisation du tanin sont à l'origine de la différence d'affinité entre les PRPs et les tanins. Le dimère B3 (catéchine(4 $\beta$ -8)-catéchine) forme des complexes de plus grande taille avec les PRPs que le monomère catéchine¹³³. La constante de dissociation entre un peptide de 19 résidus et le PGG ou l'épigallocatechine galloylée varie d'un facteur dix autour du millimolaire (~10 et 1 respectivement), alors que le nombre de sites de liaison est estimé à 2 dans le premier cas et 8 pour le second¹⁸². Ces résultats traduisent non seulement la relation entre la taille du tanin et l'affinité, mais aussi avec le nombre de sites, car plus un polyphénol a une taille importante moins il y aura de sites sur la protéine.

#### ***V. 1. 1. d. Paramètres contrôlant le complexe tanin-protéine***

Les critères définissant l'affinité ou la taille du complexe tanin-protéine peuvent être regroupés en deux classes : ceux qui sont intrinsèques à la nature de la protéine et du polyphénol, mais aussi les conditions dans lesquelles ont lieu ces complexations.

Citons donc tout d'abord quelques paramètres intrinsèques à la nature de la protéine et du polyphénol. La taille de la protéine, qui va d'une cinquantaine d'acides aminés (IB7 = 59)

à une centaine (IB6 = 118) est à prendre en considération. Une protéine trop longue pourrait avoir tendance à se replier, masquant ainsi des sites de fixation, ce qui limitera la reconnaissance des sites d'interaction entre tanin et protéine. En outre, la fixation étant « aspécifique » plus la protéine est longue plus le nombre de sites potentiels est important¹, car l'augmentation du nombre de prolines suffit à augmenter le nombre de sites de liaison¹⁸³. Apparaît là une notion de rendement de tanins fixés en fonction du nombre de résidus aminés et de la conformation de la protéine. Un compromis doit être fait entre la nécessité d'un fort taux de prolines et un repliement potentiel de la protéine entraînant un masquage des sites. De même, la taille du polyphénol (soit le degré de polymérisation pour les tanins condensés) est importante. L'augmentation de leur poids moléculaire favorise la complexation avec les protéines^{16,68,184,185}. Le nombre de groupements hydroxyles sur les cycles A et B ainsi que les galloylations sont importants. Ainsi, plus il y a de OH plus l'affinité est importante⁸¹ et donc le nombre d'unités monomériques fait accroître la quantité d'hydroxyles.

Ensuite viennent les paramètres concernant les conditions de solubilisation du complexe. La liste suivante sera non exhaustive. L'acidité et la force ionique modifient l'interaction protéine-polyphénol et réduisent la viscosité de la salive. La valeur du pH est très importante pour l'interaction¹⁸⁶ de sorte que plus le pH est bas et plus le complexe semble stable. A pH 8,2 il n'y a pas de précipitation mais formation d'un complexe soluble¹⁸⁷. Par une titration en fonction du pH sur plusieurs protéines, de Freitas et Mateus¹³³ montrent que la fixation des tanins sur les PRPs et la précipitation du complexe sont maximales à pH 3,5 et à pH 5, correspondant souvent au point isoélectrique des protéines¹⁸⁸ où les forces de répulsions électrostatiques sont minimisées. L'effet du pH est dépendant de la concentration en composés ioniques¹⁷⁸. A faible pH et faible concentration en protéines, une augmentation de la force ionique augmente la complexation. La charge globale de la protéine est écrantée par les ions qui réduisent les répulsions électrostatiques entre les protéines, augmentant la force hydrophobe¹⁸⁹. De plus, dans la salive humaine, des ions tant inorganiques qu'organiques sont présents. Martin *et al.*¹⁹⁰ miment un mélange d'ions et trouvent que l'effet du pH est très sensible à leur présence et dépend de la nature des ions présents, de telle sorte que la présence d'ions inorganiques (K⁺, Na⁺) favorise la précipitation du complexe tanins-protéines¹⁸⁸. L'influence du solvant sur la solubilité des polyphénols agit sur l'association avec les protéines : une faible solubilité favorise la fixation des tanins car les liaisons hydrogène faites avec le solvant doivent être supprimées avant la liaison aux protéines³. Des composés secondaires peuvent également intervenir sur la fixation des tanins et l'affinité avec les

protéines. Les polysaccharides peuvent aussi interférer en agissant de manière compétitive avec les protéines sur la fixation des polyphénols.

## **V. 1. 2. Stratégie de l'étude**

L'interaction protéine-tanin est encore mal connue à ce jour. Seule la non spécificité des sites de liaisons est admise dans l'ensemble de la littérature. En revanche, la nature des liaisons est très discutée. La stœchiométrie des complexes ainsi que la constante de dissociation a principalement été étudiée par mesure des variations de déplacements chimiques en RMN ou par compétition par marquage radioactif. Aucune approche structurale du complexe n'a été entreprise.

Nous proposons de valider l'existence d'un complexe entre le peptide IB7₁₄ et les tanins B3 par spectrométrie de masse, RMN et dynamique moléculaire. La stœchiométrie du complexe, le nombre de sites de liaison des tanins sur le peptide IB7₁₄ ainsi que la nature de l'interaction sont ainsi mis en évidence à l'aide de ces différentes techniques.

D'autre part, des mesures de constantes de dissociation, de nombre de sites sur la protéine entière ainsi que la taille du complexe, ont été réalisées pour le système IB7-B3.

## **V. 2. MATERIELS ET METHODES**

Les produits utilisés ont été préalablement présentés lors des chapitres précédents.

### **V. 2. 1. L'ESI couplée à un détecteur à trappe ionique**

Un spectromètre de masse associé à un détecteur à trappe ionique équipé d'une interface nanospray (LCQ Deca XP, Thermo Finnigan, laboratoire IBGC, Université de Bordeaux 2) est utilisé pour l'étude de l'interaction peptide-tanin.

Des solutions de peptide (50 picomoles), auxquelles sont ajoutées des quantités croissantes de B3 (de 0 à 200 picomoles), sont mises en solution dans un mélange eau/éthanol/acide acétique 89/10/1 (v/v/v), à pH 3,5. La température du capillaire dans

l'interface atteint 150°C, et le voltage induisant la nébulisation est calé à 1,2 kV. Les espèces doublement chargées correspondant aux différents complexes peptide-tanin sont isolées dans la trappe ionique avec une précision de  $\pm 3$  Da. L'énergie de collision nécessaire à la perturbation du complexe se situe entre 18 et 25 % par rapport à l'échelle instrumentale, en fonction de la taille des complexes.

## **V. 2. 2. Dichroïsme Circulaire**

Les conditions d'acquisition ont été décrites dans le chapitre IV.2.2.a.

Les spectres de DC ont été enregistrés pour des solutions contenant initialement 0,2 mM de peptide IB7₁₄ auxquels ont été progressivement rajoutées différentes concentrations de tanin B3 (de 0,025 à 0,6 mM).

## **V. 2. 3. Dynamique Moléculaire du complexe IB7₁₄-B3**

La dynamique moléculaire a été utilisée à différents buts : une première approche concerne l'étude structurale et dynamique du complexe IB7₁₄/B3 ; puis, elle nous a également permis d'identifier les régions lipophiles ou hydrophiles du peptide et enfin elle nous a aidés à remonter à la taille du complexe.

### ***V. 2. 3. a. Dynamique Moléculaire***

La minimisation du dimère B3 a été effectuée grâce au fichier de contraintes RMN établi par Barathieu⁶³. La conformation de plus basse énergie du peptide IB7₁₄ utilisée pour cette étude est celle obtenue par DM à 1000 K, en utilisant les contraintes obtenues pour le peptide seul (chapitre IV). Deux types d'expériences sont alors réalisés pour l'étude du complexe IB7₁₄/B3, l'une en maintenant le fichier de contraintes et l'autre pour laquelle les contraintes sont enlevées et le peptide complètement re-minimisé. Deux rapports molaires IB7₁₄/B3 sont testés, à 1/1 et 1/4. L'eau comme solvant implicite est systématiquement utilisée (GB/SA) avec l'option extended cut-off (cf chapitre IV). Quatre calculs de 2 ns sont effectués pour un rapport molaire de 1/1 et deux de 5 ns pour un rapport molaire de 1/4. Chaque calcul diffère par la position aléatoire des tanins autour du peptide de sorte tous les résidus du peptide soient en contact avec les tanins grâce à la multiplicité des expériences. Dans chaque cas, une conformation est prélevée toutes les 10 ps. Après les dynamiques

effectuées sous contraintes pour un rapport équimoléculaire, une analyse par cluster est réalisée avec le logiciel XCluster 1.1¹⁹¹, qui est intégré à MacroModel. Pour cela, nous utilisons un critère de distance, le RMSD, en comparant tous les atomes d'une structure à l'autre. Cette approche nous conduit à classer les conformères par famille selon l'homologie de structure.

### V. 2. 3. b. *Calcul de surface : amphiphilicité*

Les potentiels moléculaires de lipophilicité (MLP : Molecular Lipophilicity Potentials) ont été calculés par un programme développé au laboratoire (Michel Laguerre) d'après une idée de Audry *et al.*¹⁹² en utilisant une fonction exponentielle¹⁹³. Les constantes fragmentales atomiques utilisées ont été définies par Broto *et al.*¹⁹⁴. Les cartes MLP ont été calculées au niveau atomique grâce au programme MLPP¹⁹⁵.

### V. 2. 3. c. *Détermination du coefficient de diffusion par dynamique moléculaire et taille du complexe formé*

Les expériences sont réalisées dans de l'eau explicite. Dans le cas d'IB7₁₄ seul, le peptide est placé dans une boîte contenant 5140 molécules d'eau et de dimension 54 x 54 x 54 Å³. Pour le complexe IB7₁₄-3B3, il est placé dans une boîte de 5353 molécules d'eau et de dimension 55 x 55 x 55 Å³. Le peptide seul a été soumis à une dynamique de 2 ns à 300 K et le complexe à une dynamique de 3 ns à 300 K. Les expériences sont réalisées à pression constante (1 bar).

Les valeurs de coefficient de diffusion (D) sont obtenues à partir de la relation d'Einstein (Eq. [5]). La pente de la courbe, divisée par 6, représentant le carré du déplacement moyen de la molécule en fonction du temps est égale à D. Cette relation n'est valable que dans le cas d'une diffusion isotrope, correspondant à un système dilué.

$$D = \frac{1}{6N} \lim_{t \rightarrow \infty} \frac{d}{dt} \sum_{i=1}^N \langle [r_i(t) - r_i(0)]^2 \rangle \quad [5]$$

## V. 2. 4. RMN

Les conditions d'acquisition ainsi que la préparation des échantillons ont été décrites dans le chapitre IV.3.1.

### *V. 2. 4. a. Variation du déplacement chimique du peptide IB7₁₄ et de la protéine IB7*

Les spectres ont été enregistrés sur un spectromètre 500 MHz muni d'une sonde HRMAS. Des cartes TOCSY sont réalisées pour chaque rapport molaire IB7₁₄/B3 (de 1/0 à 1/7 pour une concentration de IB7₁₄ de 1,5 mM) afin d'attribuer la totalité des protons du peptide. Les conditions d'acquisition et de traitement des données sont identiques à celles utilisées pour l'obtention de la structure d'IB7₁₄ (§ IV-2-2). Les mêmes expériences sont réalisées pour des complexes IB7/B3 pour des rapports molaires variant de 1/0 à 1/21, pour une concentration en protéine IB7 fixe de 1 mM. Le dimère B3 est ajouté progressivement à la solution de peptide sous forme de poudre afin d'éviter toute dilution. Le mélange est soniqué afin de dissoudre complètement les tanins.

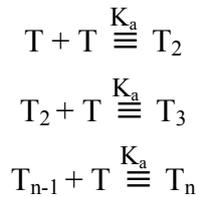
### *V. 2. 4. b. Variation du coefficient de diffusion du peptide IB7₁₄ et de la protéine IB7*

Les expériences haute résolution liquide sont effectuées sur un DPX Avance Bruker 400 MHz muni d'une sonde BBI avec un lock ²H, équipée de gradient z. L'échantillon est alors placé dans des tubes de 5 mm, nécessitant des volumes de 400 µl. La séquence utilisée est stebpgp1s1d (Bruker). L'impulsion  $\pi/2$  proton est de 5,8 µs à -6 dB. La valeur effective du gradient est incrémentée par pas de 5% (2, 5, 10,..., 95%) pour un total de 20 expériences. Dans le cas du peptide IB7₁₄ les paramètres de la séquence (cf chapitre II) ont été calibrés de la manière suivante :  $\Delta = 150$  ms ;  $\delta = 1,3$  ms. Les expériences de mesure du coefficient de diffusion transversal sont réalisées pour des rapports molaires IB7₁₄/B3 variant de 1/0 à 1/8. De la même manière, les paramètres concernant l'évolution du coefficient de diffusion transversal dans le cas de la protéine IB7 ont été calibrés comme suit :  $\Delta = 180$  ms ;  $\delta = 1,5$  ms (rapport 0 à 2,5) 1,8 ms (rapport 5 à 21) ; pour des rapports molaires de 1/0 à 1/21 des complexes IB7/B3.

### V. 2. 4. c. Mesures des constantes d'association et de dissociation

#### Auto-association des tanins

Les tanins sont connus pour leurs propriétés d'auto-association. Cette auto-association a été étudiée pour d'autres modèles de tanins tel que le galloyle glucose^{114,128} et le modèle le plus couramment utilisé pour étudier ces phénomènes est un modèle isodesmique^{114,196}. Ce modèle suppose que les molécules « libres », T (tanins), s'empilent pour former des « dimères », (T-T, deux unités reliées de manière non-covalente) « trimères », (T-T-T reliées non-covalamment) . . . avec la même constante d'association, selon le schéma suivant :



$$\text{Avec } K_a = \frac{[T_2]}{[T]^2} = \frac{[T_3]}{[T_2][T]} = \frac{[T_n]}{[T_{n-1}][T]}$$

De ces équations on peut écrire :

$$[T_0] = \frac{[T]}{(1 - K_a[T])^2}$$

Où  $[T_0]$  est la concentration totale en tanin ( $[T_0] = [T] + 2[T_2] + 3[T_3] + \dots + n[T_n]$ ) et  $[T]$  est la concentration en tanin disponible pour l'interaction ( $[T] = [T] + [T_2] + [T_3] + \dots + [T_n]$ ).

En utilisant le modèle d'empilement de tanin, Baxter *et al.*¹¹⁴ ont montré que le calcul de la constante d'association ( $K_a$ ) des tanins entre eux peut être effectué par le suivi de la variation de déplacements chimiques des protons du dimère B3, ceci en fonction de la concentration en dimère dans la solution. La relation utilisée afin de déterminer  $K_a$  est donnée par l'équation [6] :

$$\Delta\delta = \Delta\delta_{\max} K_a [T_0] \left\{ \frac{2}{1 + (4K_a [T_0 + 1])^{1/2}} \right\}^2 \quad [6]$$

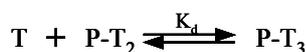
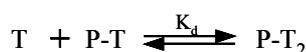
Où  $\Delta\delta$  est la variation en déplacement chimique observée,  $\Delta\delta_{\max}$  la variation maximale,  $K_a$  la constante d'association et  $[T_0]$  la concentration en tanin totale.

Les calculs ont été effectués sur le logiciel Microsoft Excel. Pour le paramétrage des courbes on utilise l'équation [6] où  $K_a$  et  $\delta_{\max}$  sont les paramètres ajustables (les courbes de titration sont montrées en annexe 3).

### Association du complexe Tanin-Protéine

La formation du complexe entre peptide et tanin induit une variation de l'environnement électronique des protons du peptide. On observe alors une variation des déplacements chimique des protons du peptide proportionnelle à la concentration de B3 ajouté. Une titration de l'interaction est alors possible.

La titration par mesure du déplacement chimique utilise une équation décrite par Charlton *et al.* et Baxter *et al.*^{128,130,182}, avec un modèle à plusieurs sites de liaison, de la manière suivante :



$$\text{Avec } K_d = \frac{n[T] \times [P]}{[T-P]_n}$$

Il est alors possible de retrouver la valeur de la constante de dissociation ( $K_d$ ) ainsi que le nombre de sites de liaisons sur le peptide en fonction des variations de déplacements chimiques des protons en RMN, comme montré dans l'équation [7] :

$$\Delta\delta = \frac{\Delta\delta_{\max}}{2} \left\{ 1 + \frac{K_d}{n[P]} + \frac{[T]}{n[P]} - \left[ \left( 1 + \frac{K_d}{n[P]} + \frac{[T]}{n[P]} \right)^2 - 4 \frac{[T]}{n[P]} \right]^{1/2} \right\} \quad [7]$$

où  $\Delta\delta$  est la variation de déplacement chimique du peptide (ppm),  $\Delta\delta_{\max}$  le maximum de variation (ppm),  $K_d$  la constante de dissociation,  $[P]$  la concentration totale en peptide,  $n$  le nombre de sites de liaisons pour le polyphénol et  $[T]$  la concentration totale en tanins capable de fixer le peptide. La concentration  $[T]$  est mesurée d'après l'équation [6]. Les

calculs sont effectués sur le logiciel Microsoft Excel. Pour le paramétrage des courbes utilisant l'équation [7],  $K_d$ ,  $\Delta\delta_{\max}$  et  $n$  sont les paramètres ajustables.

## V. 3. RESULTATS

### V. 3. 1. Mise en évidence d'un complexe IB7₁₄-*n*B3 par spectrométrie de masse

L'ionisation par électrospray (ESI), qui est connue pour apporter des informations directes d'assemblages supramoléculaire en phase gazeuse^{197,198}, a été utilisée sous infusion statique avec une seringue nanospray. Cette technique assure une consommation minimale de produit.

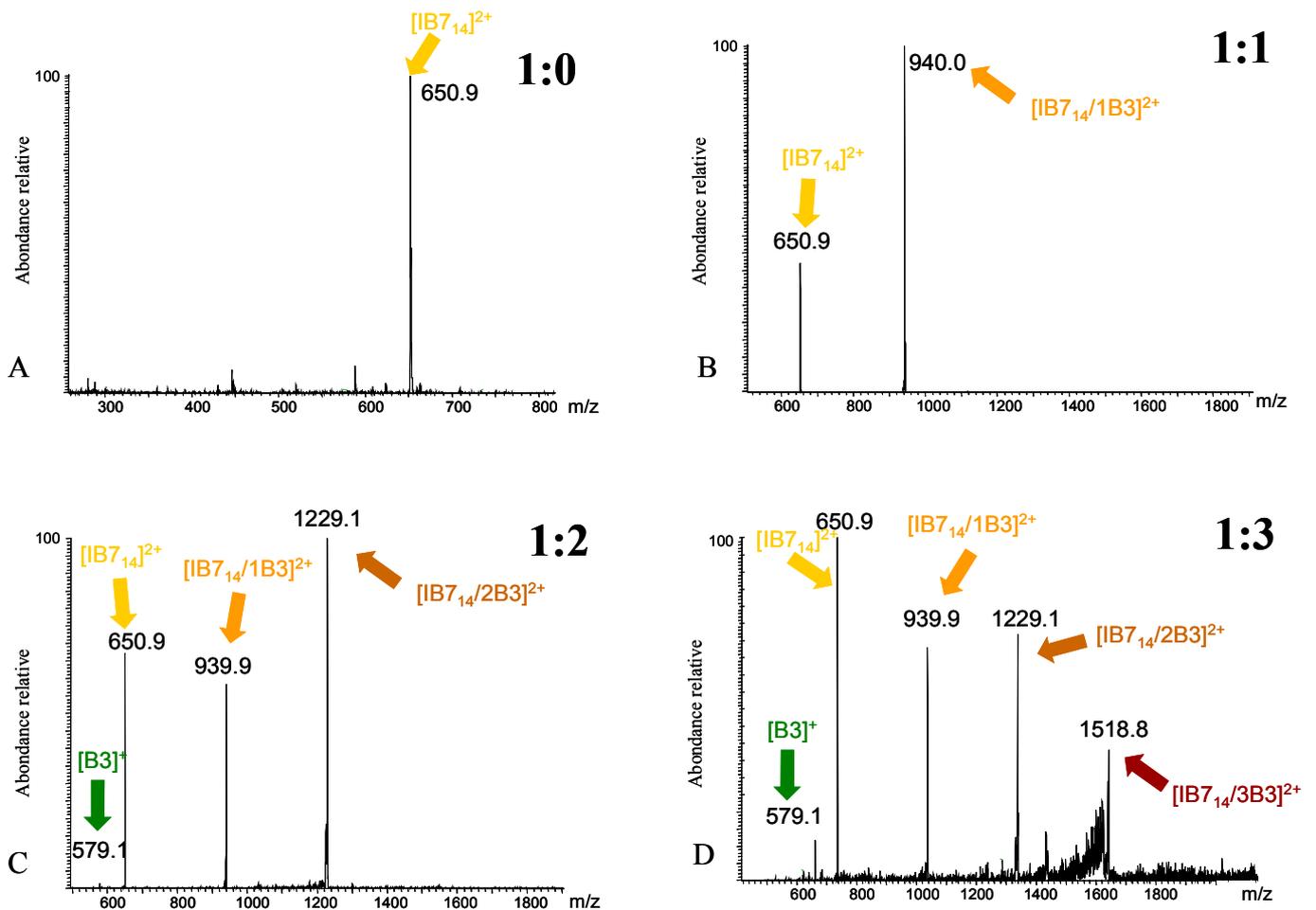
L'analyse du spectre de masse du peptide IB7₁₄ seul dans ces conditions d'ionisation conduit à un pic de rapport  $m/z$  de 650,9 qui correspond à la masse doublement protonée du peptide (figure V-1, A, masse exacte calculée = 650,83).

Le spectre de masse d'un mélange peptide-tanins conduit à l'observation de complexes non covalents IB7₁₄-B3. La trappe à ions nous permet de sélectionner sélectivement une espèce de masse précise (à plus ou moins 3 daltons), dans le but d'enrichir spécifiquement un composé. D'un mélange IB7₁₄-B3 d'un rapport molaire 1/4, ont pu être isolées dans la trappe à ions, les masses correspondant aux complexes de stœchiométrie 1/1, 1/2 et 1/3 (figure V-1, B à D respectivement). Un faible signal pouvant émaner du complexe 1/4 est également isolé, en revanche aucun spectre n'a pu être retenu par fragmentation ultérieure afin de s'assurer de son identification (par CID).

L'espèce doublement chargée correspondant au complexe IB7₁₄-B3 1/1 d'un rapport  $m/z$  de 940,0 (masse exacte de 939,91) a ainsi pu être isolée. Afin de lever l'indétermination entre l'existence du complexe formé par un peptide et un dimère ou simplement un artefact, nous l'avons fragmenté par collision énergétique. La conséquence de cette fragmentation est la réapparition d'un pic à la masse du peptide doublement chargé ( $m/z$  650,9, figure V-1, B). Le signal correspondant au dimère B3 mono-protoné n'apparaît pas sur le spectre, en raison du faible rendement d'ionisation du polyphénol. Ce complexe 1/1 IB7₁₄-B3 commence à se

désassembler pour une énergie de collision de 15% et se retrouve entièrement fragmenté pour une énergie de collision de 19% (à noter que l'espèce doublement chargée IB7₁₄ est fragmentée pour une énergie maximale de 28%). Ainsi nous confirmons l'existence du complexe équimolaire entre ces deux composés.

Des expériences similaires ont été réalisées pour toutes les autres masses sélectionnées, dont le résultat de la fragmentation est montré figure V-1, C et D. Le complexe 1/2 IB7₁₄-B3 est observé au rapport m/z 1229,1 (masse calculée de 1228,98). Cette masse est répartie en une espèce IB7₁₄ doublement chargée [IB7₁₄ + 2H]²⁺ et le tanin B3 mono chargé [B3 + H]⁺. La figure C montre le résultat issu de la fragmentation où l'on retrouve le complexe 1/1 (939,9), le peptide doublement protonné (650,9) ainsi que le signal du B3 (579,1). Le comportement du complexe 1/3 est similaire sous l'effet du CID, de telle sorte que la fragmentation fait apparaître les complexes 1/2 (1229,1), 1/1 (939,9) ainsi que les espèces mono-moléculaires [IB7₁₄ + 2H]²⁺ (650,9) et [B3 + H]⁺ (579,1) (figure D). Il est intéressant de noter que les complexes 1/1 et 1/2 commencent à se dissocier pour une énergie de collision de 15 et 18% respectivement, alors que le complexe 1/3 n'est altéré que pour une énergie de 22%.

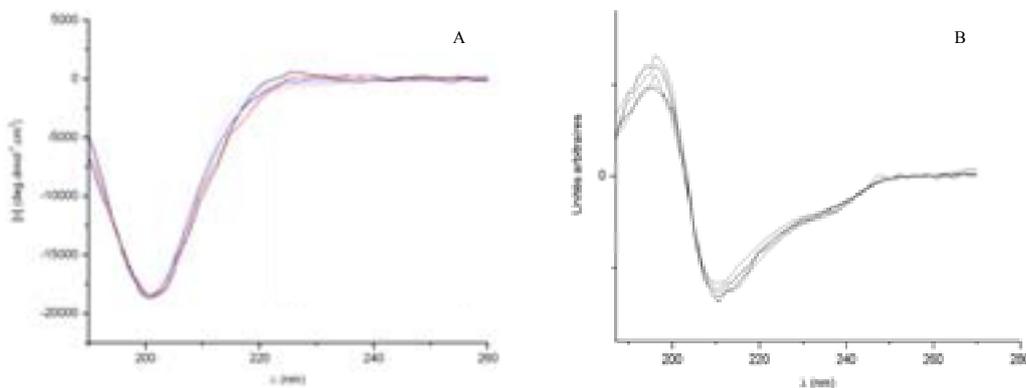


**Figure V-1 :** Spectrogramme de masse obtenu par électrospray avec une source nanospray pour IB7₁₄ seul (A) et dans un mélange IB7₁₄-B3 à différents rapports molaires 1/1 (B), 1/2 (C) et 1/3 (D).

Grâce à ces résultats nous montrons l'existence de différents degrés de complexation ainsi que de la multiplicité des sites de liaison. Mais nous ne pouvons pas conclure sur le nombre de sites de manière quantitative car nous ne sommes plus dans des conditions naturelles. Il est intéressant de remarquer que dans la chambre d'ionisation, le complexe tanin-protéine se trouve sous forme de gouttelettes dé-solvatées. Or, enlever l'eau maintient une forte affinité entre ces deux produits ce qui nous permet de conclure sur la nature des liaisons intermoléculaires. Celles-ci sont davantage de type liaison hydrogène et interaction électrostatique que des interactions de type hydrophobe^{80,178}.

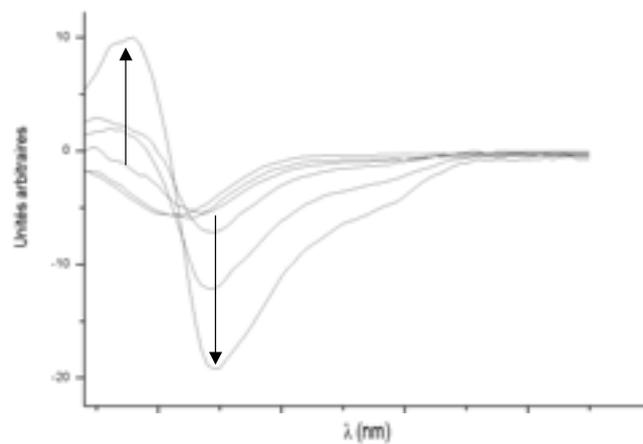
### V. 3. 2. Dichroïsme Circulaire : Influence de B3 sur la structure secondaire du peptide IB7₁₄

La figure V-2 représente les spectres dichroïques des acteurs de l'astringence séparément. La structure secondaire du peptide IB7₁₄ (A) est un mélange à 34% d'hélice de type II et 66% de conformation aléatoire (chapitre IV). Les spectres (B) sont caractéristiques de tanins avec deux bandes coton spécifiques des centres chiraux des unités catéchiques.



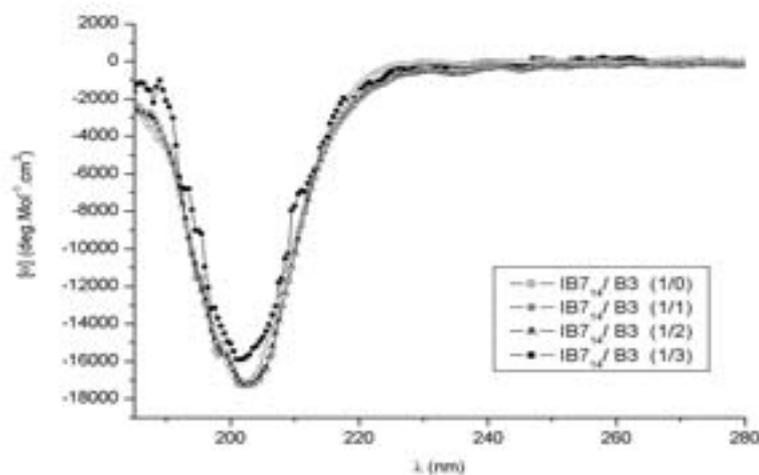
**Figure V-2 :** Spectres dichroïques du peptide IB7₁₄ seul à différentes concentrations (A) et du dimère B3 seul pour des concentrations variant de 0,025 à 0,6 mM (B) dans de l'eau.

L'ajout de différentes quantités successives de dimère B3 (correspondant à des concentrations de 0,025 à 0,6 mM) à une quantité fixe de peptide (0,200 mM) donne les spectres dichroïques suivants, figure V-3 :



**Figure V-3 :** Spectres dichroïques de différents mélanges IB₇₁₄-B3 (de 1/0 à 1/3). Les flèches indiquent les concentrations croissantes en B3.

Sur la figure V-3 nous pouvons observer la contribution du peptide ainsi que celle du B3. Evidemment, plus la concentration en B3 augmente plus le spectre ressemble au signal du tanin seul. Afin de remonter à la structure secondaire du peptide IB₇₁₄ il est indispensable de supprimer la réponse dichroïque du tanin. Pour cela, les spectres du tanin seul aux même concentrations que celles utilisées pour les différents rapports molaires (spectres figure V-2 (B)), sont soustraits aux spectres obtenus figure V-3. Nous obtenons ainsi la figure V-4, normalisée par la concentration et ne représentant plus que la réponse dichroïque du peptide.



**Figure V-4 :** Spectre dichroïque d'un mélange dimère-peptide. A une quantité fixe de peptide IB₇₁₄ (200 μM) différentes quantités de dimère B3 ont été rajoutées (IB₇₁₄/B3 1/0, 1/1, 1/2, 1/3). Les résultats sont représentés en ellipticité molaire corrigée de la contribution en dimère B3.

Toutes les courbes sont superposées à l'exception d'une légère déviation du spectre représentant le peptide au sein d'un mélange 1/3 (IB7₁₄/B3). Cependant, il faut noter que l'ajout successif de B3 produit une dilution du peptide de plus en plus importante. De sorte que dans ce dernier cas, la déviation du spectre est principalement due à une diminution du rapport signal/bruit provoquée par la dilution.

Les spectres ont été déconvolués de la même manière que pour le peptide seul en utilisant des protéines de structure secondaire connues. On retrouve une répartition identique au peptide seul (comme définit chapitre IV) avec 34% d'hélice de type II et 66% de pelote statistique étendue. L'addition de B3 ne modifie donc pas l'arrangement conformationnel du peptide. L'autre hypothèse quant à ce résultat serait l'absence d'affinité des deux composés l'un pour l'autre auquel cas aucune interaction n'est observable. Mais les autres techniques utilisées, présentées par la suite, confirment l'existence d'un complexe de forte affinité entre tanin et peptide riche en proline.

### V. 3. 3. RMN

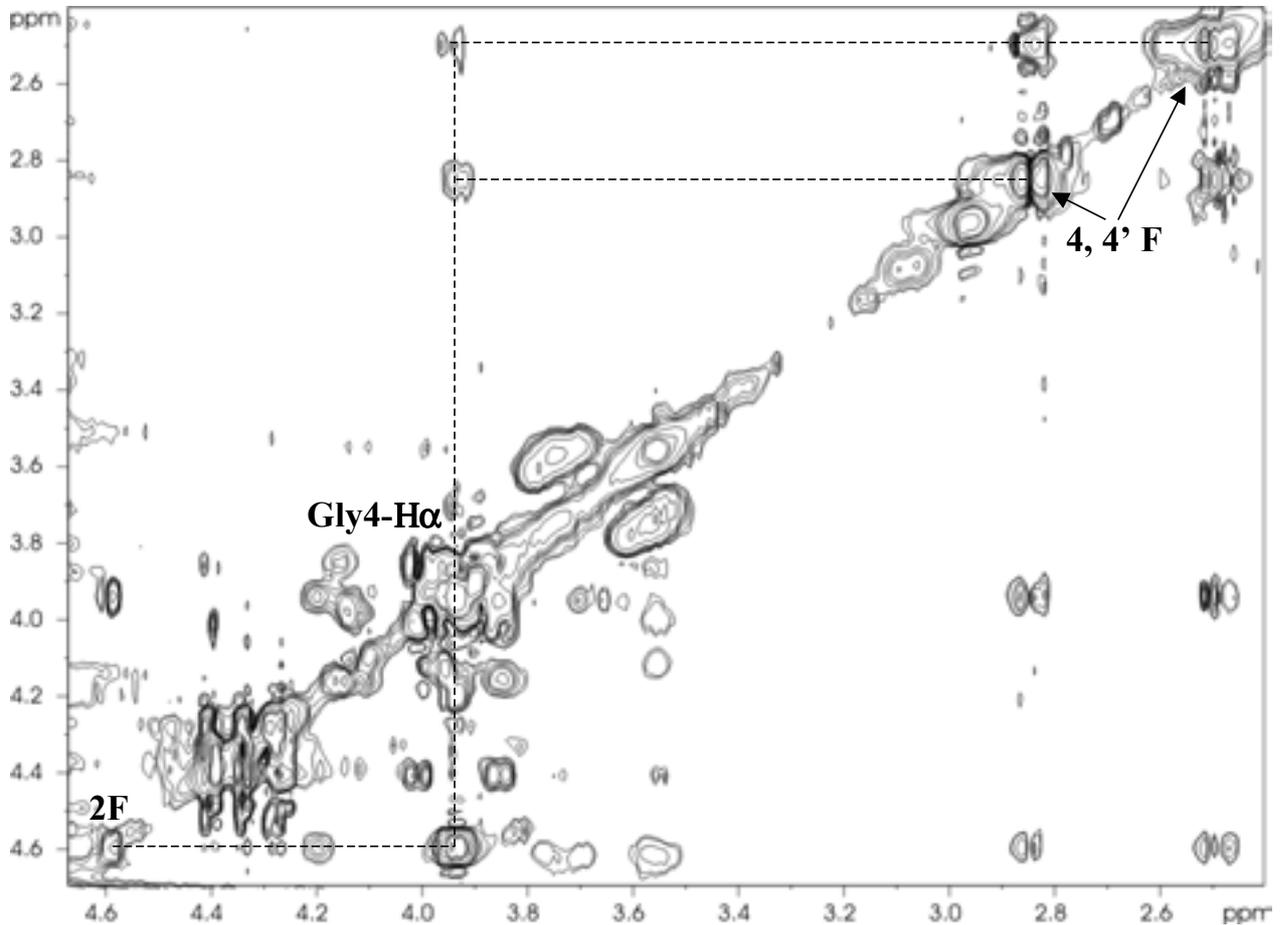
Afin de déterminer la conformation spatiale du peptide en présence de tanin B3, nous avons utilisé la batterie classique des expériences bi-dimensionnelles suivie d'une dynamique moléculaire. De plus, les variations de valeurs de déplacement chimique des protons du peptide ainsi que des coefficients de diffusion transversale du complexe, obtenues par RMN apportent des informations sur les paramètres thermodynamiques et sur la taille du complexe.

#### V. 3. 3. a. Informations structurales

Une procédure similaire à l'étude structurale du peptide seul a été entreprise pour le cas des complexes tanin-protéines. Aucune variation significative des contraintes RMN n'a été observée pour le peptide en absence ou en présence de B3. Les mêmes fichiers de contraintes (NOEs et ³J) sont donc utilisés pour déterminer la structure de départ du peptide en dynamique moléculaire. Quant au dimère B3, nous utilisons le fichier de contraintes défini par Barathieu¹⁹⁹.

En revanche, sur les cartes ROESY apparaissent des taches de corrélations intermoléculaires entre les résidus Gly4 et Lys5 du peptide IB7₁₄ et le cycle pyrane du dimère B3 (annexe 4). La figure V-5 montre quelques-uns un des effets NOEs entre Gly4 du peptide

et les protons 4F et 4'F du B3. 5 nouvelles contraintes inter-molécules sont rajoutées afin de réaliser une dynamique moléculaire (§ V.3.4).



**Figure V-5 :** Agrandissement du spectre ROESY en mode contour, représentant la région H $\alpha$ -H $\alpha$ . Seuls les effets NOEs entre Gly4-H $\alpha$  et 2F, 4F et 4'F du B3 sont présents dans cette partie du spectre.

### V. 3. 3. b. Association IB7₁₄-B3

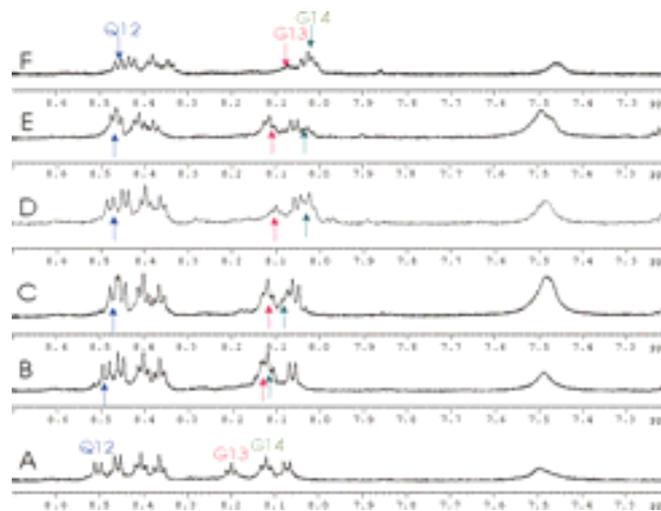
Afin de valider l'existence d'un complexe et d'affirmer sa stœchiométrie d'autres types d'expériences ont été entreprises. En effet, la démonstration d'une possible interaction par RMN n'est pas nouvelle. Etant donné que la valeur des déplacements chimiques dépend de l'environnement électronique des noyaux, le contact de B3 au peptide se traduirait par un décalage en déplacement chimique des protons impliqués dans l'interaction. De la même

manière, la valeur des coefficients de diffusion translationnels permettent de remonter à la constante de dissociation ainsi qu'aux nombres de sites de liaison des polyphénols sur le peptide.

Différentes solutions dans lesquelles le rapport IB7₁₄/B3 varie (de 1/0, 1/7), ont été soumises à des expériences de RMN mono ou bidimensionnelles.

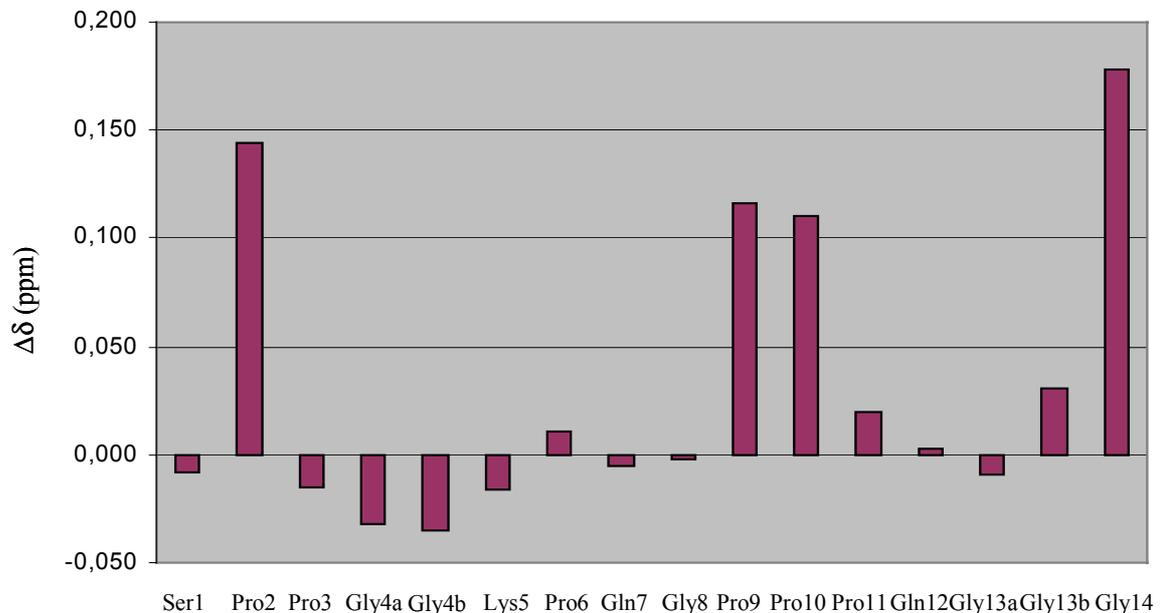
### **Variation des déplacements chimiques des protons de IB7₁₄ en présence de B3**

D'une manière générale, les valeurs des déplacements chimiques des protons peptidiques sont influencées par la présence de tanins, influence qui dépend de la concentration. Un exemple de l'évolution des déplacements chimiques des protons NH du peptide IB7₁₄ en fonction du rapport molaire B3/IB7₁₄ est présenté dans la figure V-6. La glutamine 12 est très peu influencée par la présence de dimère B3 contrairement aux glycine 13 et 14 qui subissent un fort décalage vers les hauts champs. En revanche, les protons du B3 ne subissent aucune variation de leur déplacement chimique au cours de la titration.



**Figure V-6 :** Spectre RMN 1D représentant la région des groupements amines du peptide IB7₁₄. A) spectre du peptide seul ; B) spectre du peptide en présence de B3 pour un rapport molaire de 1/1 ; C) 1/2 ; D) 1/3 ; E) 1/4,5 ; F) 1/6,6.

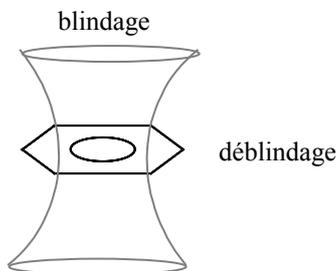
Les déplacements chimiques de tous les protons du peptide à chaque rapport molaire sont répertoriés en annexe 5. La figure V-7 répertorie l'ensemble des variations de déplacements chimiques ( $\Delta\delta$ ) des protons  $H\alpha$  du peptide.



**Figure V-7 :** Variations des déplacements chimiques des protons  $H\alpha$  de chaque résidu du peptide IB7₁₄.  $\Delta\delta$  représente la différence entre le déplacement chimique du peptide seul et du peptide complexé au tanin B3 (rapport molaire 1/7).

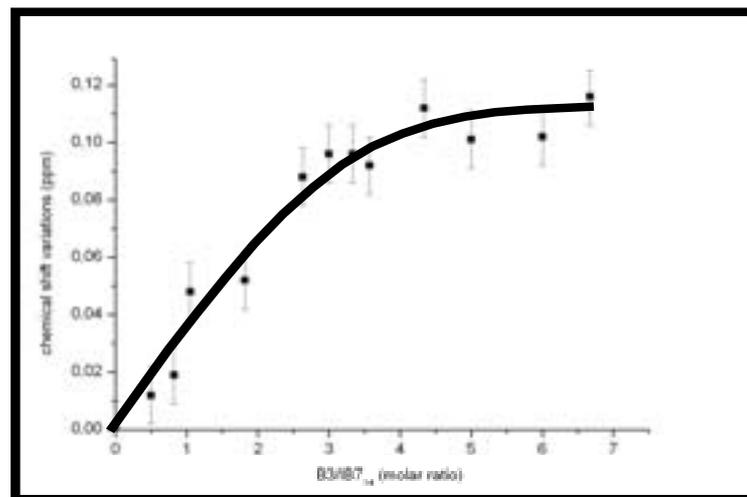
D'après la figure V-7, nous pouvons remarquer que parmi tous les protons du peptide certains semblent davantage influencés par la présence de B3. Les prolines précédentes des prolines (incluses dans un motif X-**Pro**-Pro) subissent une forte variation de déplacement chimique avec l'ajout de B3, contrairement aux prolines isolées. Etant donné que la variation de placement chimique traduit un changement d'environnement électronique, elle traduit le contact avec le tanin ou alors un changement conformationnel du peptide. Etant donné qu'il n'y a aucune modification des  3J  comme des effets NOE, cette variation découle donc de la fixation des tanins sur le peptide. Nous vérifions ainsi l'hypothèse déjà formulée concernant l'importance des motifs X-**Pro**-**Pro** pour l'interaction avec les tanins¹⁵⁵. D'autre part, si nous considérons ces protons par groupes nous distinguons 3 groupes majeurs subissant un fort décalage en présence de B3 et dépendant de leur position sur le peptide : de Pro2 à Pro6 pour

un premier groupe, de Pro9 à Pro11 pour le second, et Gly13 et 14 pour le dernier. Il est tout à fait intéressant de noter que les déplacements chimiques des prolines et des glycines C-ter sont blindés alors que ceux des autres résidus sont déblindés.



En s'approchant du peptide, les groupements aromatiques du polyphénol modifient la densité électronique des protons du peptide de telle sorte que certains vont se retrouver blindés (comme les prolines 2, 9 et 10 ainsi que les glycines 13 et 14, alors que d'autres seront déblindés. Les protons éloignés des zones de contacts ne seront quant à eux que faiblement perturbés par la présence de B3.

La figure V-8 représente la variation de déplacement chimique ( $\Delta\delta$ ) du  $H\alpha$  de la proline 9 en fonction du rapport molaire IB7₁₄/B3 :



**Figure V-8 :** Evolution de la variation du déplacement chimique ( $\Delta\delta$ ) du proton  $H\alpha$  de la proline 9 du peptide IB7₁₄ en fonction de la quantité de dimère B3 ajouté.

Au-delà d'un rapport molaire B3/IB7₁₄ de 3-4, l'addition de procyanidine ne modifie plus la valeur des déplacements chimiques de manière significative. La courbe a été traitée selon l'équation [7] afin de remonter à la constante de dissociation du complexe IB7₁₄/nB3 et aux nombres de sites de liaison sur le peptide. Au préalable, nous avons défini la constante d'association des dimères B3 entre eux en utilisant la méthode décrite par Baxter *et al.*¹¹⁴. Les

courbes ont été paramétrées selon l'équation [6]. Par la variation du déplacement chimique nous obtenons une constante d'association  $K_a = 4 \text{ M}^{-1}$  bien en dessus des concentrations utilisées pour l'interaction avec les protéines.

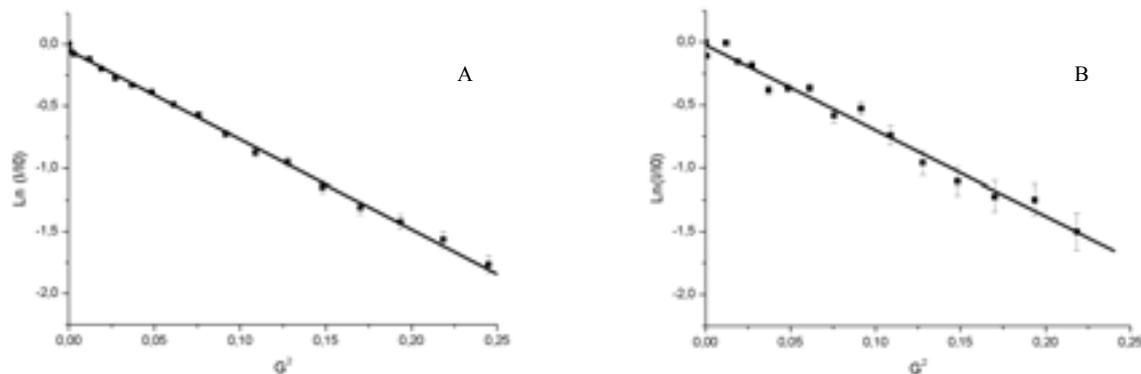
Nous obtenons alors la constante de dissociation pour le proton  $\text{H}\alpha$  de la proline 9 telle que  $K_d = 1,9 \cdot 10^{-4} \text{ M}$ , avec un nombre de site,  $n = 3,2$ . L'analyse a été répétée pour d'autres résidus dont les résultats sont donnés tableau V-1 :

Résidus	$K_d$ (mM)	$\Delta\delta_{\max}$	n	$\chi^2$
Pro2- $\text{H}\alpha$	8.15	0.228	3.0	0.028
Pro2- $\text{H}\beta$	3.00	0.085	3.5	0.015
Pro6- $\text{H}\alpha$	0.28	0.067	2.6	0.003
Pro9- $\text{H}\alpha$	0.19	0.111	3.2	0.013
Gly13-NH	2.46	0.186	3.2	0.018

Tableau V-1 : Données thermodynamiques du complexe IB7₁₄-B3.  $K_d$ , n,  $\Delta\delta_{\max}$  sont obtenus par les variations de déplacements chimiques en RMN, après un paramétrage par la méthode des moindres carrés selon l'équation [7]. Seulement les résultats pour quelques résidus sont représentés.

### **Variation du coefficient de diffusion transversal du complexe en fonction du rapport molaire IB7₁₄/B3**

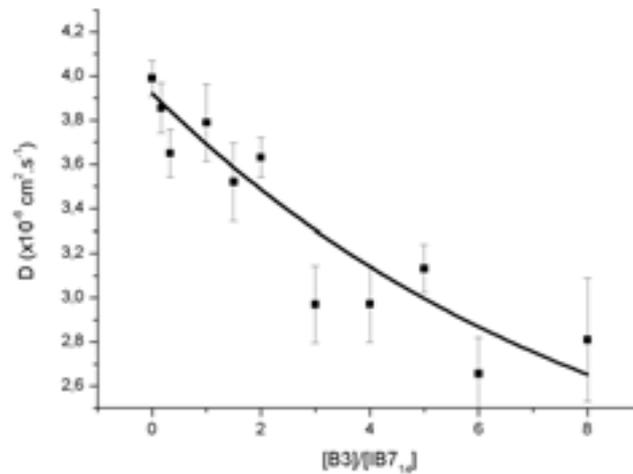
Les coefficients de diffusion sont mesurés grâce à l'application d'une série de gradients, comme explicité dans le chapitre II.3.4. L'équation [4] (chapitre II) montre la relation existant entre l'intensité du signal obtenu et la puissance du gradient appliqué. La figure V-9 montre ainsi l'évolution du signal en fonction du gradient appliqué lors d'une expérience où le peptide est seul (A) et pour un rapport molaire de 1/8, IB7₁₄/B3 (B).



**Figure V-9 :** Variation de l'intensité du He de la lysine en fonction des gradients appliqués lorsque le peptide est seul à 1,5 mM (A) ou en présence de 12 mM de B3 (B). La pente de la droite nous permet d'accéder au coefficient de diffusion transversal respectivement  $4,0$  et  $2,8 \times 10^{-6} \text{ cm}^2 \cdot \text{s}^{-1}$ .

D'après l'équation de Stokes-Einstein (Eq. [3], chapitre II),  $D$  étant inversement proportionnel à la taille de la molécule, nous pouvons retrouver le rayon de giration du complexe. Dans le cas présent, plus il y a de B3 qui se fixent sur le peptide plus le complexe augmente en taille donc plus  $D$  doit diminuer. Un coefficient de diffusion de  $4,0 \cdot 10^{-6} \text{ cm}^2 \cdot \text{s}^{-1}$ , obtenu pour le peptide seul (1,5 mM) implique un rayon de  $5,5 \text{ \AA}$ . Alors qu'en présence de 12 mM de B3 (rapport molaire de  $1/8 \text{ IB}_{714}/\text{B3}$ ) le coefficient de diffusion descend à  $2,8 \cdot 10^{-6} \text{ cm}^2 \cdot \text{s}^{-1}$  ce qui implique un rayon de giration de  $7,8 \text{ \AA}$ .

Il est intéressant de noter que la titration de  $D$  en fonction du rapport molaire  $\text{B3}/\text{IB}_{714}$  nous ramène aussi à la constante de dissociation du complexe ainsi qu'aux nombres de sites de fixation, de la même manière que la variation de déplacements chimiques. La figure V-10 reporte l'évolution du coefficient de diffusion du peptide en fonction de la quantité de B3 ajouté à la solution peptidique. Dans la mesure où la température ainsi que la viscosité du milieu sont considérées comme constantes, cette évolution traduit ainsi la variation de la taille du complexe, amenant au nombre de B3 fixés sur le peptide.



**Figure V-10 :** Variation du coefficient de diffusion du peptide IB7₁₄ en fonction de la quantité de dimère B3 ajouté. La courbe en noir représente le meilleur paramétrage des données selon l'équation [7].

Un même type d'équation que précédemment où la variation de déplacement chimique est remplacée par la variation du coefficient de diffusion, défini par Luo *et al.*¹⁹⁶, est utilisé pour traiter les données. Nous avons obtenu une constante de dissociation  $K_d \approx 2$  mM et un nombre de sites,  $n = 3$ .

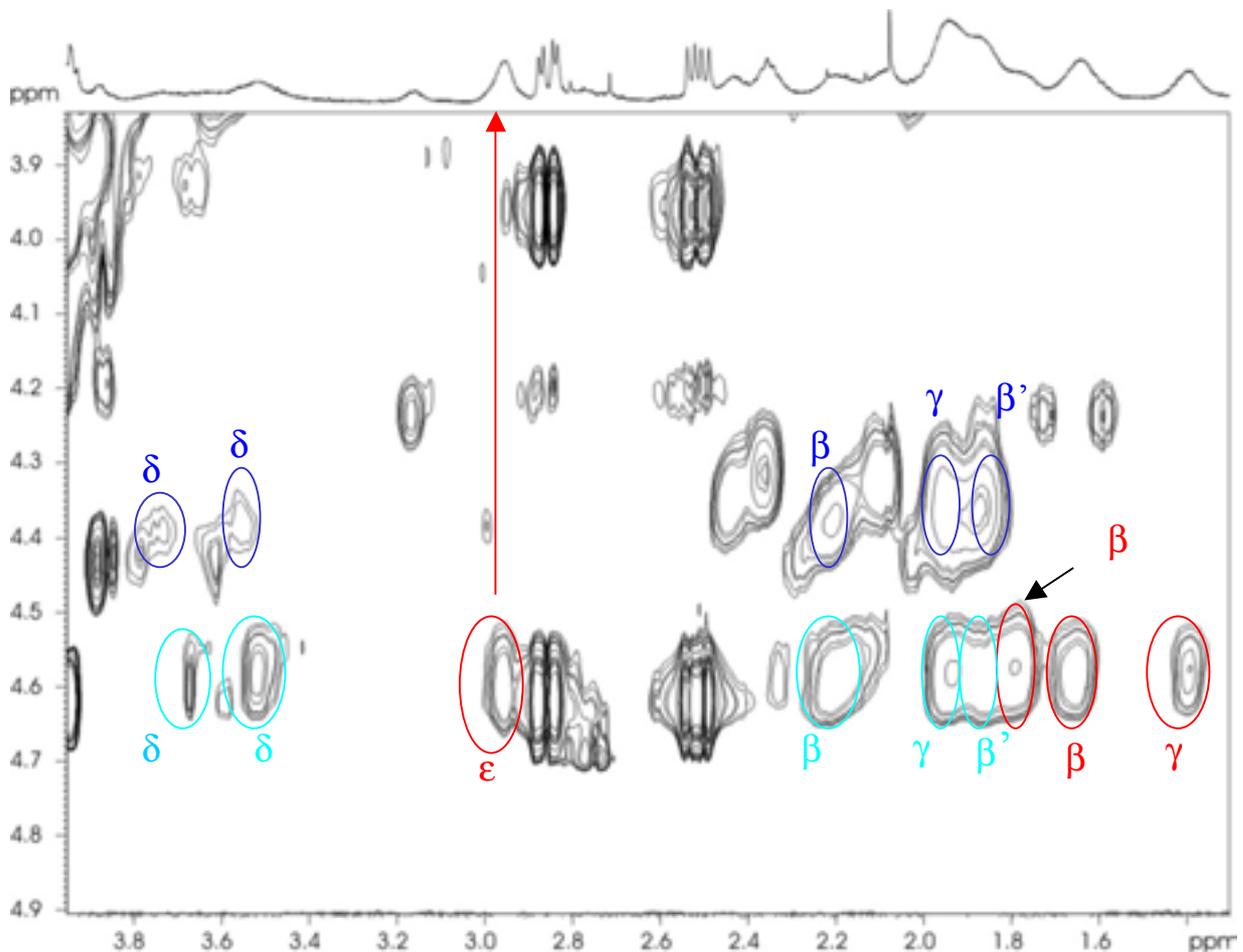
### V. 3. 3. c. Association IB7-B3

L'analyse de l'interaction IB7-B3 s'est focalisée sur l'aspect thermodynamique ( $K_d$  et nombre de site) ainsi que sur la taille globale du complexe. Ces différents paramètres ont été obtenus par mesure du coefficient de diffusion et de la variation du déplacement chimique de groupes de protons à différents rapports molaires B3/IB7 (de 0/1 à 21/1).

#### Variation des déplacements chimiques de groupes de protons de IB7 en présence de B3

L'attribution des protons des 59 résidus aminés de la protéine n'a pas encore été réalisée. Néanmoins, l'attribution complète des protons de IB7₁₄ a servi de référence pour la détermination des déplacements chimiques de groupes d'acides aminés identiques de IB7. En effet, chaque résidu de la protéine n'est pas identifiable individuellement, en raison du

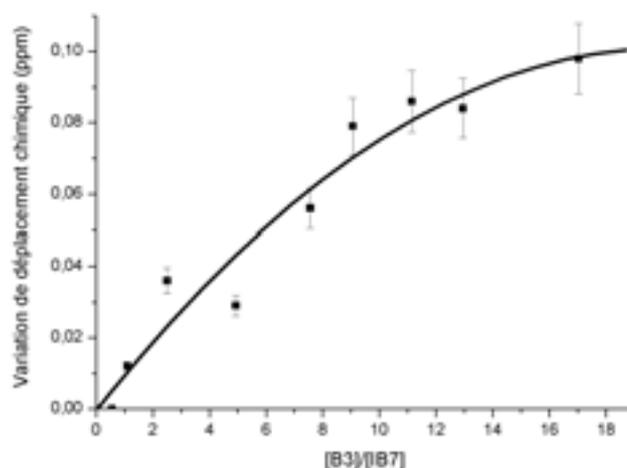
chevauchement, donc nous définissons des groupes correspondant à un type d'acide aminé comme le montre la figure V-11. Ainsi un groupe de lysines a pu être identifié (en rouge sur la figure) principalement par la valeur du proton  $\gamma$ , très blindé et sortant à 1,40 ppm. Deux groupes de prolines sont également différenciés, les prolines précédant une proline (Pro-p, en bleu clair sur la figure) dont la valeur du  $H\alpha$  est environ de 4,6 ppm et les prolines précédant un autre résidu (Pro-x, en bleu foncé sur la figure) dont le  $H\alpha$  résonne à environ 4,4.



**Figure V-11 :** Agrandissement d'une région de la carte TOCSY de la protéine IB7, pour un rapport molaire B3/IB7 de 13, montrant un groupe de lysines (en rouge) et deux groupes de prolines (en bleu). La flèche en rouge indique la projection du proton  $\epsilon$  de la lysine sur le spectre 1D.

La variation du déplacement chimique de groupes lysine, Pro-p et Pro-x ainsi que glutamine de IB7, en fonction du rapport  $[B3]/[IB7]$  a été suivie. Seule la courbe de titration

de Pro-p est présentée en figure V-12. Les signaux des prolines sont beaucoup plus éclatés à un rapport protéine-tanin élevé. Ceci montre l'importance des résidus prolines dans l'interaction puisque les valeurs des déplacements chimiques sont fortement modifiées.



**Figure V-12 :** Variation du déplacement chimique du H $\alpha$  du groupe Pro-p, en fonction du rapport molaire B3/IB7. La courbe représente le meilleur paramétrage des données selon l'équation [7].

Pour chacun des groupes d'acides aminés,  $K_d$  et le nombre de sites ont été déterminés en paramétrant les points expérimentaux avec l'équation [7]. L'ensemble des valeurs est reporté dans le tableau V-2.

Résidus	$K_d$ (mM)	n	$\chi^2$
NH Gln	0,19	9,9	0,03
H $\alpha$ Pro-p	240	11,5	0,07
H $\beta$ Lys	0,1	16,7	0,01
H $\beta'$ Pro-x	0,19	16,4	0,01

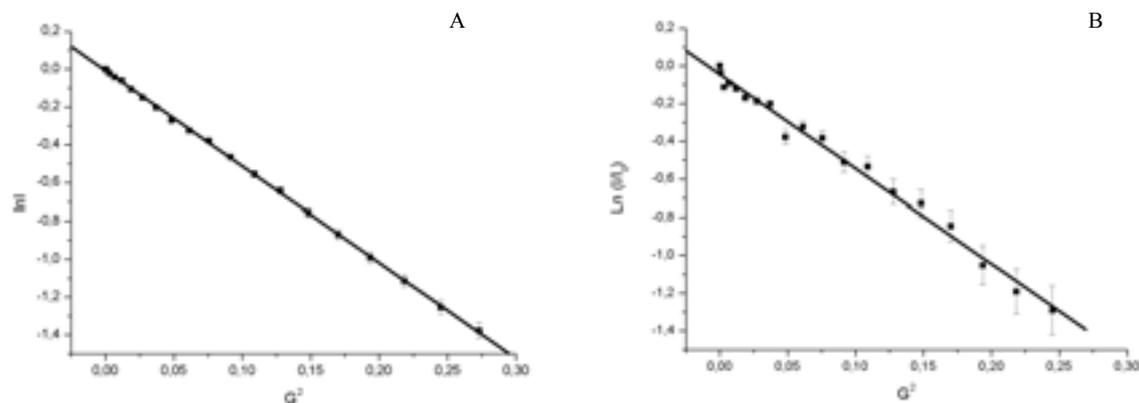
**Tableau V-2 :** Données thermodynamiques du complexe IB7/B3.  $K_d$  et n sont obtenus par les variations de déplacements chimiques en RMN, après un paramétrage par les moindres carrés selon l'équation [7].

Dans le cas de la glutamine, le paramétrage de la variation de déplacement chimique du NH indique un nombre de sites de liaisons de 10. Un résultat comparable est obtenu pour le H $\alpha$  de la Pro-p ( $n = 11$ ). En revanche, en considérant le proton H $\beta$  de la lysine, le nombre de sites de liaisons est alors d'environ 17. Enfin, les variations de déplacement chimique du proton H $\beta'$  de la Pro-x conduisent à une valeur de nombres sites de liaisons de 16,5.

Aucune variation des déplacements chimiques des protons de B3 n'est observable.

### **Variation du coefficient de diffusion transversal du complexe en fonction du rapport molaire IB7/B3**

Les coefficients de diffusion de la protéine ont été mesurés pour chaque rapport IB7/B3 (de 1/0 à 1/21). En revanche, à forte concentration en B3, le signal de la protéine devient difficilement perceptible, donc les derniers points ne sont pas reportés sur la courbe de titration. Le massif choisi correspond aux H $\epsilon$  de la lysine comme illustré par la flèche en rouge sur la figure V-11. En effet ces protons résonnent dans une région relativement isolée, ce qui facilite leur intégration. La figure V-13 montre la variation de l'intensité du signal du proton H $\epsilon$  de la lysine en fonction de la puissance des gradients appliqués pour la protéine seule à 1 mM (A) et en présence de B3 à 17 mM (B).



**Figure V-13 :** Variation de l'intensité du H $\epsilon$  de la lysine en fonction des gradients appliqués lorsque le peptide est seul à 1 mM (A) ou en présence de 17 mM de B3 (B). La pente de la courbe nous permet d'accéder au coefficient de diffusion transversal respectivement  $1,7$  et  $1,2 \times 10^{-6} \text{ cm}^2 \cdot \text{s}^{-1}$ .

Le coefficient de diffusion translationnel diminue avec l'ajout de quantités croissantes de B3 dans la solution de protéine, de  $1,7$  à  $1,2 \text{ E}^{-10} \text{ m}^2 \cdot \text{s}^{-1}$ , comme illustré dans la

figure V-14. D'après l'équation de Stokes-Einstein (Eq. [3], chapitre II) ceci implique une augmentation de la taille du complexe de 11500 Da.

De plus, la figure V-14 présente l'évolution du coefficient de diffusion de la protéine IB7 en fonction du rapport molaire IB7/B3 (obtenu de la même manière que pour le peptide seul ou complexé).

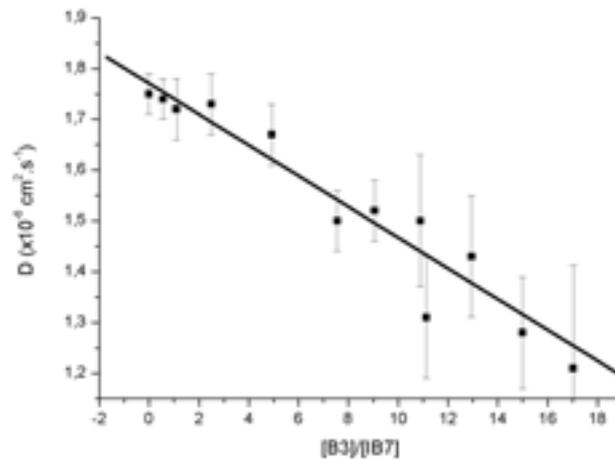


Figure V-14 : Variation du coefficient de diffusion de la protéine IB7 en fonction de la quantité de dimère B3 ajouté.

Le paramétrage de la variation de D par l'équation [7] conduit à une valeur de  $K_d$  de 0,7 mM, et un nombre de sites de liaison de 17.

### V. 3. 4. Dynamique Moléculaire

#### V. 3. 4. a. Dynamique moléculaire sous contraintes RMN

Nous avons réalisé, dans un premier temps, un calcul de structure sous contraintes RMN pour un rapport équimolaire 1/1 IB7₁₄/B3, de la même manière que précédemment (chapitre IV.2.3). La dynamique moléculaire est réalisée à 500 K et après minimisation, seulement 40 conformères sont retenus dans une échelle d'énergie d'environ 100 kJ.mole⁻¹.

Une analyse par cluster nous a informés de l'existence de 5 familles de structures répertoriées dans la figure V-15.

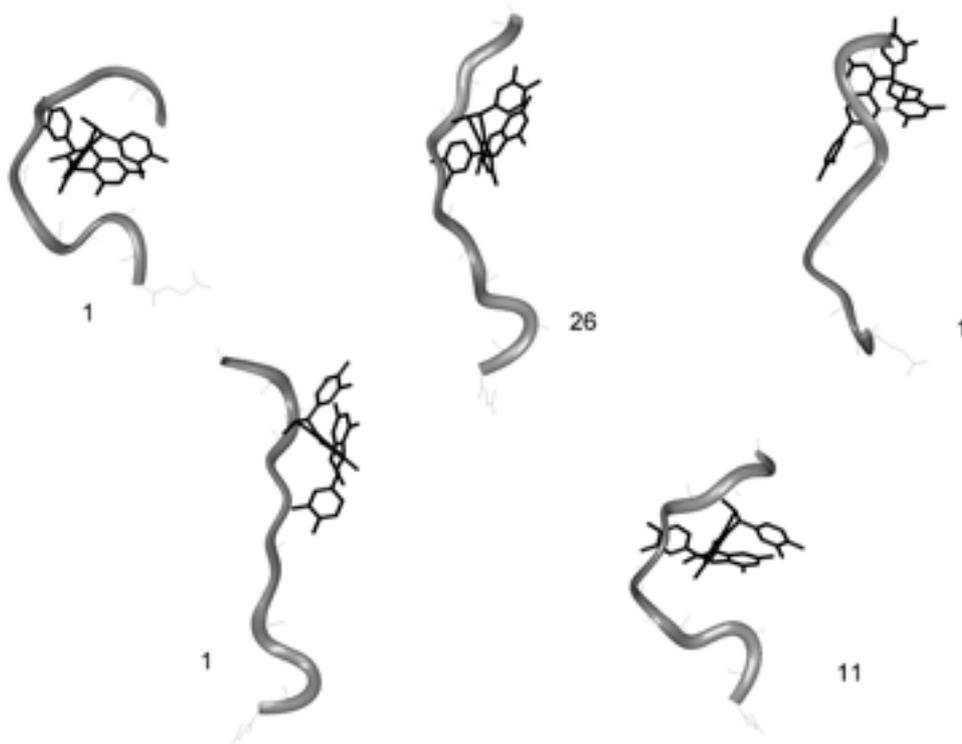


Figure V-15 : Conformations de plus basse énergie obtenue par dynamique moléculaire. Les 40 structures sont réparties en cinq familles, dont trois ne sont représentées que par une molécule et les deux autres par 11 et 26 respectivement.

Les 5 familles de structures sont inégales en population constitutive : trois d'entre elles sont composées par seulement une espèce et les deux autres sont composés de 11 et 26 conformères respectivement. La figure V-16 montre les 26 structures appartenant à la famille majoritaire.

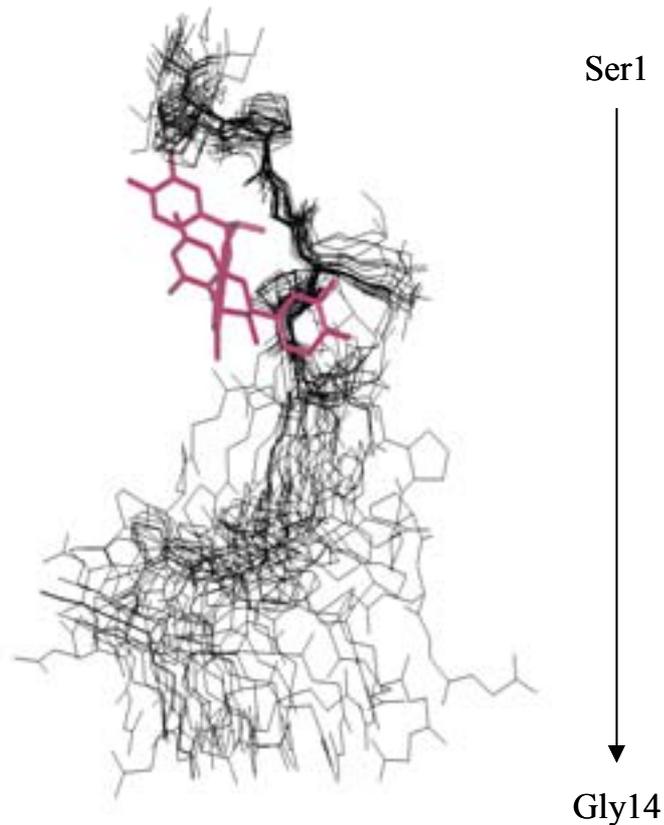


Figure V-16 : Superposition des 26 conformations du peptide IB7₁₄ en présence de B3, pour un rapport équimoléculaire. La dynamique moléculaire a été effectuée durant 2 ns, sous contraintes RMN. Le B3 (en bâtonnet mauve) est positionné par rapport aux effets NOEs observés en RMN.

Toutes ces structures sont superposables entre les résidus Pro2 et Pro6. Si l'on compare cette superposition avec celle effectuée en l'absence de B3 (figure IV-11), on remarque que la souplesse des conformations est extrêmement réduite lorsque le peptide est lié au tanin.

#### ***V. 3. 4. b. Dynamique moléculaire sans contraintes RMN***

Etant donné que nous n'avons que cinq contraintes NOEs intermoléculaires, toutes concernant un même site, nous avons choisi de réaliser une dynamique sans considérer ces effets NOEs.

Un premier jeu de calcul a été effectué pour un rapport équimoléculaire 1/1 IB7₁₄/B3. La conformation initiale du peptide est celle préalablement définie par dynamique en considérant les contraintes RMN. 5 dynamiques sont réalisées au cours desquelles le tanin est placé au départ en différents endroits autour du peptide, de manière à envelopper totalement le peptide par des tanins grâce à la multiplicité des expériences. Pour les 5 positions, une dynamique est effectuée, précédée par une minimisation permettant de « coller » le B3 au peptide. 3 dynamiques se concluent par une fixation forte des B3 autour du peptide. Ainsi, les régions affines fixent le B3 alors que les « non-affines » le rejette. 2 sites de liaisons sont ainsi mis en évidence : l'un se situerait de Pro2 à Pro6 et l'autre proche des résidus Pro9 à Pro11. Le premier site est le même que celui mis en évidence par RMN. Durant les 2 autres calculs le dimère est éjecté du peptide après seulement 1 ns.

Etant donné que la spectrométrie de masse propose un complexe multi-ligand, avec au moins trois tanins B3 fixés simultanément sur le peptide, nous avons réalisé un autre jeu de dynamique avec plus de 3 tanins par molécule de peptide : le système considéré comporte 4 dimères pour un peptide IB7₁₄. Les 4 dimères B3 sont positionnés à chaque extrémité du peptide de manière à englober une majeure partie de la surface du peptide. 2 dynamiques de 5 ns sont effectuées pour des positions de départ en B3 différentes. Au cours des 2 expériences indépendantes, un des quatre B3 est rejeté au bout de 2 ns, alors que les trois autres tanins se fixent au peptide. Les positions prises par les B3 sont quasiment identiques d'une dynamique à l'autre : une première procyanidine se lie entre les résidus Pro2 et Pro6, une seconde entre les Pro9 et Pro10 et la troisième entre les Gly13 et Gly14 (figure V-17).

Sur la figure V-17, il est notable que les tanins ne se fixent que sur l'une des faces du peptide.

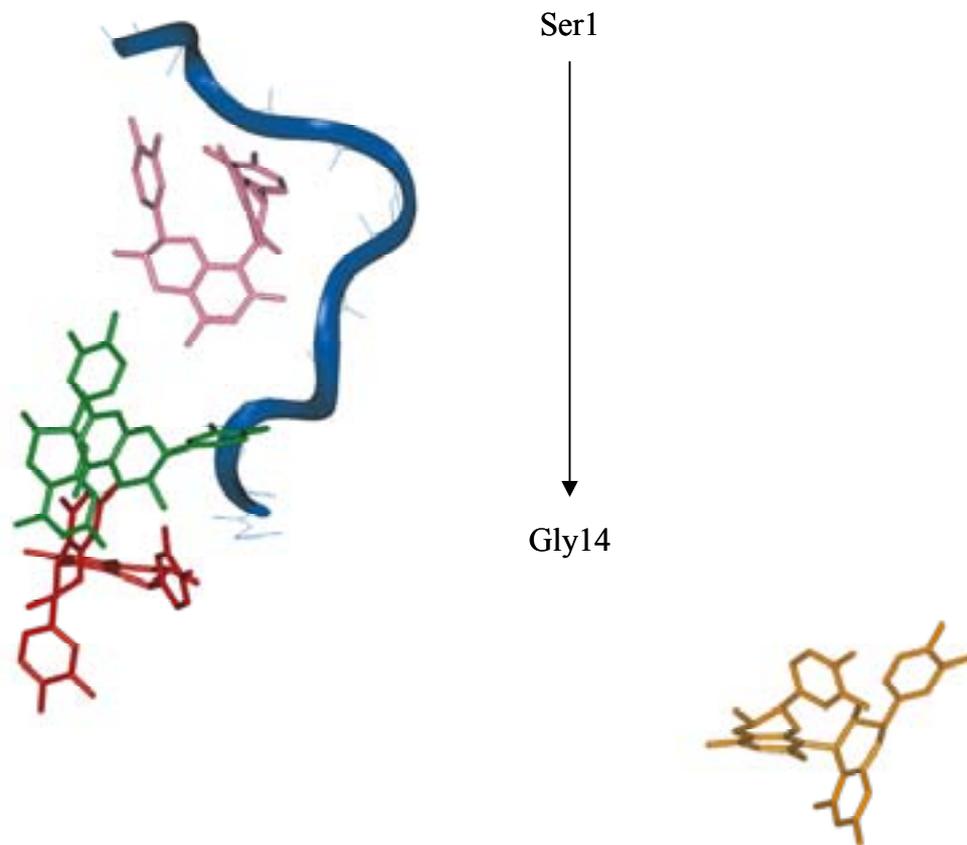


Figure V-17 : Structure du complexe IB7₁₄-B3 (rapport molaire 1/4) après une dynamique de 5 ns. La protéine est représentée en bleu (en « ruban »). Les trois B3 fixés sont représentés en bâtonnet (rose, vert et rouge). Le quatrième tanin B3 est éjecté du complexe dès 2 ns de dynamique.

La figure V-18 montre la superposition des conformations adoptées par le peptide à la fin de ces deux dynamiques ainsi que la conformation obtenue sous contrainte RMN.

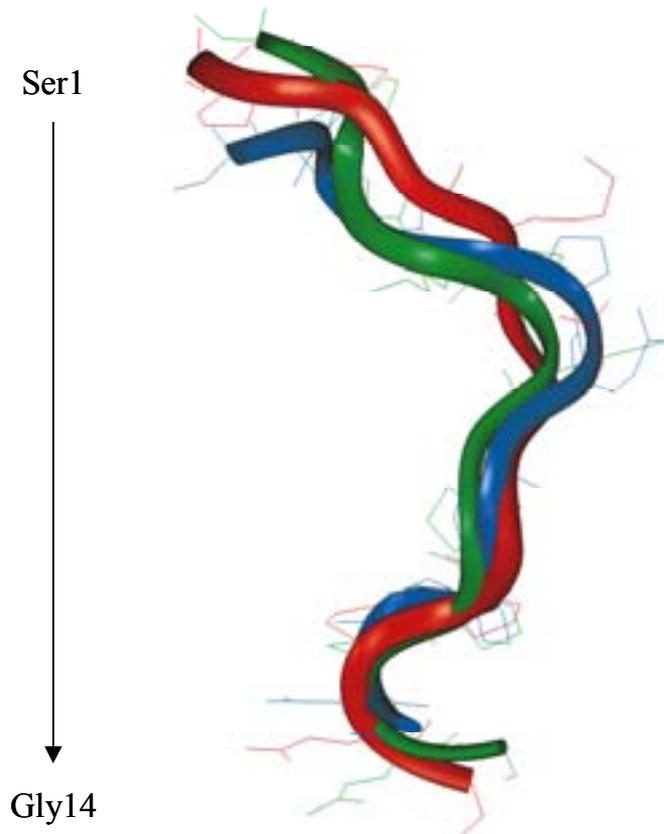


Figure V-18 : Superposition en « ruban » du peptide IB7₁₄ à la fin de la dynamique moléculaire sans contraintes (vert et bleu) et sous contraintes RMN (rouge). Les calculs ont été réalisés en présence de 4 tanins (non montrés).

Plusieurs résultats découlent de ces dynamiques :

- (i) Les trois conformations sont quasiment superposables.
- (ii) Le peptide semble stabilisé en une conformation privilégiée, probablement par formation de liaisons hydrogène entre les carbonyles des Pro2 et Pro6 et les fonctions OH des groupements phénols et catéchols des procyanidines. Nous pouvons également observer la formation de liaisons hydrogène entre un carbonyle du peptide et le 3-OH du cycle pyrane.
- (iii) De plus, les molécules de B3 maintiennent leur conformation durant toute la dynamique, *i.e.*, elles restent en conformation compacte, qui est la conformation majoritaire en solvant aqueux¹²⁹.

### V. 3. 4. c. Surfaces de lipophilie du peptide IB7₁₄

Etant donné que les tanins se fixent sur une même face du peptide, nous avons décidé de calculer la répartition hydrophile ou lipophile à la surface du peptide.

La structure de départ choisie pour les calculs est celle où trois tanins B3 sont en contact avec le peptide. La figure V-19 montre le profil amphiphile du peptide atome par atome, avec une partie hydrophile (en bleu) et une partie lipophile (en rouge). Sur la figure, l'interface entre ces deux zones est colorée en rose. Il est tout à fait remarquable que la structure du peptide soit séparée en deux faces opposées selon les propriétés amphiphiles. De plus la zone hydrophile du peptide est largement majoritaire et compte pour deux tiers.

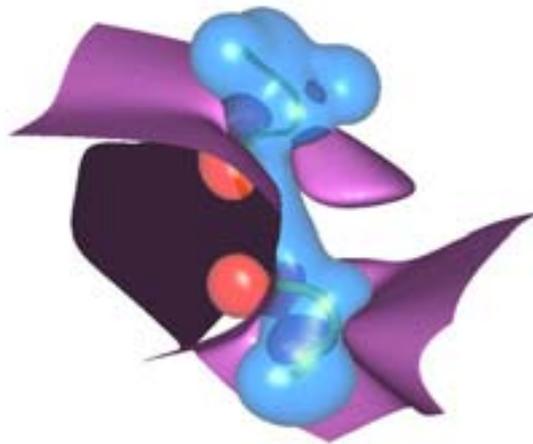


Figure V-19 : Représentation de la nature amphiphile du peptide IB7₁₄. Le domaine hydrophile, majoritaire, est en bleu et le domaine lipophile est en rose. L'interface entre ces deux régions est représentée en mauve.

Dès lors que le peptide et les dimères sont mélangés, dans 5 cas sur 6 les procyanidines B3 se localisent dans les poches hydrophiles (figure V-20).

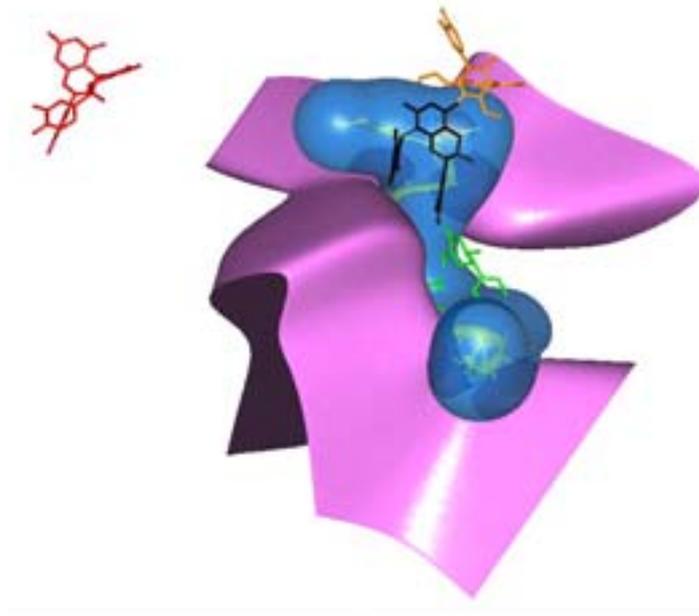


Figure V-20 : Répartition des dimères B3 en fonction des domaines d'hydrophilicité. Les procyanidines sont localisées dans les domaines hydrophiles (en bleu).

La comparaison entre les domaines d'hydrophilies du peptide et la répartition de B3 autour du peptide est significative. Les dimères sont toujours fixés dans les poches hydrophiles donc lorsque tous les sites sont occupés, il ne peut se fixer d'autres B3 : la fixation simultanée d'un maximum de 3 dimères par peptide provient de l'encombrement stérique. De cette manière nous sommes convaincus de la liaison des polyphénols principalement sur les atomes hydrophiles (tels que la majorité des H $\alpha$  des prolines, et les NH des glycines 13 et 14).

#### ***V. 3. 4. d. Détermination du coefficient de diffusion du peptide IB7₁₄ et du complexe IB7₁₄/B3 par Dynamique Moléculaire (DM)***

Par DM, le coefficient de diffusion de IB7₁₄ en absence de B3 varie de 3,5 à 1,0 .10⁻⁶ cm²s⁻¹. En revanche, le rayon de giration du peptide seul, mesuré à la fin de la

dynamique est toujours de  $7,8 \pm 0,1 \text{ \AA}$ . L'incertitude sur les valeurs de coefficient de diffusion obtenues serait due à la capacité du peptide à former des liaisons hydrogène avec le solvant, en conservant une taille identique d'une expérience à l'autre. Le peptide seul semble donc être instable dans l'eau et par conséquent le coefficient de diffusion est peu précis. En effet la courbe déduite de la relation d'Einstein (annexe 6) est loin d'être linéaire. La diffusion n'est donc pas isotrope. Il est de ce fait difficile de conclure tant sur la dimension du peptide que sur la valeur exacte de coefficient de diffusion d'après ces résultats. En revanche, en présence de B3, le coefficient de diffusion est assez précis de l'ordre de  $2,2 \pm 0,4 \times 10^{-6} \text{ cm}^2\text{s}^{-1}$  soit un rayon de  $9,0 \pm 0,1 \text{ \AA}$ . La condition de diffusion isotrope est respectée, car la courbe de variation du déplacement moyen de la molécule en fonction du temps est linéaire (annexe 6).

## V. 4. DISCUSSION SUR LA NATURE DU SITE ET L'AFFINITE

La multiplicité des techniques offre une compréhension de l'interaction Tanins-Protéines à différents niveaux : la stœchiométrie et l'affinité du complexe IB7₁₄/B3, comparées au complexe IB7/B3, ont été déterminées par SM, RMN et DM. La structure du peptide au sein du complexe, ainsi que l'identification des sites ont été mis en évidence par CD, RMN et DM. De là, la taille du complexe a été abordée par calcul de coefficients de diffusion par RMN et DM. Enfin, la nature de l'interaction a été démontrée par SM, RMN et DM. Ces différents aspects sont détaillés ci-dessous.

### V. 4. 1. La stœchiométrie et l'affinité du complexe

La stœchiométrie du complexe peptide IB7₁₄-tanins a été déterminée par trois différentes techniques : la SM, la RMN et la DM. Le peptide IB7₁₄ peut fixer trois dimères B3. Tout d'abord, en utilisant les conditions d'ionisation douces (ESI), la spectrométrie de masse nous a permis de mettre en évidence des complexes de stœchiométries 1/1, 1/2, et 1/3. Ces complexes sont très stables aux vues de l'énergie nécessaire à leur fragmentation et des conditions dénaturantes utilisées lors de l'expérience (bas pH et éthanol).

Les expériences de RMN montrent une dépendance des valeurs de déplacement chimique ainsi que de coefficients de diffusion en fonction du rapport IB7₁₄/B3. Le paramétrage de ces données par un modèle de fixation non coopératif développé par Baxter *et al.*¹²⁸ permet de mettre également en évidence la formation d'un complexe de stœchiométrie 1 peptide/3 tanins B3. De la même manière, la valeur de la constante de dissociation n'a pu être mesurée qu'approximativement : elle serait de l'ordre du millimolaire. Les résultats apparaissent tout à fait subjectifs étant donné la courbe de titration de certains résidus qui est inachevée.

Il est important de noter que  $K_d$  varie d'un facteur 50 autour du mM. En revanche, le nombre de sites de liaison du B3 sur le peptide est fixe, autour de 3. La large gamme de  $K_d$  obtenue peut avoir différentes raisons : d'une part la mesure des déplacements chimiques a été effectuée sur des cartes bidimensionnelles (TOCSY) où la résolution n'est pas excellente (2,5 Hz/pt). D'autre part, et ce qui nous semble le point majeur, l'équation [7] utilisée, a été définie pour un modèle où les sites sont indépendants les uns des autres. Or, il y a fort à croire que la fixation de tanins sur le peptide suit un modèle coopératif auquel cas l'équation n'est pas adaptée. L'amplitude des valeurs observées proviendrait donc d'une différence d'affinité d'un site à l'autre. De même, nous avons pu également montrer en spectrométrie de masse, que l'énergie nécessaire pour dissocier le complexe de stœchiométrie 1/3 est nettement supérieure à celle requise pour dissocier les complexes 1/1 et 1/2, suggérant l'existence d'une fixation à effet coopératif. L'utilisation d'un modèle prenant en compte un mode de fixation coopératif a effectivement déjà été suggéré^{182,200}. Par exemple, Hagerman *et al.* ont montré le mode coopératif de fixation de tanins par des titrations de fixation compétitive où le signal de SAB (marqué à l'iode 25) fixé aux polyphénols est déplacé en présence d'autres protéines¹. Mais le modèle coopératif n'a pas été défini, et il n'a pas été utilisé pour cette étude. Les valeurs de  $K_d$  que nous avons obtenues sont cependant en bon accord avec les valeurs déjà publiées de 2 à 33 mM pour un polyphénol similaire (l'épigallocatechine galloylée) et un peptide 19-mer¹⁸². En outre, afin d'éviter la précipitation de leurs complexes, les auteurs travaillent dans 10-20% de DMSO, alors que nous avons choisi l'utilisation de la HRMAS, technique permettant d'enregistrer des spectres de solutions colloïdales sans perte de signal ni de résolution. D'autre part, nous avons mesuré des coefficients de diffusion par RMN du peptide seul ou en présence de B3. Ces mesures nous ont permis d'accéder à une stœchiométrie identique du complexe de 1/3 (IB7₁₄/B3) et d'obtenir une constante de dissociation du complexe de l'ordre du millimolaire. Enfin, la dynamique moléculaire

effectuée sur un système comprenant 4 tanins B3 pour une molécule de peptide, conduit également à un complexe final de stœchiométrie 1/3.

En ce qui concerne la protéine entière IB7, des mesures de variations de déplacement chimique et de coefficients de diffusion en fonction de l'ajout de B3 ont été effectuées par RMN, à titre d'expériences préliminaires. La constante de dissociation obtenue, serait, comme dans le cadre du peptide, de l'ordre du millimolaire. Par mesure des variations de déplacements chimique, la stœchiométrie du complexe IB7/B3 obtenue est soit 1/11 soit 1/17. Par mesure de coefficients de diffusion, la stœchiométrie serait plutôt de l'ordre de 17. En revanche, les expériences de diffusion ont été effectuées à l'aide d'une sonde BBI statique. Or, à forte concentration en B3, la solution était sous forme colloïdale, ce qui suggère que l'on ne voit peut-être pas les signaux des agrégats. Il est probable que deux complexes de différentes stœchiométries existent. De plus, le modèle implique que les sites sont indépendants entre eux, or la fixation est certainement aussi à effet coopératif. Etant donné que la protéine est constituée de trois fois la séquence du peptide IB7₁₄, et que IB7₁₄ est apte à fixer trois B3, la protéine pourrait fixer au moins 3x3 tanins B3. Les résultats obtenus sont cohérents avec cette observation. L'adaptation des expériences de diffusion en rotation à l'angle magique tout comme l'accès à la structure tridimensionnelle de la protéine nous permettraient de déterminer avec précision la stœchiométrie et l'affinité du complexe.

#### **V. 4. 2. Sites de fixation des B3 sur le peptide**

D'après les résultats présentés précédemment trois tanins B3 sont capables de fixer le peptide IB7₁₄, trois sites devraient être mis en évidence. En effet, trois sites d'interaction ont clairement été identifiés par spectrométrie de RMN et par DM : Pro2-Pro6, Pro9-Pro11 et Gly13-Gly14. La variation de déplacement chimique de chaque proton de tous les acides aminés d'IB7₁₄ a été suivie par RMN. Les variations les plus importantes sont observées pour les acides aminés de ces trois sites. D'autre part, des expériences de type ROESY, réalisées pour un rapport molaire de 1/3, mettent en lumière des effets NOEs intermoléculaires entre le dimère B3 et les protons H $\alpha$  de la Gly4 et NH de la Lys5. Ces effets NOE sont en parfait accord avec un site Pro2-Pro6.

Lorsqu'on réalise des expériences de DM pour un rapport molaire 1/1, en faisant varier la position de départ de B3 de manière aléatoire, on retrouve le dimère B3 fixé au peptide à deux endroits différents : entre Pro2 et Pro6 et entre Pro9 et Pro11. Le premier site

est alors en accord avec les effets NOEs observés par RMN, et le second en accord avec les variations importantes de déplacements chimiques observées pour les prolines 9 et 10. Les dynamiques moléculaires partant d'un système IB7₁₄/B3 1/4 conduisent, quant à elles, toutes à la fixation de trois B3 sur le peptide alors que le quatrième est systématiquement éjecté du complexe. Les deux premiers sites étant ceux préalablement mis en évidence quant au troisième, son existence est confirmée par l'observation de fortes variations de déplacements chimiques pour les protons des glycines C-terminales, Gly13 et Gly14.

La présence d'effets NOE entre le dimère B3 et le site Pro2-Pro6 met en évidence l'existence d'un contact entre le B3 et ce site, dont la durée de vie est suffisamment longue pour être observée par RMN. En revanche, aucun effet NOE n'est observé entre le dimère et les deux autres sites alors qu'il y a toutefois une variation importante des déplacements chimiques. Cette observation pourrait s'expliquer par une différence d'orientation des tanins par rapport au peptide ou bien par une différence de dynamique de fixation des tanins entre les trois sites suggérant des différences d'affinité.

### **V. 4. 3. Structure de IB7₁₄ et IB7 en présence de tanins B3**

L'influence de la présence de tanins B3 sur la structure secondaire de la protéine IB7 et du peptide IB7₁₄ a été déterminée par dichroïsme circulaire. Il a été montré dans le chapitre IV que le peptide et la protéine ont exactement la même répartition de structure secondaire à 36% d'hélice de type II et 64% de pelote statistique. Les mêmes valeurs sont obtenues en présence de tanins B3. Les dimères B3 se fixent donc à la protéine et au peptide sans perturber leur conformation secondaire.

En revanche, concernant le peptide IB7₁₄, les contraintes issues de la RMN (valeurs de déplacements chimiques, constantes de couplages ³J et effets NOEs) à l'aide de la modélisation moléculaire nous ont permis d'accéder à la conformation spatiale du peptide. Seul en solution, le peptide IB7₁₄ adopte une conformation en deux hélices de type II de part et d'autre de la glycine 8, qui apparaît comme un axe de libre rotation. La flexibilité de ce résidu autorise un grand nombre de conformations possibles, maintenant toutefois les deux parties en hélice figées d'une conformation à l'autre. La présence de B3 semblerait bloquer la flexibilité en se fixant au peptide. Les deux parties en hélice seraient maintenues, mais le degré de liberté autour de la glycine 8 serait moindre. Au lieu d'obtenir un nuage de

conformations lorsque le peptide est seul, seulement une d'entre elles est favorisée en présence de trois tanins, en passant par un nuage de conformations restreint en présence d'une seule molécule de B3. L'ensemble de ces résultats est résumé dans la figure V-21. La conformation privilégiée en présence de 3 B3 est retrouvée, que l'on réalise des dynamiques avec ou sans les contraintes découlant des expériences de RMN. En effet, ces structures sont quasiment superposables. Ainsi, une conformation est privilégiée, pour laquelle tous les carbonyles semblent pointer sur la même face du peptide.

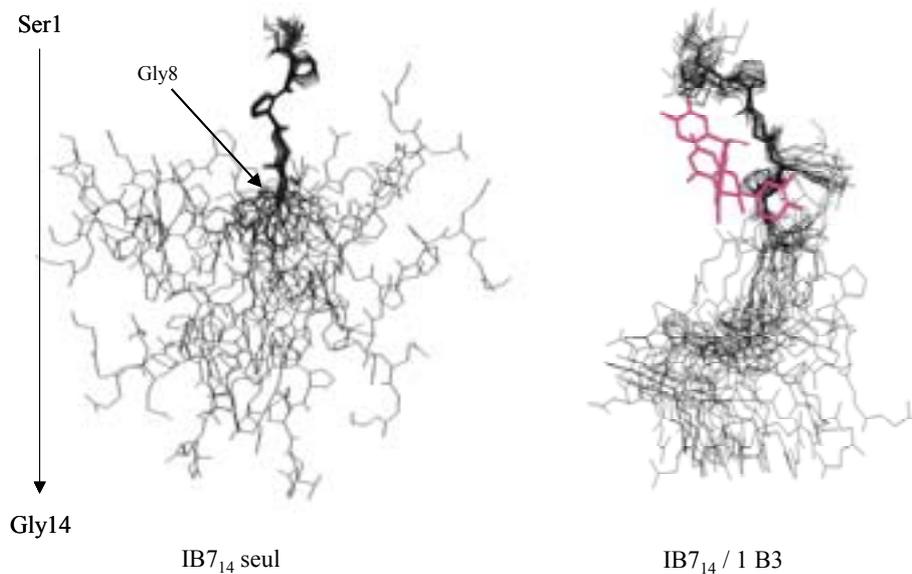


Figure V-21 : Restriction des conformations adoptées par le peptide en présence de B3.

Afin de valider la variation de la flexibilité autour de la glycine 8 en absence ou en présence de tanin il serait intéressant de mesurer les temps de relaxation longitudinaux des résidus du peptide. Ainsi, nous aurions accès à la dynamique du peptide atome par atome en fonction de l'ajout de B3.

#### V. 4. 4. Taille des complexes IB7₁₄/3B3 et IB7/nB3

La mesure du coefficient de diffusion apporte des informations sur la taille des molécules. Pour le peptide IB7₁₄, la différence de coefficient de diffusion est mesurée par RMN, entre le peptide seul et en présence de tanins B3 en concentration saturante, ce qui implique une augmentation de la taille du complexe. Le rayon de giration passe de 5,5 Å pour le peptide seul à 7,8 Å pour le complexe IB7₁₄-3B3. Ce résultat a été comparé à un calcul du

coefficient de diffusion par DM, où le rayon de giration du complexe obtenu était d'environ 8 à 9 Å. L'écart obtenu pour la mesure de la taille du complexe par RMN n'est que de 14 % par rapport à la DM. La condition de diffusion isotrope est respectée, car la courbe de variation du déplacement moyen de la molécule en fonction du temps est linéaire (annexe 6). En revanche, les calculs de la taille du peptide lorsqu'il est seul sont de 5,5 Å par RMN et 7,8 Å par DM. L'écart de 30 % entre ces valeurs peut être expliqué par la possibilité pour le peptide de faire des liaisons hydrogène avec l'eau lorsqu'il est seul en solution, alors qu'en présence de tanins, tous les groupements potentiels de liaison sont occupés par les B3. La relation d'Einstein n'est pas vérifiée (pente non linéaire, annexe 6), due au comportement non isotrope du peptide.

En ce qui concerne la variation de coefficient de diffusion pour le complexe IB7-B3, il diminue de 31% (de  $1,7$  à  $1,2 \times 10^{-6} \text{ cm}^2\text{s}^{-1}$ ) lorsque le rapport  $[B3]/[IB7]$  varie de 0 à 17. Il est enfin possible d'établir le rayon approximatif du complexe IB7-B3, d'après la relation [3], en considérant que celui-ci est sphérique. Le rayon serait de 18 Å pour le complexe dont le coefficient de diffusion est  $1,2 \times 10^{-6} \text{ cm}^2\text{s}^{-1}$ . En considérant l'ensemble du complexe tanin-protéine comme une entité sphérique, ceci correspond à une augmentation de masse du complexe de 11500 (D proportionnel à  $M^{1/3}$ ). Cette augmentation de masse peut correspondre à une protéine fixant 20 molécules de tanins (masse molaire de B3 :  $577 \text{ g}\cdot\text{mol}^{-1}$ ) ou bien à l'ensemble de deux protéines reliées entre elles par 10 tanins (masse molaire de IB7 :  $5766 \text{ g}\cdot\text{mol}^{-1}$ ). En considérant les variations de déplacement chimique des groupes de protons des acides aminés de la protéine, deux valeurs de nombre de sites de liaisons potentiels ont été trouvées :  $n = 10-11$  et  $n = 16-18$ . En revanche, le nombre de sites de liaisons obtenus à partir de la courbe de titration du coefficient de diffusion de IB7 ( $n = 17$ ) se rapproche de la seconde valeur. Afin d'accéder à l'intégralité de la courbe de titration (les derniers points n'ont pas été pris en compte en raison du manque de sensibilité), l'adaptation des expériences de diffusion sur une sonde HRMAS éluderait le problème relié à la formation de particules colloïdales. Pour confirmer l'hypothèse de la fixation de 20 tanins sur IB7 ou bien la formation d'un début de réseau entre 2 protéines et 10 tanins, une mesure par électrophorèse pourrait être envisagée. Les deux entités migreraient de manière différente selon le poids moléculaire. Il serait ainsi possible de choisir avec davantage de certitude l'une ou l'autre des hypothèses.

## V. 4. 5. Nature de l'interaction

La SM (ESI), la RMN ainsi que la DM nous ont permis de valider la nature de l'interaction *via* des liaisons hydrogène entre certains hydroxyles du tanin B3 et le peptide IB7₁₄. Tout d'abord, les calculs de modélisation moléculaire effectués sur le complexe IB7₁₄/B3, en accord avec les effets NOEs obtenus par RMN, nous permettent de visualiser des liaisons hydrogène entre les groupements carbonyles de certaines prolines et les groupements hydroxyles des cycles A, D ou B, E de la procyanidine B3. La conformation naturelle de B3 est une structure très compacte avec les groupements hydroxyles OH pointant vers l'extérieur de la structure ce qui facilite la formation de liaisons hydrogène. De plus, la nature amphiphile du peptide a été déterminée par des calculs d'hydrophobicité, tels que le peptide est constitué d'une face hydrophile et d'une face hydrophobe. Or il est intéressant de constater que les B3 ne se fixent que sur une face du peptide IB7₁₄, identifiée comme la face hydrophile. Enfin, l'observation de complexes en SM est en accord avec l'existence de liaisons hydrogène^{197,198,201}. En effet, au niveau de la chambre d'ionisation, le complexe IB7₁₄/B3 est sous forme désolvatée ce qui exclut une interaction de nature hydrophobe entre ces composés. Tous ces résultats sont donc en accord avec une interaction majeure pilotée par liaisons hydrogène ou encore par interaction de type électrostatique peut être appuyées par une interaction de type hydrophobe. Hagerman et Butler décrivent également par des expériences de compétition en milieu aqueux, l'interaction procyanidine-peptide gouvernée par des liaisons hydrogène¹ en accord avec Artz *et al.*¹⁷⁹. La nature de l'interaction est pourtant largement discutée. Haslam²⁰² décrit la complexation des polyphénols sur les protéines comme un exemple de reconnaissance moléculaire où les polyphénols agissent comme des ligands multidentates dont la principale force d'interaction provient d'effets hydrophobes. L'interaction est alors traduite comme une association entre les cycles phénols hydrophobes et les cycles pyrrolidines des prolines, sans déterminer la structure du complexe^{129,176}. Il est toutefois important de noter que les auteurs travaillent dans des solutions à 10-20% de DMSO en vue d'améliorer la solubilité de leurs échantillons. Ces conditions peuvent modifier les propriétés hydrophobes/philes des tanins comme des peptides. De plus, ils considèrent les résidus proline et glycine comme des résidus hydrophobes, ce qui va dans le sens de leur conclusion. En revanche, Rose *et al.*²⁰³ précisent que les prolines sont des résidus relativement polaires, ce qui conforte nos résultats où la fixation a lieu principalement sur les prolines.



**Chapitre VI CONCLUSIONS ET  
PERSPECTIVES**



La caractérisation de l'astringence, perçue comme une sensation tactile lors de la dégustation des vins rouges, a été entrevue d'un point de vue physico-chimique par le biais d'une analyse complète de l'interaction Tanins-Protéines. Les Protéines Riches en Proline de la salive humaine précipitent les tanins alimentaires ce qui a pour conséquence la diminution du potentiel lubrifiant de la salive.

## **VI. 1. CONCLUSION**

La synthèse de protéines riches en proline a été mise au point [Simon, 2002 #107]. L'obtention de B3 a fait l'objet du sujet de thèse de Karine Barathieu⁶³.

Pour la première fois une analyse structurale de PRP a été effectuée. L'influence de la présence de tanins B3 sur l'adaptation tridimensionnelle du peptide a pu être déterminée. La compréhension du mode d'interaction entre IB7₁₄ et B3 a été menée par une approche « multi-techniques »²⁰⁴. En effet, la spectrométrie de masse, la RMN et la dynamique moléculaire nous ont permis d'accéder à la stœchiométrie du complexe ainsi qu'à la constante de dissociation : 3 tanins B3 fixent un peptide IB7₁₄ avec un  $K_d$  de l'ordre du millimolaire. Cependant, le modèle utilisé prenant en compte une fixation indépendante de chaque tanin sur le peptide paraît ne pas être bien adapté. La mise au point d'un modèle mathématique prenant en compte une fixation coopérative des tanins est envisagée. Tous les B3 se fixent sur les zones hydrophiles du peptide au moyen de liaisons hydrogène et dans une moindre mesure d'interactions de type électrostatique. L'analyse structurale complète de la protéine IB7 n'a pu être effectuée. Cependant, une première approche thermodynamique de la formation du complexe IB7-B3 a pu être entreprise par RMN. La stœchiométrie du complexe est alors ~11 ou ~17, avec la même constante d'affinité (~mM) que le peptide IB7₁₄. Ces résultats montrent que la capacité de fixation des tanins sur ce type de protéine est en rapport avec leur taille et donc du nombre de résidus proline qu'elle contient : la seule fonction connue pour les PRPs basiques étant de précipiter les tanins, elles agissent comme des « éponges » à tanins.

Un modèle d'étude de l'interaction Tanin-Protéine a donc été mis au point. La détermination de la stœchiométrie et de la constante d'affinité du peptide pour le B3 pourrait être réalisée pour un grand nombre de tanins. En effet, comparer ces différents paramètres avec le même peptide permettrait de doser la capacité d'un tanin à fixer les protéines. La

typicité d'un vin pourrait être étudiée par le biais de l'affinité des tanins sur ce modèle de peptide. Une autre approche serait de considérer la fixation de tanins sur différentes protéines afin de sélectionner celles paraissant les plus aptes à exercer des fonctions de protéines de collage lors de la clarification des vins.

## **VI. 2. PERSPECTIVES**

En ce qui concerne l'analyse de l'interaction IB7₁₄-B3, il serait intéressant d'avoir des données dynamiques du peptide en absence et en présence de B3 : en mesurant les temps de relaxation de chaque atome du peptide nous devrions comprendre la variation de la flexibilité autour de la glycine 8 en présence de tanin, par une diminution de la dynamique. Dans un autre registre, la mesure de l'échange isotopique par spectrométrie de masse des protons du peptide pourrait apporter de précieuses informations : non seulement les protons accessibles au solvant en présence de tanins mettraient en évidence les sites de fixation, mais également un suivi de la cinétique de l'échange pourrait rendre compte de la dynamique de ces différents sites.

Afin d'obtenir des données physico-chimiques plus précises sur le complexe protéine entière/tanin, il serait nécessaire, à l'avenir, d'adapter les expériences de mesures de coefficients de diffusion en utilisant la technique HRMAS, d'une part, et d'attribuer l'ensemble des résonances de la protéine, d'autre part. Ceci est envisagé en utilisant d'autres méthodes de RMN permettant d'identifier tous les protons malgré un chevauchement important des résonances en travaillant sur une protéine doublement enrichie en ¹³C et en ¹⁵N (RMN 3D). Cette approche devrait nous permettre d'obtenir un fichier de contraintes de la protéine en absence ou présence de tanins, afin d'accéder à la structure tridimensionnelle de la protéine IB7.

L'analyse de l'interaction Tanin-Protéine pourrait également être envisagée d'un autre point de vue : l'aspect cinétique de l'agrégation. En effet, un échantillon contenant peptide et tanin peut attendre plusieurs jours avant de précipiter.

Afin de comprendre l'importance de la stéréochimie et du degré de polymérisation des tanins dans le cadre de l'astringence, il serait intéressant d'étudier cette interaction sur

d'autres modèles de tanins. En effet, la conformation « fermée » du B3 pourrait être comparée à celle d'un tanin ayant une liaison C₄-C₆ « ouverte », ou avec des tanins ayant des poids moléculaires plus importants comme le trimère C1, dont la mise au point de la synthèse fait l'objet de la thèse d'I. Tarascon au LCOO... Ainsi, nous pourrions établir une gamme d'affinité en fonction de la régiosélectivité (liaison C4-C8 ou C4-C6), de la stéréosélectivité (unité catéchine ou épicatechine), du degré de polymérisation ou encore du nombre de groupements galloyles sur le tanin.



## **Chapitre VII ANNEXES**



## ANNEXE 1 : ABBAQUES DES DEPLACEMENTS CHIMIQUES EN FONCTION DES ACIDES AMINES EN SOLUTION

Acides Aminés	NH	H $\alpha$	H $\beta$	Autres
Arg	8,27	4,38	1,89; 1,79	H $\gamma$ : 1,72; H $\delta$ : 3,31; HN:7,17; 6,62
Asn	8,75	4,75	2,83; 2,75	HN:7,59
Asp	8,41	4,77	2,84; 2,75	-
Gly	8,39	3,97	-	-
Gln	8,41	4,37	2,13;2,01	H $\gamma$ : 2,38; $\delta$ HN: 6,87
Lys	8,41	4,36	1,87; 1,75	H $\gamma$ : 1,45; H $\delta$ : 1,70; H $\epsilon$ : 3,02; HN:7,52
Pro	-	4,47	2,30; 1,98	H $\gamma$ : 2,03; H $\delta$ : 3,65
Ser	8,38	4,50	3,89; 3,89	-

Tableau 1 : Valeurs standards des déplacements chimiques  $^1\text{H}$  des acides aminés constitutants IB7, en solution dans l'eau¹⁰².

## ANNEXES 2 : EFFETS NOES ET CONSTANTES DE COUPLAGE $^3\text{J}$ RAPPORTEES AUX ANGLES DIEDRES

Résidus 1	Résidus 2	Distance
$^2\text{Pro-H}\alpha$	$^3\text{Pro-H}\delta_2$	$2,2 \pm 0,4$
$^2\text{Pro-H}\alpha$	$^3\text{Pro-H}\delta_3$	$2,2 \pm 0,4$
$^3\text{Pro-H}\alpha$	$^4\text{Gly-HN}$	$2,2 \pm 0,4$
$^4\text{Gly-H}\alpha_2$	$^5\text{Lys+HN}$	$2,2 \pm 0,4$
$^4\text{Gly-H}\alpha_3$	$^5\text{Lys+HN}$	$2,2 \pm 0,4$
$^5\text{Lys+ H}\alpha$	$^6\text{Pro-Q}\delta_2$	$2,2 \pm 0,4$
$^5\text{Lys+ H}\alpha$	$^6\text{Pro-Q}\delta_3$	$2,2 \pm 0,4$
$^5\text{Lys+ H}\alpha$	$^6\text{Pro-Q}\gamma_2$	$3,5 \pm 0,9$
$^5\text{Lys+ H}\alpha$	$^6\text{Pro-Q}\gamma_3$	$5,0 \pm 0,5$

⁵ Lys+ H $\alpha$	⁶ Pro-Q $\beta_2$	3,5 $\pm$ 0,9
⁵ Lys+ H $\alpha$	⁶ Pro-Q $\beta_3$	3,5 $\pm$ 0,9
⁵ Lys+-H $\beta_2$	⁶ Pro-Q $\delta_2$	5,0 $\pm$ 0,5
⁵ Lys+-H $\beta_2$	⁶ Pro-Q $\delta_3$	5,0 $\pm$ 0,5
⁵ Lys+-H $\beta_3$	⁶ Pro-Q $\delta_2$	5,0 $\pm$ 0,5
⁵ Lys+-H $\beta_3$	⁶ Pro-Q $\delta_3$	5,0 $\pm$ 0,5
⁵ Lys+-H $\gamma_2$	⁶ Pro-Q $\delta_2$	3,5 $\pm$ 0,9
⁵ Lys+-H $\gamma_2$	⁶ Pro-Q $\delta_3$	3,5 $\pm$ 0,9
⁵ Lys+-H $\gamma_3$	⁶ Pro-Q $\delta_2$	3,5 $\pm$ 0,9
⁵ Lys+-H $\gamma_3$	⁶ Pro-Q $\delta_3$	3,5 $\pm$ 0,9
⁶ Pro-H $\alpha$	⁷ Gln-HN	2,2 $\pm$ 0,4
⁶ Pro-H $\beta_2$	⁷ Gln-HN	3,5 $\pm$ 0,9
⁶ Pro-H $\beta_3$	⁷ Gln-HN	3,5 $\pm$ 0,9
⁷ Gln-H $\alpha$	⁸ Gly-HN	2,2 $\pm$ 0,4
⁷ Gln-H $\beta_2$	⁸ Gly-HN	3,5 $\pm$ 0,9
⁷ Gln-H $\beta_2$	⁸ Gly-HN	3,5 $\pm$ 0,9
⁷ Gln-H $\beta_3$	⁸ Gly-HN	3,5 $\pm$ 0,9
⁷ Gln-H $\beta_3$	⁸ Gly-HN	3,5 $\pm$ 0,9
¹⁰ Pro-H $\alpha$	¹¹ Pro-Q $\delta_2$	2,2 $\pm$ 0,4
¹⁰ Pro-H $\alpha$	¹¹ Pro-Q $\delta_3$	2,2 $\pm$ 0,4
¹⁰ Pro-H $\beta_2$	¹³ Gly-H $\alpha_1$	3,5 $\pm$ 0,9
¹⁰ Pro-H $\beta_2$	¹³ Gly-H $\alpha_2$	3,5 $\pm$ 0,9
¹⁰ Pro-H $\beta_3$	¹³ Gly-H $\alpha_1$	3,5 $\pm$ 0,9
¹⁰ Pro-H $\beta_3$	¹³ Gly-H $\alpha_2$	3,5 $\pm$ 0,9
¹¹ Pro-H $\alpha$	¹² Gln-HN	2,2 $\pm$ 0,4
¹¹ Pro-H $\beta_2$	¹² Gln-HN	3,5 $\pm$ 0,9
¹¹ Pro-H $\beta_3$	¹² Gln-HN	3,5 $\pm$ 0,9
¹¹ Pro-H $\gamma_2$	¹² Gln-HN	3,5 $\pm$ 0,9
¹¹ Pro-H $\gamma_3$	¹² Gln-HN	3,5 $\pm$ 0,9

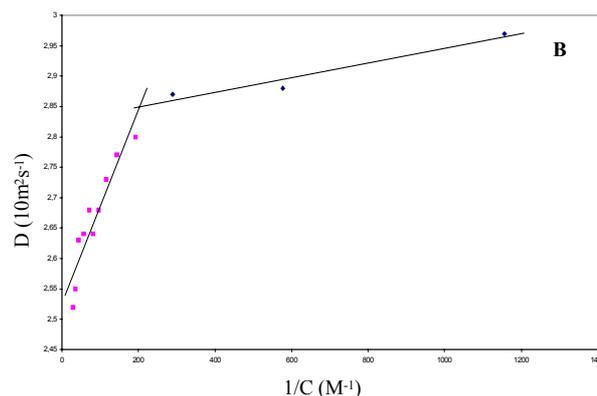
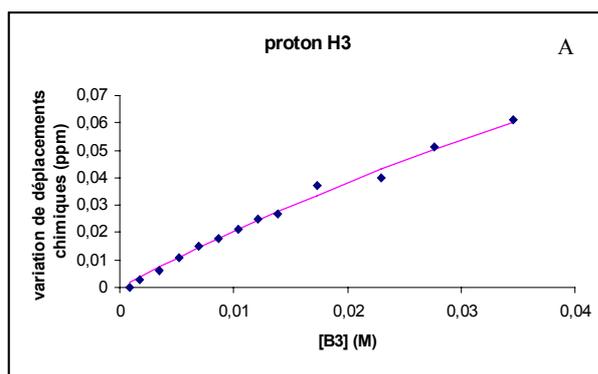
$^{12}\text{Gln-H}\alpha$	$^{13}\text{Gly-HN}$	$2,2 \pm 0,4$
$^{12}\text{Gln-H}\alpha$	$^{14}\text{Gly-HN}$	$3,5 \pm 0,9$
$^{13}\text{Gly-H}\alpha_2$	$^{14}\text{Gly-HN}$	$3,5 \pm 0,9$

Tableau 2 : Distances entre les protons du peptide IB7₁₄ en fonction des effets NOEs observés. Les effets NOEs ont été déterminés par une expérience ROESY pour un rapport molaire 1/3 IB7₁₄/B3 en solution dans l'eau (avec 10 % de D₂O) à pH 3,5 et 298 K.

	$^3J_{\text{NH-C}\alpha}$ (Hz)	Angle dièdre ( $\theta$ ) en °
S ₁	8,3	$\pm 153$
G ₄	5,9	$\pm 133/\pm 24,5$
K ₅	7,0	$\pm 142$
Q ₇	6,6	$\pm 139/\pm 13,4$
G ₈	5,7	$\pm 131/\pm 27,4$
Q ₁₂	7,0	$\pm 142$
G ₁₃	5,7	$\pm 131/\pm 27,4$
G ₁₄	5,7	$\pm 131/\pm 27,4$

Tableau 3 : Constantes de couplages  3J  (en Hertz) entre les protons amines et les protons  $\alpha$  des acides aminés voisins. Ces mesures ont été lues sur le spectre 1D pour le peptide IB7₁₄ seul en solution dans l'eau (avec 10 % de D₂O) à pH 3,5 et 298 K. Aucune variation significative n'apparaît en présence de tanins B3.

### ANNEXES 3 : AUTO-ASSOCIATION DU TANIN B3



Variation du déplacement chimique d'un proton du B3 (A) et variation du coefficient de diffusion transversal du B3 (B) en fonction de la concentration en B3.

## ANNEXES 4 : EFFETS NOES INTERMOLECULAIRES IB7₁₄/B3

Résidus Peptidiques	Résidus B3	Distance
⁴ Gly-H $\alpha$ ₁	2F	2
⁴ Gly-H $\alpha$ ₂	2F	5
⁴ Gly-H $\alpha$ ₁	4F	6
⁴ Gly-H $\alpha$ ₁	4'F	6
⁵ Lys+-HN	2F	5

Tableau 4 : Distances entre protons en accord avec les effets NOEs intermoléculaires, obtenues pour un rapport molaire IB7₁₄/B3 de 1/3 dans H₂O/D₂O (90/10) à 27°C.

## ANNEXES 5 : VARIATION DU DEPLACEMENT CHIMIQUE DES ¹H PEPTIDIQUES LORS DE L'ADDITION DE PROCYANIDINE B3

ratio IB7 ₁₄ /B3	résidu	NH	H $\alpha$	H $\beta$	Autres
s					

0.00	<b>¹Ser</b>	4.390	3.991,3.842	
0.50		4.400	4.007,3.860	
0.82		4.394	3.999,3.851	
1.05		4.390	3.998,3.849	
1.94		4.400	4.403,3.865	
2.60		ND	ND	
2.92		4.402	4.005,3.849	
3.33		4.403	4.004,3.856	
3.60		4.400	4.001,3.851	
5.00		ND	ND	
5.56		ND	ND	
5.88		4.398	4.005,3.854	
6.70		4.398	4.003,3.851	
0.00	<b>²Pro</b>	4.691	1.874,2.290	$H\gamma = 1.994, H\delta = 3.574, ND$
0.50		4.662	1.872,2.287	$H\gamma = 1.993, H\delta = 3.56, 3.743$
0.82		4.643	1.864,2.271	$H\gamma = 1.984, H\delta = 3.554, 3.731$
1.05		4.642	1.862,2.267	$H\gamma = 1.983, H\delta = 3.552, 3.73$
1.94		4.624	1.868,2.254	$H\gamma = 1.977, H\delta = 3.555, 3.733$
2.60		4.631	1.869,2.271	$H\gamma = 1.985, H\delta = 3.557, 3.747$
2.92		4.618	1.866,2.247	$H\gamma = 1.973, H\delta = 3.555, 3.702$
3.33		4.614	1.858,2.248	$H\gamma = 1.978, H\delta = 3.552, 3.733$
3.60		4.595	1.852,2.237	$H\gamma = 1.970, H\delta = 3.551, 3.744$
5.00		4.607	1.851,2.230	$H\gamma = 1.961, H\delta = 3.54, 3.71$
5.56		4.578	1.842,2.237	$H\gamma = 1.96, H\delta = 3.533, ND$
5.88		4.551	1.847,2.244	$H\gamma = 1.957, H\delta = 3.538, ND$
6.70		(4.56)	1.843,2.224	$H\gamma = 1.959, H\delta = 3.554, 3.744$
0.00	<b>³Pro</b>	4.390	1.921,2.251	$H\gamma = 2.021, H\delta = 3.805, 3.629$
0.50		4.397	1.917,2.281	$H\gamma = 2.02, H\delta = 3.796, 3.629$
0.82		4.394	1.909,2.29	$H\gamma = 2.02, H\delta = 3.795, 3.619$
1.05		4.395	1.907,2.287	$H\gamma = 2.019, H\delta = 3.784, 3.628$

1.94		4.406	1.916,2.288	H $\gamma$ = 2.02, H $\delta$ =3.781,3.613
2.60		4.390	1.918,2.282	H $\gamma$ = 2.019, H $\delta$ =3.782,3.617
2.92		4.404	1.928,2.283	H $\gamma$ = 2.023, H $\delta$ =3.785,3.614
3.33		4.404	1.92,2.298	H $\gamma$ = 2.023, H $\delta$ =3.782,3.613
3.60		4.401	1.916,2.291	H $\gamma$ = 2.014, H $\delta$ =3.782,3.614
5.00		4.410	1.921,2.294	H $\gamma$ = 2.018, H $\delta$ =3.789,3.618
5.56		4.405	1.899,2.285	H $\gamma$ = 2.017, H $\delta$ = ND ,3.62
5.88		4.406	1.918,2.292	H $\gamma$ = 2.015, H $\delta$ =3.779, ND
6.70		4.405	1.819,2.286	H $\gamma$ = 2.016, H $\delta$ =3.78, 3.617
0.00	⁴ Gly	8.355	3.916,3.860	
0.50		8.362	3.919,3.860	
0.82		8.362	3.915,3.868	
1.05		8.261	3.926,3.865	
1.94		8.361	3.934,3.869	
2.60		8.360	(3.889)	
2.92		8.360	3.945,3.867	
3.33		8.366	3.942,3.864	
3.60		8.366	3.946,3.870	
5.00		8.361	(3.905)	
5.56		8.360	(3.883)	
5.88		8.358	(3.895)	
6.70		8.364	3.948,3.871	
0.00	⁵ Lys	8.067	4.598	1.759 H $\gamma$ = 1.410, H $\delta$ =1.65, H $\epsilon$ =2.963
0.50		8.062	4.607	1.775 H $\gamma$ = 1.413, H $\delta$ =1.659, H $\epsilon$ =2.963
0.82		8.057	4.601	1.761 H $\gamma$ = 1.407, H $\delta$ =1.661, H $\epsilon$ =2.965
1.05		8.056	4.621	1.764 H $\gamma$ = 1.409, H $\delta$ =1.658, H $\epsilon$ =2.957
1.94		8.055	4.600	1.765 H $\gamma$ = 1.408, H $\delta$ =1.651, H $\epsilon$ =2.966
2.60		8.054	4.616	1.778 H $\gamma$ = 1.402, H $\delta$ =1.649, H $\epsilon$ =2.954
2.92		8.052	4.603	1.768 H $\gamma$ = 1.41, H $\delta$ =1.659, H $\epsilon$ =2.965
3.33		8.053	4.606	1.769 H $\gamma$ = 1.408, H $\delta$ =1.652, H $\epsilon$ =2.966

3.60	8.050	4.609	1.767	H $\gamma$ = 1.407, H $\delta$ = 1.647, H $\epsilon$ = 2.963
5.00	8.045	4.611	1.767	H $\gamma$ = 1.404, H $\delta$ = 1.65, H $\epsilon$ = 2.956
5.56	8.047	4.600	1.769	H $\gamma$ = 1.408, H $\delta$ = 1.646, H $\epsilon$ = 2.95
5.88	8.043	4.621	1.764	H $\gamma$ = 1.403, H $\delta$ = 1.644, H $\epsilon$ = 2.959
6.70	8.039	4.614	1.766	H $\gamma$ = 1.41, H $\delta$ = 1.659, H $\epsilon$ = 2.962
0.00	<b>⁶Pro</b>	4.385	1.902,2.262	H $\gamma$ = 2.003, H $\delta$ = 3.787,3.61
0.50		4.384	1.889,2.271	H $\gamma$ = 2.006, H $\delta$ = 3.775,3.609
0.82		4.384	1.899,2.278	H $\gamma$ = 2.009, H $\delta$ = 3.767,3.574
1.05		4.382	1.88,2.267	H $\gamma$ = 1.998, H $\delta$ = 3.774,3.569
1.94		4.380	1.87,2.25	H $\gamma$ = 1.971, H $\delta$ = 3.752,3.541
2.60		4.379	1.82,2.261	H $\gamma$ = 1.999, H $\delta$ = 3.746,3.564
2.92		4.386	1.877,2.255	H $\gamma$ = 1.996, H $\delta$ = 3.716,3.552
3.33		4.380	1.874,2.263	H $\gamma$ = 1.969, H $\delta$ = 3.741,3.549
3.60		4.378	1.871,2.25	H $\gamma$ = 1.997, H $\delta$ = 3.726,3.548
5.00		4.386	1.889,2.256	H $\gamma$ = 1.976, H $\delta$ = 3.691,3.551
5.56		4.382	1.897,2.248	H $\gamma$ = 1.981, H $\delta$ = ND, 3.542
5.88		4.373	1.886,2.246	H $\gamma$ = ND, H $\delta$ = 3.73.688,3.53
6.70		4.374	1.87,2.248	H $\gamma$ = 1.993, H $\delta$ = 3.67, 3.549
0.00	<b>⁷Gln</b>	8.456	4.275	H $\gamma$ = 2.361
0.50		8.457	4.277	H $\gamma$ = 2.364
0.82		8.455	4.274	H $\gamma$ = 2.364
1.05		8.444	4.282	H $\gamma$ = 2.363
1.94		8.447	4.282	H $\gamma$ = 2.37
2.60		8.445	4.277	H $\gamma$ = 2.367
2.92		8.445	4.281	H $\gamma$ = 2.367
3.33		8.451	4.285	H $\gamma$ = 2.369
3.60		8.444	4.279	H $\gamma$ = 2.359
5.00		8.443	4.285	H $\gamma$ = 2.361
5.56		8.441	4.278	H $\gamma$ = 2.358
5.88		8.441	4.275	H $\gamma$ = 2.364

6.70		8.434	4.280	1.976,2.093	H $\gamma$ =2.363
0.00	<b>⁸Gly</b>	8.400	3.942,3.895		
0.50		8.401	3.945		
0.82		8.400	3.951		
1.05		8.398	3.948		
1.94		8.400	3.955		
2.60		8.401	3.939		
2.92		8.400	3.954		
3.33		8.399	3.950		
3.60		8.398	3.946		
5.00		8.396	3.951		
5.56		8.397	3.941		
5.88		8.395	3.948		
6.70		8.393	3.944		
0.00	<b>⁹Pro</b>		4.839	1.812,2.421	H $\gamma$ = ND, H $\delta$ =3.5, ND
0.50			4.827	1.855,2.418	H $\gamma$ = 2.043, H $\delta$ =3.518, ND
0.82			4.820	1.908,2.373	H $\gamma$ = 2.093, H $\delta$ =3.517, ND
1.05			4.791	1.841,2.388	H $\gamma$ = 2.076, H $\delta$ =3.516,3.738
1.94			4.787	1.803,2.387	H $\gamma$ = 2.088, H $\delta$ =3.525, ND
2.60			4.751	1.835,2.378	H $\gamma$ = 2.015, H $\delta$ =3.557, ND
2.92			4.743	1.830,2.361	H $\gamma$ = 2.012, H $\delta$ =3.554, ND
3.33			4.743	1.843,2.367	H $\gamma$ = 2.006, H $\delta$ =3.546,3.718
3.60			4.747	1.833,2.369	H $\gamma$ = 2.003, H $\delta$ = ND, 3.744
5.00			4.746	1.832,2.329	H $\gamma$ = 2.009, H $\delta$ =3.558,3.734
5.56			4.727	1.815,2.340	H $\gamma$ = 1.998, H $\delta$ =3.550,3.729
5.88			4.737	1.829,2.387	H $\gamma$ = 2.007, H $\delta$ = ND, ND
6.70			4.723	1.820,2.314	H $\gamma$ = 2.008, H $\delta$ =3.558,3.741
0.00	<b>¹⁰Pro</b>		4.746	1.845,2.383	H $\gamma$ = 2.014, H $\delta$ =3.565,3.741
0.50			4.742	1.846,2.3	H $\gamma$ = 2.013, H $\delta$ =3.564,3.741
0.82			4.736	1.839,2.364	H $\gamma$ = 2.02, H $\delta$ =3.563,3.74

1.05		4.733	1.833,2.353	$H\gamma = 2.018, H\delta = 3.747, 3.561$
1.94		4.724	1.825,2.339	$H\gamma = 1.996, H\delta = 3.552, 3.736$
2.60		4.714	1.841, ND	$H\gamma = 2.02, H\delta = ND, ND$
2.92		4.676	1.869,2.291	$H\gamma = 1.991, H\delta = 3.551, 3.748$
3.33		4.652	1.868,2.367	$H\gamma = 1.989, H\delta = 3.564, 3.744$
3.60		4.638	1.859,2.367	$H\gamma = 2.051, H\delta = 3.548, 3.743$
5.00		4.639	1.850, ND	$H\gamma = 2.013, H\delta = 3.547, 3.739$
5.56		4.633	1.869,2.266	$H\gamma = 1.979, H\delta = 3.542, ND$
5.88		4.641	1.857,2.33	$H\gamma = 1.978, H\delta = 3.548, 3.743$
6.70		4.636	1.858,2.364	$H\gamma = 1.973, H\delta = 3.55, ND$
<hr/>				
0.00	<b>¹¹Pro</b>	4.370	1.864,2.271	$H\gamma = 1.984, H\delta = 3.603, 3.76$
0.50		4.365	1.862,2.25	$H\gamma = 1.972, H\delta = 3.592, 3.731$
0.82		4.375	1.861,2.25	$H\gamma = 1.975, H\delta = 3.591, 3.73$
1.05		4.363	1.87,2.251	$H\gamma = 1.972, H\delta = 3.591, 3.73$
1.94		4.368	1.867,2.251	$H\gamma = 1.962, H\delta = 3.579, 3.752$
2.60		4.370	1.87,2.24	$H\gamma = 1.972, H\delta = 3.591, 3.772$
2.92		4.368	1.865,2.219	$H\gamma = 1.962, H\delta = 3.569, 3.762$
3.33		4.365	1.874,2.238	$H\gamma = 1.951, H\delta = 3.576, 3.753$
3.60		4.351	1.873,2.227	$H\gamma = 1.958, H\delta = 3.578, 3.76$
5.00		4.357	1.869,2.213	$H\gamma = 1.963, H\delta = 3.573, 3.756$
5.56		4.357	1.86,2.219	$H\gamma = 1.97, H\delta = 3.569, 3.765$
5.88		4.358	1.864, ND	$H\gamma = 1.969, H\delta = 3.573, 3.775$
6.70		4.350	ND, 2.219	$H\gamma = 1.971, H\delta = 3.579, 3.754$
<hr/>				
0.00	<b>¹²Gln</b>	8.493	4.302	$H\gamma = 2.38$
0.50		8.494	4.300	$H\gamma = 2.37$
0.82		8.483	4.301	$H\gamma = 2.37$
1.05		8.481	4.301	$H\gamma = 2.361$
1.94		8.471	4.302	$H\gamma = 2.361$
2.60		8.481	4.296	$H\gamma = 2.366$
2.92		8.474	4.303	$H\gamma = 2.362$

3.33	8.476	4.303	1.915,2.092	H $\gamma$ = 2.364
3.60	8.476	4.296	1.914,2.066	H $\gamma$ = 2.354
5.00	8.470	4.297	1.94,2.097	H $\gamma$ = 2.349
5.56	8.470	4.297	1.896,2.062	H $\gamma$ = 2.348
5.88	8.466	4.268	1.893,2.077	H $\gamma$ = 2.363
6.70	8.470	4.299	1.903,2.063	H $\gamma$ = 2.354

---

0.00	¹³ Gly	8.189	4.120,3.988
0.50		8.140	4.122,3.959
0.82		8.132	4.109,3.996
1.05		8.130	4.133,3.950
1.94		8.117	4.107,3.987
2.60		8.129	4.135,3.962
2.92		8.074	4.106,3.987
3.33		8.099	4.126,3.988
3.60		8.098	4.117,3.958
5.00		8.064	4.126,3.957
5.56		8.045	(4.071)
5.88		8.043	(4.064)
6.70		8.056	4.129,3.989

---

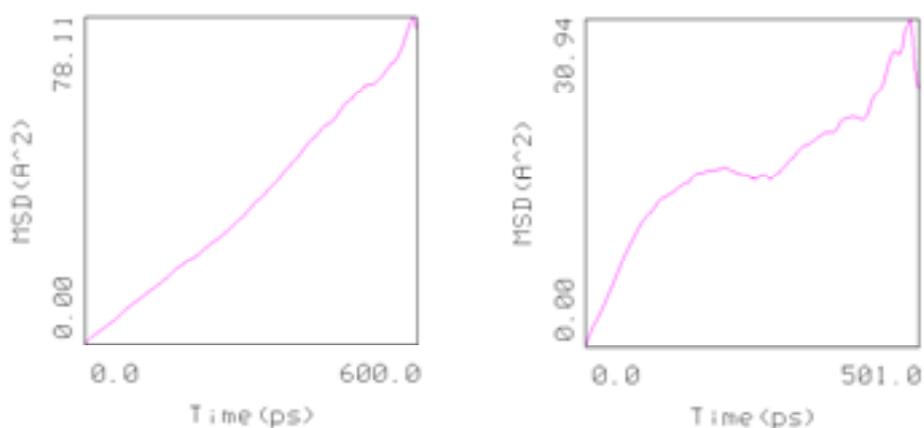
0.00	¹⁴ Gly	8.123	3.932
0.50		8.124	3.939
0.82		8.121	3.933
1.05		8.101	3.91
1.94		8.079	3.938
2.60		8.066	3.881
2.92		8.073	3.897
3.33		8.069	3.855
3.60		8.069	3.888
5.00		8.047	3.888
5.56		8.026	3.849

---

5.88	8.025	3.847
6.70	7.905	3.754

Tableau 5 : Valeurs des déplacements chimiques des protons du peptide IB7₁₄ en fonction de l'ajout de concentrations croissantes en B3.

**ANNEXES 6 : COURBE DE DIFFUSION DU PEPTIDE IB7₁₄ PAR DM, REPRESENTANT LE CARRE DU DEPLACEMENT MOYEN DE LA MOLECULE EN FONCTION DU TEMPS. LA FIGURE DE GAUCHE REPRESENTE LE PEPTIDE COMPLEXE AUX TANINS, CELLE DE DROITE, LE PEPTIDE SEUL**



Le déplacement quadratique moyen (MSD) est représenté en fonction du temps. Les mesures ont été prises en fin de la dynamique.



## BIBLIOGRAPHIE

1. Hagerman, A. E. & Butler, L. G. *Journal of Biological Chemistry* 256, 4494-4497 (1981).
2. McManus, J. P. et al. *Journal of Chemical Society Perkins Trans. II*, 1429-1443 (1985).
3. Spencer, C. M. et al. *Phytochemistry* 27, 2397-3409 (1988).
4. Bate-Smith, E. C. *Food* 23, 124 (1954).
5. Joslyn, M. A. & Goldstein, J. L. *Astringency of fruit and fruit products in relation to phenolic content*. (eds. Mrack, M. E. & Stewart, G. F.) (Advances in food research, Chichester, 1964).
6. Bartoshuk, L. M. *The American Journal of Clinical Nutrition* 31, 1068-1077 (1978).
7. Green, B. G. *Acta Psychology Amsterdam*, 84-119 (1993).
8. Breslin, P. A. S., Glimore, M. M., Beauchamp, G. K. & Green, B. G. *Chemical senses* 18, 405-417 (1993).
9. Critchley, H. D. & Rolls, E. T. *Chemical senses* 21, 135-145 (1996).
10. Bartoshuk, L. M. *Food quality and preference* 4, 21-32 (1993).
11. Haslam, E. & Lilley, T. H. *Critical Reviews in Food Science and Nutrition* 27, 1-40 (1988).
12. Lee, C. B. & Lawless, H. T. *Chemical senses* 16, 225-238 (1991).
13. Lawless, H. T. & Corrigan, C. J. *Chemical senses* 19, 141-154 (1994).
14. Fischer, U., Boulton, R. B. & Noble, A. C. *Food quality and preference* 5, 55-64 (1994).
15. Singleton, V. I. & Esau, P. *Phenolic substances in grapes and wines, and their significance* (Academic Press, New York, 1969).
16. Bate-Smith, E. C. *Phytochemistry* 12, 907-912 (1973).
17. Wu, A. M., Csako, G. & Herp, A. *Molecular and Cellular Biochemistry* 137, 39-55 (1994).
18. Mandel, I. D. *Journal of Dental Research* 66, 623-627 (1987).
19. Kauffman, D. L. & Keller, P. J. *Archives of Oral Biology* 24, 249-256 (1979).
20. Bennick, A. *Journal of Dental Research* 66, 457-461 (1987).
21. Prinz, J. F. & Lucas, P. W. *Journal of Oral Rehabilitation* 27, 991-994 (2000).
22. Kallithraka, S., Bakker, J. & Clifford, M. N. *Journal of Sensory Studies* 13, 29-43 (1998).

23. Singleton, V. L. & Noble, A. C. *Wine flavor and phenolic substances* (eds. Charalambous, G. & Katz, I.) (American Chemical Society, Washington, 1976).
24. Thorngate, J. H. & Singleton, V. L. *American Journal of Enology and Viticulture* 45, 259-262 (1994).
25. Ricardo da Silva, J. M., Rosec, J. P., Bourzeix, M., Mourgues, J. & Moutounet, M. *Vitis* 31, 55-63 (1992).
26. Fox, P. C., Bodner, L., Tabak, L. A. & Levine, M. J. *Journal of Dental Research* 64, 327 (1985).
27. Tsai, H. & Bobek, L. A. *Critical Review in Oral Biology and Medicine* 9, 480 (1998).
28. Iontcheva, I., Oppenheim, F. G. & Troxler, R. F. *Journal of Dental Research* 76, 734 (1997).
29. Bennick, A. *Molecular and Cellular Biochemistry* 45, 83-99 (1982).
30. Kauffman, D. L., Wong, T. Y., Bennick, A. & Keller, P. J. *Biochemistry* 21, 6558-6562 (1982).
31. Maeda, N. *Biochemistry Genetic* 23, 455-464 (1985).
32. Stubbs, M. J. et al. *Archive of Oral Biology* 43, 753-770 (1998).
33. Wong, R. S. C. & Bennick, A. *Journal of Biological Chemistry* 255, 5943-5948 (1980).
34. Bennick, A., McLaughlin, A. C., Grey, A. A. & Madapallimattam, G. *Journal of Biological Chemistry* 256, 4741-4746 (1981).
35. Hay, D. I., Carlson, E. R., Schluckebier, S. K., Moreno, E. C. & Schlesinger, D.H. *Calcified Tissue International* 40, 126-132 (1987).
36. Hay, D. I., Moreno, E. C. & Schlesinger, D. H. *Inorganic Perspectives in biology and Medicine* 2, 271-285 (1979).
37. Moreno, E. L., Varaghese, K. & Hay, D. I. *Calcified Tissue International* 28, 7-16 (1979).
38. Hatton, M. N., Loomis, R. E., Levine, M. J. & Tabak, L. A. *Biochemistry Journal* 230, 817-820 (1985).
39. Bergey, E. J., Levine, M. J., Reddy, M. S., Bradway, S. D. & Al-Hashimi, I. *Biochemistry Journal* 234, 43-48 (1986).
40. Ayad, M., Van Wuyckhuysse, B. C. & Minaguchi, K. *J. Dent. Res.* 79, 979-982 (2000).
41. Robinovitch, M. R. et al. *Oral Diseases* 7, 86-93 (2000).
42. Beeley, J. A. *Oral Diseases* 7, 69-70 (2001).
43. Mehansho, H., Ann, D. K., Butler, L. G., Rogler, J. & Carlson, D. M. *Journal of Biological Chemistry* 262, 12344-12350 (1987).
44. Carlson, D. M. *Biochimie* 70, 1689-1695 (1988).

45. McArthur, C., Sanson, G. D. & Beal, A. M. *Journal of Chemical Ecology* 21, 663-691 (1995).
46. Mehansho, H. et al. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* 80, 3948 (1983).
47. Lu, Y. & Bennick, A. *Archive of Oral Biochemistry* 43, 717-728 (1998).
48. Warner, T. F. & Azen, E. A. *Medical Hypotheses* 26, 99-102 (1988).
49. Robinson, R. et al. *Biochemistry Journal* 263 (1989).
50. Azen, E. A., Hellekant, L. M., Sabatini, L. N. & Warner, T. F. *Journal of Dental Research* 69, 1724-1730 (1990).
51. Azen, E. A. *Biochemistry Genetic* 9, 69-85 (1973).
52. Kauffman, D. L., Bennick, A., Blum, M. & Keller, P. J. *Biochemistry* 30, 3351-3356 (1991).
53. Saitoh, E., Isemura, S. & Sanada, K. *Journal of Biochemistry* 93, 883-888 (1983).
54. Saitoh, E., Isemura, S. & Sanada, K. *Journal of Biochemistry* 94, 1991-1999 (1983).
55. Shibata, A. et al. *International Journal of Peptide and Protein Research* 23, 158-165 (1984).
56. Bertuccioli, M. & Ferrari, S. in *Colloids and mouthfeel in wines* (ed. Lallemand) 51-54 (1999).
57. Swain, T. & Bate-Smith, E. C. in *Comparative Biochemistry* (eds. Florin, M. & Mason, H. S.) (Academic Press, London, 1962).
58. Haslam, E., Lilley, T. H., Cai, Y., Martin, R. & Magnolato, D. *Planta Medica* 55, 1-8 (1989).
59. Prieur, C., Rigaud, J., Cheynier, V. & Moutounet, M. *Phytochemistry* 36, 781-784 (1994).
60. Ribéreau-Gayon, P. *Annales de Physiologie Végétale* 6, 211-242 (1964).
61. Souquet, J. M., Cheynier, V., Brossaud, F. & Moutounet, M. *Phytochemistry* 43, 509-512 (1996).
62. Foo, L. Y. & Porter, L. J. *Phytochemistry* 19, 1747-1754 (1980).
63. Barathieu, K. (Université Bordeaux 1, Bordeaux, 2002).
64. Feeny, P. P. *Ecology* 15, 565 (1970).
65. Feeny, P. P. *Recent Advances in phytochemistry* 10, 1 (1976).
66. Yokazawa, T. et al. *Journal of Agricultural and Food Chemistry* 46, 2143-2150 (1998).
67. Hagerman, A. E., Rice, M. E. & Ritchard, N. T. *Journal of Agricultural and Food Chemistry* 46, 2590-2595 (1998).

68. Okuda, T., Yoshida, T. & Hatamo, T. in *Phenolics compounds in food and their effects on health II: antioxydant and cancer prevention* (eds. Huang, M. T., Ho, C. T. & Lee, C. Y.) (1992).
69. Mitjavila, S., Lacombe, C., Carrera, G. & Derache, R. *Journal of Nutrition* 107, 2113-2121 (1977).
70. Ahmed, A. E., Smtihard, R. & Ellis, M. *British Journal of Nutrition* 65, 189-197 (1991).
71. Oterdoom, H. J. *Lancet* 2, 330 (1985).
72. Chung, K. T., Wong, T. Y., Wei, C. I., Huang, Y. W. & Lin, Y. *Critical Reviews in Food Science and Nutrition* 38, 421-464 (1998).
73. Mehansho, H., Clements, S., Sheares, B. T., Smith, S. & Carlson, D. M. *Journal of Biological Chemistry* 260, 4418-4423 (1985).
74. Butler, L. G., Mehansho, H. & Carlson, D. M. *Annual Review of Nutrition* 7, 423 (1987).
75. Butler, L. G. in *Toxicants of plant origin* (ed. Cheeke, P. R.) 95-121 (1989).
76. Haslam, E. *Journal of Chemical Society, Chemical Communication* 309 (1981).
77. Saucier, C. (Victor Segalen Bordeaux 2, 1997).
78. Ribéreau-Gayon, P. *Traité d'oenologie. 2. Chimie du vin, stabilisation et traitement* (1998).
79. McManus, J. P., Davis, K. G., Lilley, T. H. & Haslam, E. *Journal of Chemical Society. Chemical Communications* 7, 309-311 (1981).
80. Luck, G. et al. *Phytochemistry* 37, 357-371 (1994).
81. Hagerman, A. E. in *Phenolic compounds in food and their effects on health I. Analysis, occurrence and chemistry* (eds. Ho, C. T., Lee, C. Y. & Huang, M. T.) 236-247 (Washington, 1992).
82. Cid, H., Vargas, V., Bunster, M. & Bustos, S. *FEBS* 198, 140-144 (1986).
83. Asano, K., Shinagawa, K. & Hashimoto, N. *American Society of Brewing Chemists* 40, 147 (1982).
84. Haslam, E. *Biochemical Journal* 139, 285-288 (1974).
85. Murray, N. J., Williamson, M. P., Lilley, T. H. & Haslam, E. *European Journal of Biochemistry* 219, 923-935 (1994).
86. Chen, Y. H., Yang, J. T. & Chau, K. H. *Biochemistry* 13, 3350-3359 (1974).
87. Berova, N., Nakanishi, K. & Woody, R. W. (ed. Wiley-VCH) 1-877 (2000).
88. Johnson, W. C. J. *Proteins* 7, 205-214 (1990).
89. Greenfield, N. J. *Analytical Biochemistry* 235 (1996).
90. Fasman, G. D. *Biopolymers* 37, 339-362 (1995).

91. Cascio, M. & Wallace, B. A. *Analytical Biochemistry* 227, 90-100 (1995).
92. Shibata, A. et al. *Biochemistry* 31, 5728-5733 (1992).
93. Castano, S., Cornut, I., Büttner, K., Dasseux, J. L. & Dufourcq, J. *Biochimica et Biophysica Acta* 1416, 161-175 (1999).
94. Castano, S., Desbat, B., Laguerre, M. & Dufourcq, J. *Biochimica et Biophysica Acta* 1416 (1999).
95. Sreerama, N., Venyaminov, S. Y. & Woody, R. W. *Protein Science* 8, 370-389 (1999).
96. Hillenkamp, F., Karas, M., Ingeldoh, A. & Stahl, B. in *Biological Mass Spectrometry* (eds. Burligame, A. L. & McCloskey, J. A.) (Elsevier, Amsterdam, 1990).
97. Yamashita, M. & Fenn, J. B. *Journal of Physical Chemistry* 88, 4451-4671 (1984).
98. Rusconi, F., Schmitter, J. M., Rossier, J. & le Maire, M. *Analytical Chemistry* 70, 3046-3052 (1998).
99. Fukushima, E. & Roeder, S. B. W. *Experimental pulse NMR* (ed. Addison-Wesley) (Massachusetts, 1981).
100. Harris, R. K. *Nuclear magnetic resonance spectroscopy* (ed. Longman) (1986).
101. Canet, D. *La RMN. Concepts et méthodes* (ed. InterEditions) (1991).
102. Wuthrich, K. *NMR of protein and nucleic acids* (Wiley, New York, 1986).
103. Nagayama, K., Kumar, A., Wüthrich, K. & Ernst, R. R. *Journal of Magnetic Resonance* 40, 321-334 (1980).
104. Braunschweiler, L. & Ernst, R. R. *Journal of Magnetic Resonance* 53, 521-528 (1983).
105. Noggle, J. H. & Shirmer, R. E. *The nuclear overhauser effects* (Academic Press, New York, 1971).
106. Bothner-By, A. A., Stephens, R. L., Lee, J. M., Warren, C. D. & Jeanloz, R. W. *Journal of American Chemical Society* 101, 811-813 (1984).
107. Bax, A. & Davis, D. G. *Journal of Magnetic Resonance* 63, 207-213 (1985).
108. Eccles, C., Güntert, P., Billeter, M. & Wüthrich, K. *J. Biomol.* 1, 111-130 (1991).
109. Bystrov, V. F. *Program NMR Spectroscopy* 10, 41 (1976).
110. Piantini, U., Sorensen, O. W. & Ernst, R. R. *Journal of American Chemical Society* 104 (1982).
111. Rance, M. et al. *Biochemical and Biophysical Research Communication* 117, 263-268 (1983).
112. Pardi, A., Billeter, M. & Wüthrich, K. *Journal of Molecular Biology* 180, 741-751 (1984).

113. Günter, P., Mumenthaler, C. & Wüthrich, K. *Journal of Molecular Biology* 273, 283-298 (1997).
114. Baxter, N. J., Williamson, M. P., Lilley, T. H. & Haslam, E. *Journal of Chemical Society (Faraday Trans)* 92, 231-234 (1996).
115. Stejskal, E. O. & Tanner, J. E. *Journal of Chemistry Physic* 42, 288 (1965).
116. Wu, D., Chen, A. & Johnson, C. S. *Journal of Magnetic Resonance* 115, 260-264 (1995).
117. Guéron, M., Plateu, P. & Decors, M. *Program NMR Spectroscopy* 23, 135-209 (1991).
118. Piotto, M., Saudek, V. & Sklenar, V. *Journal of Biomolecular NMR* 2 2, 661-666 (1992).
119. Mirabel, M., Glories, Y., Pianet, I. & Dufourc, E. J. *Journal de Chimie Physique* 96, 1629-1634 (1999).
120. Weiner, S. J. et al. *Journal of American Chemical Society* 106, 765 (1984).
121. Brooks, C. L., Montgomery, Pettitt & Karplus, M. *Journal of Chemical Physic* 83, 5897-5908 (1985).
122. Jacoby, S. L. S., Kowalik, J. S. & Pizzo, J. T. *Iterative methods for nonlinear optimization problems* (New Jersey, 1972).
123. Fletcher, R. & Reeves, C. M. *Computer J.* 7, 149 (1964).
124. Rabanal, F., Ludevid, M. D., Pons, M. & Giralt, E. *Biopolymers* 33, 1019-1028 (1993).
125. Goetz, M., Schmitter, J. M., Geoffre, S. & Dufourc, E. J. *Journal of Peptide Science* 5, 245-250 (1999).
126. Schwyzer, R. et al. *Helvetica Chimica Acta* 55, 253-254 (1972).
127. Murray, N. J. & Williamson, M. P. *European Journal of Biochemistry* 219, 915-921 (1994).
128. Baxter, N. J., Lilley, T. H., Haslam, E. & Williamson, M. P. *Biochemistry* 36, 5566-5577 (1997).
129. Hatano, T., Yoshida, T. & Hemingway, R. W. in *Plant polyphenol 2 : Chemistry, Biology, Pharmacology, Ecology* (eds. Gross, G. G., Hemingway, R. W. & Yoshida, T.) 509-526 (Plenum Publishers, New York, 1999).
130. Charlton, A. J. et al. *FEBS Letters* 382, 289-292 (1996).
131. Bacon, J. R. & Rhodes, M. J. C. *Journal of Agricultural and Food Chemistry* 48, 838-843 (2000).
132. Kallithraka, S., Bakker, J. & Clifford, M. N. *Journal of the Science of Food and Agriculture* 81, 261-268 (2000).
133. De Freitas, V. & Mateus, N. *Journal of Agricultural and Food Chemistry* 49, 940-945 (2001).

134. Merrifield, R. B. *Journal of American Chemical Society* 85, 2149-2154 (1963).
135. Dourtoglou, V., Gross, B., Lambropoulou, V. & Zioudrou, C. *Synthesis*, 572-574 (1984).
136. Fields, C. G., Lloyd, D. H., MacDonald, R. L., Otteson, K. M. & Noble, R. L. in *Peptides 1990* (eds. Giralt, E. & Andreu, D.) 65-66 (ESCOM, Leiden, 1991).
137. Barany, G. & Merrifield, R. B. in *The Peptides* (eds. Gross, E. & Meienhofer, J.) 1-284 (Academic Press, New York, 1979).
138. König, W. & Geiger, R. in *Chemistry and biology of peptides* (ed. Meienhofer, J.) 343-350 (Sciences Publishers, MI, 1972).
139. Rich, D. H. & Singh, J. in *The Peptides* (ed. Meienhofer, J.) 241-261 (Academic Press, New York, 1979).
140. Fields, C. G., Lloyd, D. H., MacDonald, R. L., Otteson, K. M. & Noble, R. L. *Peptide Research* 4, 95-101 (1991).
141. Knorr, R., Trzeciak, A., Bannwarth, W. & Gillesen, D. *Tetrahedron Letters* 30, 1927-1930 (1989).
142. Atherton, E. & Sheppard, R. D. *Solid phase peptide synthesis - A practical approach* (Oxford University Press, IRL, 1989).
143. Baba, T., Sugiyama, H. & Seto, S. *Chemical & Pharmaceutical Bulletin* 21, 207-209 (1973).
144. Tam, J. P., Riemen, M. W. & Merrifield, R. B. *Peptide Research* 1, 6-18 (1988).
145. De Freitas, V. (*Oenologie - Bordeaux II*, Bordeaux, 1995).
146. Mirabel, M. in *Oenologie* (Bordeaux 2, Bordeaux, 2000).
147. Kiatgrajai, P., Wellons, J. D., Gollob, L. & White, J. D. *Journal of Organic Chemistry* 47, 2910-2912 (1982).
148. Tückmantel, W., Kozinowski, A. P. & Romanczyk, J. *Journal of American Chemical Society* 121, 12073-12081 (1999).
149. Tiffany, M. L. & Frimm, S. *Biopolymers* 6, 1767-1770 (1968).
150. Traub, W. & Yonath, A. *Journal of Molecular Biology* 16, 404-414 (1966).
151. Traub, W., Shmueli, U., Suwalsky, M. & Yonath, A. in *Conformation of Biopolymers* (ed. Ramachandran, G. N.) 449-467 (Academic Press, New York, 1967).
152. Nordbo, A. H., Darwish, S., Sorensen, K. R. & Bhatnagar, R. S. *Journal of Dental Research* 63, 227 (1984).
153. Traub, W. & Shmueli, U. *Nature* 176, 501-503 (1963).
154. Naganagowda, G. A., Gururaja, T. L. & Levine, M. J. *Journal of Biomolecular Structure & Dynamics* 16, 91-107 (1998).

155. McArthur, M. W. & Thornton, J. M. *Journal of Molecular Biology* 218, 397-412 (1991).
156. Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A. & Brodsky, B. *Biochemistry* 39, 14960-14967 (2000).
157. Van Host, G. J. & Varner, J. E. *Plant Physiol.* 74, 247-251 (1984).
158. Adzhubei, A. A. & Sternberg, M. J. E. *Journal of Molecular Biology* 229, 472-493 (1993).
159. Szabo, Z. (<http://cryst.bbk.ac.uk/pp97/assignments/projects/szabo/intro.htm>, 1997).
160. Fox, D. G., Cary, P. D. & Kneale, G. G. *Biochimica et Biophysica Acta*, 138-146 (1999).
161. Marion, D. & Wüthrich, K. *Biochemical and Biophysical Research Communication* 113, 967-974 (1983).
162. Derome, A. & Williamson, M. *Journal of Magnetic Resonance* 88, 177-185 (1990).
163. Bax, A. & Davis, D. G. *Journal of Magnetic Resonance* 65, 355-360 (1985).
164. Malliavin, T. E., Pons, J. L. & Delsuc, M. A. *Journal of Biomolecular NMR* 8, 445-452 (1996).
165. Malliavin, T. E., Pons, J. L. & Delsuc, M. A. *Bioinformatics* 14, 624-631 (1998).
166. Marion, D., Ikura, M., Tschudin, R. & Bax, A. *Journal of Magnetic Resonance* 85, 393-399 (1989).
167. Van Gunsteren, W. F. & Berendsen, H. J. C. *Molec. Simul.* 1, 173 (1988).
168. Mohamadi, F. et al. *J. Comp. Chem.* 11, 441 (1990).
169. Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrikson, T. *Journal of American Chemical Society* 112, 6127 (1990).
170. Aznar Derunes, C. (Université Bordeaux 1, Bordeaux, 2001).
171. Bax, A. *Two dimensional nuclear magnetic resonance in liquids* (Reidel, Amsterdam, 1982).
172. Kainosho & Tsuji. *Biochem.* 21, 6273 (1982).
173. Kay, B. K., Williamson, M. P. & Sudol, M. *Faseb Journal* 14, 231-41. (2000).
174. Wolfenden, R. *Science* 222, 1087 (1983).
175. Fernandez, J. & Lilley, T. H. *Journal of Chemical Society (Faraday Trans)* 88, 2503-2509 (1992).
176. Hatano, T. & Hemingway, R. W. *Journal of Chemical Society. Chemical Communications* 6, 2537 (1996).
177. Clifford, M. N. in *Interaction of food components* (eds. Birch, G. G. & Lindley, M. G.) 143-163 (Elsevier Science, London, 1986).

178. Oh, H. I., Hoff, J. E., Armstrong, G. S. & Haff, L. A. *Journal of Agricultural and Food Chemistry* 28, 394-398 (1980).
179. Artz, W. E., Bishop, P. D., Dunker, A. K., Schanus, E. G. & Swanson, B. G. *Journal of Agricultural and Food Chemistry* 35, 417-421 (1987).
180. Peleg, H., Gacon, K., Schlich, P. & Noble, A. C. *Journal of the Science of Food and Agriculture* 79, 1123-1128 (1999).
181. Siebert, K. J., Troukhanova, N. V. & Lynn, P. Y. *Journal of Agricultural and Food Chemistry* 44, 80-85 (1996).
182. Charlton, A. J. et al. *Journal of Agricultural and Food Chemistry* 50, 1593-1601 (2002).
183. Hagerman, A. E. & Butler, L. G. *Journal of Agricultural and Food Chemistry* 28, 944-947 (1980).
184. Lea, A. G. H. & Arnold, G. M. *Journal of the Science of Food and Agriculture* 29, 478-483 (1978).
185. Asano, K., Ohtsu, K., Shinagawa, K. & Hashimoto, N. *Agricultural and Biological Chemistry* 48, 1139- (1984).
186. Kallithraka, S., Bakker, J. & Clifford, G. *Journal of Agricultural and Food Chemistry* 45, 2211-2216 (1997).
187. Fickel, J., Pitra, C., Joest, B. A. & Hofmann, R. R. *Comparative Biochemistry and Physiology Part C* 122, 225-229 (1999).
188. Perez-Maldonado, R. A., Norton, B. W. & Kerven, G. L. *Journal of the Science of Food and Agriculture* 69, 291-298 (1995).
189. Oh, H. I. & Hoff, J. E. *Journal of Food Science* 52, 1267-1269 (1987).
190. Martin, M. M., Rockholm, D. C. & Martin, J. C. *Journal of Chemical Ecology* 11 (1985).
191. Shenkin, P. S. & McDonald, D. Q. *Journal of Computational Chemistry* 15, 899-916 (1994).
192. Audry, E., Dubost, J. P., Colleter, J. C. & Dallet, P. *European Journal of Medical Chemistry* 21, 71-72 (1986).
193. Fauchère, J. L., Quarendon, P. & Kaetterer, L. *Journal of Molecular Graphics* 6, 203-206 (1988).
194. Broto, P., Moreau, G. & Vandycke, C. *European Journal of Medicinal Chemistry* 19, 71-78 (1984).
195. Laguerre, M., Saux, M., Dubost, J. P. & Carpy, A. *Pharmaceutical Sciences* 3, 217-222 (1997).
196. Luo, R. S., Liu, M. L. & Mao, X. A. *Spectrochimica Acta Part A* 55, 1897-1901 (1999).
197. Loo, J. A. *Mass Spectrometry Reviews* 16, 1-23 (1997).

198. Robinson, C. V. et al. *Journal of American Chemical Society* 118, 8646-8653 (1996).
199. Barathieu, K. et al. in *XXI International Conference on Polyphenols* (ed. Hadrami, I. e.) 551-552 (Marrakech, Morocco, 2002).
200. Creighton, T. E. *Biochemistry* 19, 4308-4312 (1980).
201. Vergé, S. et al. *Tetrahedron Letters* 43, 2363-2366 (2002).
202. Haslam, E. *Journal of Natural Products* 59, 205-215 (1996).
203. Rose, G. D., Geselowitz, A. R., Lesser, G., Lee, R. & Zehfus, M. H. *Science* 229, 9834 (1985).
204. Simon, C. et al. *Biochemistry* 42, 10385-10395 (2003).

## **PUBLICATIONS**

## Synthesis and Circular Dichroism Study of the Human Salivary Proline-Rich Protein IB7

CÉCILE SIMON,^a ISABELLE PLANET^{a,b} and ERICK J. DUFOURC^{a*}

^a Institut Européen de Chimie et Biologie, FRE CNRS 2247, Pessac, France

^b Laboratoire de Chimie Organique et Organométallique, UMR CNRS 5802, Talence, France

Received 18 July 2002

Accepted 2 August 2002

**Abstract:** The solid phase synthesis of a 59 amino acid human salivary protein IB7 has been accomplished using Fmoc strategy. Because the protein contains 25 proline, 13 glycine and 9 glutamine residues the coupling time, piperidine delivery and acetic anhydride reaction time were increased. Yield after HPLC purification was 35%. Circular dichroism studies revealed that about one third of IB7 residues adopted a type II helix secondary structure, as found in collagen helices. The rest of the sequence adopts a random coil secondary structure. Copyright © 2003 European Peptide Society and John Wiley & Sons, Ltd.

**Keywords:** proline-rich protein; salivary protein; solid phase synthesis; circular dichroism; type II helix

### INTRODUCTION

Proline-rich proteins (PRP) are ubiquitous and constitutively expressed in the human body. They are involved in numerous enzymatic reactions or in biological recognition [1]. They are predominant in human saliva and represent up to 70% of the proteins secreted by the parotid gland [2]. About 20 have been identified from a single human donor and sorted into three groups: acidic, alkaline and glycosylated proteins [3]. Acidic PRPs are involved in calcium-binding and inhibit crystal growth [4]. Glycosylated proteins could have lubricating properties especially in the mouth [5]. Basic PRPs do not bear any known biological functions but bind polyphenols very efficiently [4,6]. Nine of them have been sequenced and contain about 40% proline, 21% glycine and 17% glutamine residues. The basic PRP IB7 serves as a building unit for other members of the family because its sequence is often contained

in their amino-terminal part [3]. Interestingly, IB7 itself is made of three repeated sequences and of five PGGPP patterns as sketched in Figure 1.

On the other hand, many plant species contain complex polyphenols also named tannins. They act as chemical defences [6,7] and are divided in condensed tannins, i.e. polymers of procyanidin and hydrolysable tannins (mainly gallate tannins). Grape seeds contain a large amount of condensed tannins that are partially responsible for the organoleptic properties of grapes and wines (e.g. astringency and bitterness). During ingestion, tannins interact with many kinds of proteins present in saliva, stomach, intestines and blood [8,9]. The interaction between tannins and salivary proteins is thought to be the primary source of astringency [10,11]. The high affinity of PRPs for polyphenols has been attributed to their open and extended structure and to their high proline content. Murray and coworkers [12] propose by analogy with PRPs studied so far that IB7 would also possess a random coil structure. Other works put forward the existence of a poly(L)proline type II conformation (collagen helix) [13] owing to the high number of proline residues. The favourable interaction between tannins and IB7 has been

* Correspondence to: Dr Erick J. Dufourc, IECB-Polytechnique, 16 Avenue Pey Berland, 33607 Pessac, France.  
 e-mail: erick.dufourc@iecb-polytechnique.u-bordeaux.fr  
 Contract/grant sponsor: Conseil Interprofessionnel du Vin de Bordeaux (CIVB).

SPFGKIQGPPQGGNQPGPP  
PPFGKIQGPPQGGNKPQGGP  
PPGKIQGPPQGGDNKSR²⁹

Figure 1 Amino acid sequence of the human salivary proline-rich protein IB7. Underlined fragments are present five times in the total sequence.

proposed to be due to the short size of the protein [6] and to the presence of the repeated (5 times) PGGPP pattern in its sequence.

All the above-mentioned studies of PRP structure and interaction with polyphenols were carried out on peptides of about 20 residues long. Currently the interaction between polyphenols and salivary proteins is being studied and as a representative element of the salivary PRP family, the protein IB7 was chosen. It is a 59 residue protein and because its sequence is contained in 5 of the 9 human basic salivary proteins it is believed to be a good model protein to follow the interaction. The present study reports the challenging solid-phase chemical synthesis of the proline-rich protein IB7 as well as its secondary structure characterization by circular dichroism.

## MATERIALS AND METHODS

### Peptide Synthesis

The chemical synthesis of IB7 peptide was performed on an Applied Biosystems Peptide Synthesiser 433A (PE Biosystem, Courtaboeuf, France) using Fmoc strategy [14]. Fluorenylmethylcarbonyl (Fmoc)-L-Arg(Pmc)-Wang-resin, (Fmoc)-L-Arg(Pmc)-NovaSyn-resin, *N*- $\alpha$ -Fmoc-protected amino acids, 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HOBt) and *N*-hydroxybenzotriazole (Hbtu) were purchased from Novabiochem (Läufelfingen, Switzerland); *N*-methylpyrrolidone (NMP), piperidine, dichloromethane (DCM), dimethylformamide (DMF), diisopropylethylamine (DIEA), trifluoroacetic acid (TFA) and acetic acid were from SDS (Peypin, France). The *N*- $\alpha$ -Fmoc-amino acids were protected on the side chain as follows: *t*-butyl for serine and aspartic acid, *t*-butoxycarbonyl (Boc) for lysine, trityl (trt) for glutamine and asparagines, and 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) for arginine.

### Cleavage of the Protein from the Resin

The final peptide mixture was cleaved from the resin and removal of the protecting groups accomplished by TFA/thioanisol/Milli Q-water (93/3/2/2 by volume). Total deprotection was achieved by allowing a reaction time of 150 min at 0 °C. TFA was evaporated under reduced pressure. The crude peptide mixture was then precipitated in cold diethyl ether and the cloudy aqueous phase was freeze dried.

### Purification and Analysis

The powder was dissolved in distilled water and analysed by HPLC (Waters 2487 dual wavelength absorbance detector) using a Nova Pak C₁₈ column (5  $\mu$ m, 100 Å, 3.9  $\times$  150 mm) reverse-phase column (Saint-Quentin-en-Yvelines, France). The elution was carried out with 0.1% aqueous TFA (eluent A) and acetonitrile (0.08% TFA, eluent B) at a flow rate of 1 ml/min. Absorption was monitored at 214 and 254 nm. Maximum resolution for the synthetic peptide mixture was achieved with a 30 min linear gradient, where eluent B varied from 18% to 45%. Semi-preparative HPLC was performed with a Waters Delta Pak C₁₈ column (15  $\mu$ m, 100 Å, 300  $\times$  7.8 mm) at a flow rate of 2.5 ml/min under the same elution conditions.

### Mass Spectroscopy (MS)

Matrix assisted laser desorption and ionization time of flight (MALDI TOF) spectra were obtained on a Reflex III mass spectrometer (Bruker Daltonik, Bremen, Germany). The crude product was loaded on the target using the dried droplet method [15]. Sinapic acid was chosen as the matrix. Spectra were externally calibrated using the [M+H]⁺ ions of two peptide standards. The amount of peptide deposited on each spot was typically 3–5 pmol.

### Circular Dichroism (CD)

Far-ultraviolet (190–270 nm) circular dichroism spectra were recorded on a Mark VI Jobin-Yvon dichrograph (Longjumeau, France), calibrated using iso-androsterone (Roussel-Uclaf, France) in dioxane and camphorsulfonic acid (Sigma-Chimie, France). The protein was dissolved in different solvents: water or water/ethanol (88/12 v/v) to mimic the composition of wine. Each one was done in the absence or presence of CaCl₂. All experiments were performed at room temperature, pH 3.5, and by using cells of

1 and 0.1 mm path length. Poly(L)proline (PLP, average amino acid number of 5500, Sigma, St Louis, MO, USA) was dissolved under the same conditions, and home made poly(L)lysine (20 amino acids long, PLL₂₀) was prepared at pH 7.

## RESULTS AND DISCUSSION

### Chemical Synthesis

The feasibility of a peptide solid phase synthesis depends on the nature and frequency of amino acid residues in the protein. Hydrophobic peptides are known to be difficult to synthesize because of their poor water solubility [16]. Other amino acids may

also bring some problems. For example, proline promotes a turn formation that may prevent the attachment of the following residue. On the contrary, glycine, which is the least bulky amino acid, has less availability for peptide elongation due to its capacity for entering into the small cavities of the resin beads. The IB7 composition (Figure 1) is made of 59 amino acids among which are 25 prolines, 13 glycines and 9 glutamines. Despite this challenging composition, solid phase synthesis was carried out using the FastMoc strategy [14] in double coupling mode, with 2.5 mmol of resin and a four-fold excess of amino acids. In order to optimize the coupling yield several steps needed to be modified. The coupling time between amino acids and the peptide in elongation was extended up to 30 min per residue. Likewise,

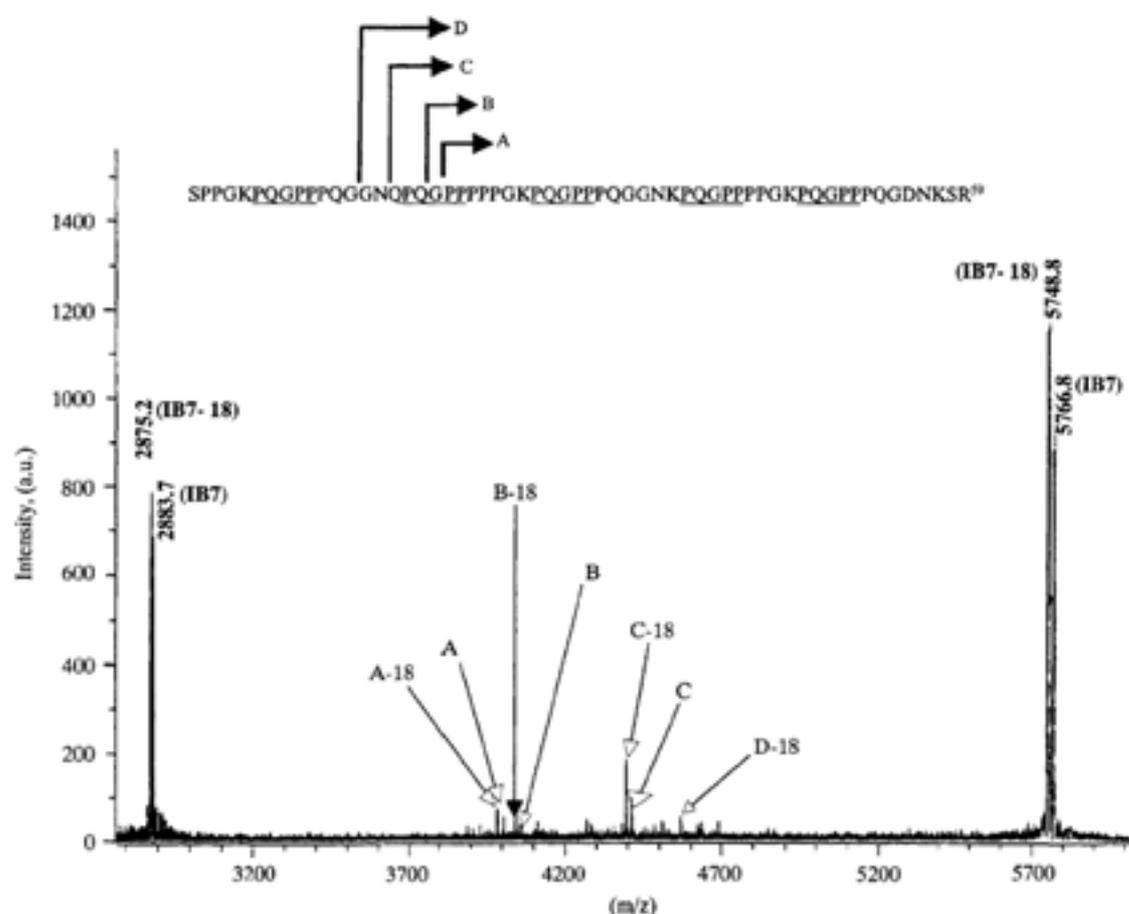


Figure 2 MALDI-TOF MS spectrum of the crude reaction mixture after dissolution in water. Major peaks correspond to the simply (1⁺) and doubly (2⁺) charged species of IB7 and IB7-18. Secondary peaks represent truncated peptides with or without a loss of 18. Some of the truncated peptides are schematized at the top of the figure.

increasing the delivery time of piperidine optimized the yield of the  $\alpha$ -Fmoc-NH₂ groups cleavage. This implementation can be done by continuously monitoring the conductivity change that follows the removal of the Fmoc groups. Because the access to the chain was expected to be difficult, the reaction time of acetic anhydride as a capper was increased (5%). This avoided the formation of hybrid peptides with missing residues. In order to increase the yield it appeared judicious to use a Nova resin rather than the Wang resin classically used for long peptides (Novabiochem, France).

### Purification

The peptide was cleaved from the resin by TFA treatment. In the same cleavage mixture ethanedithiol was added in order to remove the side chain protecting groups. MS (MALDI-TOF) was used to check the content of the crude synthesis products (Figure 2). Two major peaks were identified, each one being split into mono-protonated (1⁺) and di-protonated (2⁺) forms. IB7 has an exact mass of 5765.9 g.mol⁻¹ that corresponds to the peak of 5766.8 *m/z* for the (1⁺) form and of 2883.7 *m/z* for the (2⁺). Peaks at *m/z* of 5748.8 and 2875.2 correspond to the mass of the protein minus 18 for (1⁺) and (2⁺) forms, respectively. This deserves a comment: a central problem in the chemical synthesis of aspartic acid-containing protein is aspartimide formation [17]. Some factors promote this reaction such as acidic solvent, and the presence of Asp-Gly or Asp-Asn sequences that might involve imide formation. Because residue 55 in IB7 is aspartate, followed by asparagine the conditions favour aspartimide formation. Its presence could be avoided in decreasing the cleavage time with TFA. Unfortunately, this would render the removal of some protecting groups very difficult. Because aspartate is in position 55 in the IB7 sequence, i.e. almost at the protein end, aspartimide formation is expected to have little importance for further structural studies.

In the MS spectrum other minority peaks are present and correspond to the truncated peptides of the bulk reaction mixture. Figure 3 shows some of them, e.g. acetylated fragments, from Arg⁵⁹-Pro²⁰ to Arg⁵⁹-Gln¹⁴ with *m/z* ranging from 3984.5 to 4584.3. The presence of these peptides could be explained by the occurrence of five consecutive proline residues from position 22–18 (see Figure 3), which could prevent efficient elongation due to possible turns formation.

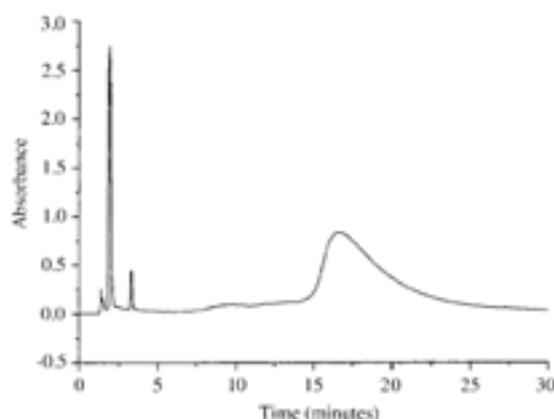


Figure 3 HPLC profile of the crude product using reverse phase chromatography on an analytic C18 column. Elution took place within 30 min using a linear gradient where eluent B (see text) varied from 18% to 45%. The first peaks in the chromatogram originated from the protecting groups and the major fraction eluted between 15 and 20 min contained a mixture of IB7 and truncated peptides.

A very broad and major peak mainly constitutes the resulting HPLC profile (Figure 4). Indeed, all of the peptides (including the truncated ones) that were obtained after cleavage from the resin were eluted at similar elution times because of their composition was the same. The purification was carried out by a fine fractionation of the major peak, and the purity in each fraction was checked by MALDI TOF mass spectrometry. Figure 4 shows a MS spectrum of a fraction that contained only IB7 and IB7-18. Both peptides were non-dissociable by the HPLC. The total yield after purification was 35%.

### Secondary Structure Analysis by Circular Dichroism

Circular dichroism was used to assess the secondary structure of the synthetic protein. The IB7 spectrum in water is presented in Figure 5 with a minimum ellipticity at 200 nm of  $[\theta] = -20000$  deg.dmol⁻¹.cm². No change was noticed when increasing the peptide concentration from 50  $\mu$ M to 2 mM. Because several studies [12,18] predicted a random coil conformation for PRP proteins IB7 was compared with poly(L)lysine (PLL₂₀) that adopts a random coil under specific conditions such as neutral pH at room temperature. As IB7 contains 40% proline, it was also compared with the left-handed type II helical structure of the poly(L)proline, PLP

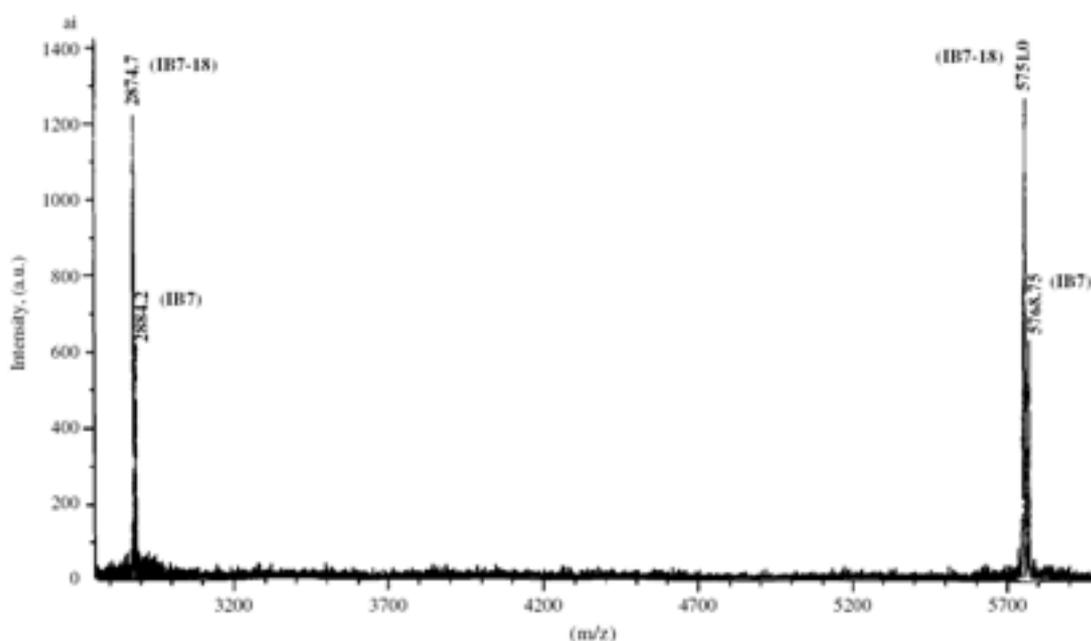


Figure 4 MALDI TOF spectrum of the purified peptide after fractionation by HPLC. The two major species IB7 and IB7-18 are well characterized by their (1⁺) and (2⁺) ions.

(see Figure 5). PLL₂₀ exhibited a minimum ellipticity at 196 nm of  $[\theta] = -12000 \text{ deg.dMol}^{-1}.\text{cm}^2$ , whereas PLP showed a minimum at 206 nm of  $[\theta] = -45000 \text{ deg.dMol}^{-1}.\text{cm}^2$ . The IB7 spectrum clearly stands in between. The IB7 curve was submitted to a spectral deconvolution that accomplishes linear combinations of random, alpha helix, beta sheet and helix II CD standard curves (FORTRAN routine, EJ Dufourc, unpublished). Use of 64% random coil, 0% alpha helix, 0% beta sheet and 36% helix II standard curves could best simulate the IB7 experimental CD data. Interestingly the proline content is 42% and the amount of proline residues engaged in sequences of three or more proline blocks is 31%. This suggests a strong relationship between the number of proline blocks and a helix of type II secondary structure.

IB7 was also dissolved in water/ethanol (88/12) in order to mimic the composition of wine. Similar results were obtained in water and water/ethanol (data not shown). Because calcium ions are known to promote secondary structure randomization [19] CD spectra were recorded in the presence of 4 mM CaCl₂. As seen in Figure 5 calcium addition did not change the observed CD profile. As a control, the CD spectrum of PLP were recorded also in the presence

of the same calcium concentration (Figure 5). No change was observed, suggesting that 4 mM CaCl₂ is not able to denature helices of type II.

## CONCLUSIONS

This is the first time, to our knowledge, that an entire proline-rich human protein has been chemically synthesized with the correct yield in spite of a large number of proline and glycine residues in the sequence. Aspartimide formation at the end of the sequence could not be avoided, but is expected not to be crucial for structural studies. A major finding is that the secondary structure of IB7 is not entirely in a random coil, as Murray and co-workers would have predicted. About one third of its amino acids are in a helix of type II conformation (collagen helix). It is noteworthy that a similar secondary structure prediction was already reported for IB9, a PRP very similar to IB7, by using CD, proton NMR spectroscopy [20] and by hydrophobicity profile analysis [13]. In order to further detail our findings for IB7, NMR studies combined with molecular modelling are under way in our laboratory.

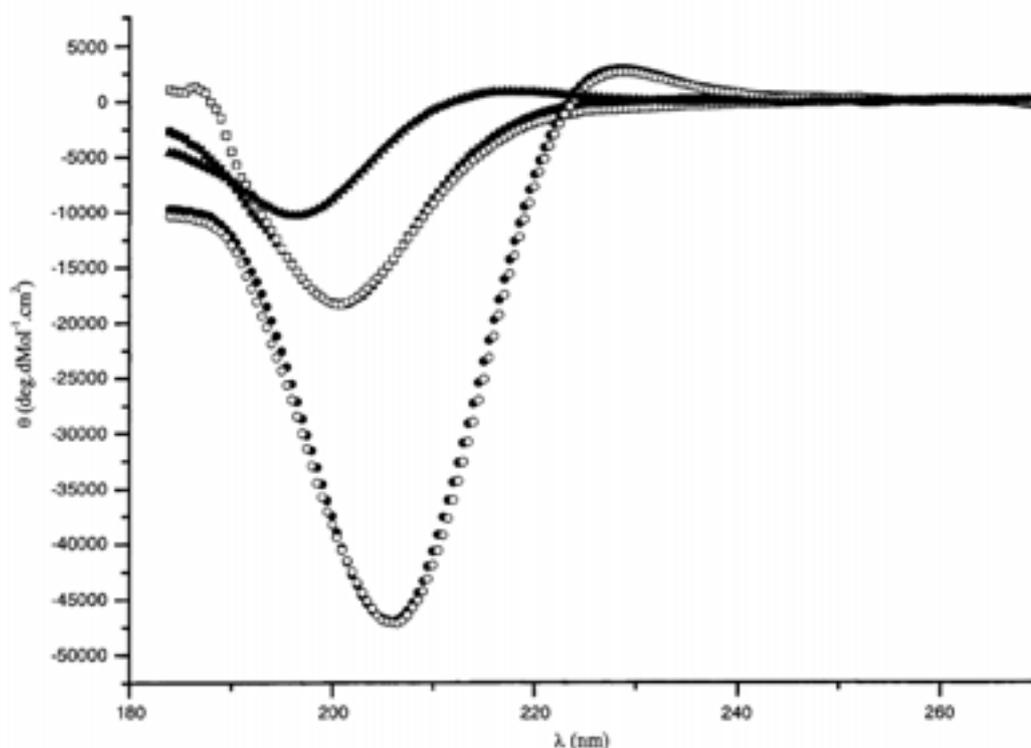


Figure 5 Far-ultraviolet CD spectrum for IB7 (200  $\mu\text{M}$ ) and standards (PLL₂₀, 20  $\mu\text{M}$ , PLP, 16  $\mu\text{M}$ ) in water. Top curve (▲) shows the random coil conformation of a poly(l)lysine of 20 residues at pH 7 and room temperature. Curve at the bottom (●) is characteristic of poly(l)proline (PLP) type II helix. In the middle is shown the spectrum of IB7 (■). Plots with empty symbols (○,□) present the effect of a final concentration of 4 mM CaCl₂.

### Acknowledgements

The authors are indebted to Professor Jean-Marie Schmitter and Katel Bathany for mass spectrometric analyses.

### REFERENCES

1. Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* 2000; **14**: 231–241.
2. Kauffman DL, Keller PJ. The basic proline-rich proteins in human parotid saliva from a single subject. *Arch. Oral Biol.* 1979; **24**: 249–256.
3. Kauffman DL, Bennick A, Blum M, Keller PJ. Basic proline-rich proteins from parotid saliva: relationships of the covalent structures of ten proteins from a single individual. *Biochemistry* 1991; **30**: 3351–3356.
4. Bennick A, McLaughlin AC, Grey AA, Madapallimatam G. The location and nature of calcium-binding sites in salivary acidic proline-rich phosphoproteins. *J. Biol. Chem.* 1981; **256**: 4741–4746.
5. Hatton MN, Loomis RE, Levine MJ, Tabak LA. Masticatory lubrication. The role of carbohydrate in the lubricating property of a salivary-albumin complex. *Biochem. J.* 1985; **230**: 817–820.
6. Lu Y, Bennick A. Interaction of tannin with human salivary proline-rich proteins. *Arch. Oral Biochem.* 1998; **43**: 717–728.
7. Beart JE, Lilley TH, Haslam E. Plant polyphenols — secondary metabolism and chemical defense: some observations. *J. Chem. Soc. Perkin Trans. II* 1985; **24**: 33–38.
8. Haslam E, Lilley TH, Cai Y, Martin R, Magnolato D. Traditional herbal medicines — the role of polyphenols. *Plant. Med.* 1989; **55**: 1–8.
9. Foley WJ, MacArthur C. *The Effects and Costs of Allelochemicals for Mammalian Herbivores: An Ecological Perspective*. Cambridge University Press: Cambridge, 1994; 370–391.

10. Bate-Smith EC. Astringency in foods. *Foods* 1954; **23**: 124.
11. Prinz JF, Lucas FW. Saliva tannin interactions. *J. Oral Rehab.* 2000; **27**: 991-994.
12. Murray NJ, Williamson MP, Lilley TH, Haslam E. Study of the interaction between salivary proline-rich proteins and a polyphenol by 1H-NMR spectroscopy. *Eur. J. Biochem.* 1994; **219**: 923-935.
13. Cid H, Vargas V, Bunster M, Bustos S. Secondary structure prediction of human salivary proline-rich proteins. *FEBS Lett.* 1986; **198**: 140-144.
14. Fields CG, Lloyd DH, MacDonald RL, Otteson KM, Noble RL. HBTU activation for automated Fmoc solid-phase peptide synthesis. *Pept. Res.* 1991; **4**: 95-101.
15. Kussmann M, Nordhoff E, Rahbek-Nielsen H, Haebel S, Rossel-Larsen M, Jakobsen L, Gobom J, Mirgorodskaya E, Kroll-Kristensen A, Palm L, Roepstorff P. Matrix-assisted laser desorption/ionization mass spectroscopy sample preparation techniques designed for various peptide and protein analytes. *J. Mass Spectr.* 1997; **32**: 593-601.
16. Goetz M, Schmitter JM, Geoffre S, Dufoure EJ. Chemical synthesis of yeast mitochondrial ATP synthase membranous subunit 8. *J. Pept. Sci.* 1999; **5**: 245-250.
17. Tam JP, Riemen MW, Merrifield RB. Mechanisms of aspartimide formation: the effects of protecting groups, acid, base, temperature and time. *Pept. Res.* 1988; **1**: 6-18.
18. Murray NJ, Williamson MP. Conformational study of a salivary proline-rich protein repeat sequence. *Eur. J. Biochem.* 1994; **219**: 915-921.
19. Tiffany ML, Frimn S. Circular dichroism of poly-L-proline in an unordered conformation. *Biopolymers* 1968; **6**: 1767-1770.
20. Shibata S, Asakura J, Isemura T, Isemura S, Saitoh E, Sanada K. Conformational study of the basic proline-rich polypeptides from human parotid saliva. *Int. J. Pept. Prot. Res.* 1984; **23**: 158-165.

## Three-Dimensional Structure and Dynamics of Wine Tannin–Saliva Protein Complexes. A Multitechnique Approach[†]

Cécile Simon,[‡] Karine Barathieu,[‡] Michel Laguerre,[‡] Jean-Marie Schmitter,[‡] Eric Fouquet,[‡] Isabelle Pianet,^{*,‡,§} and Erick J. Dufoure[‡]

*Institut Européen de Chimie et Biologie, UMR CNRS 5144, Pessac, France, and Laboratoire de Chimie Organique et Organométallique, UMR CNRS 5802, Talence, France*

*Received March 5, 2003; Revised Manuscript Received June 13, 2003*

**ABSTRACT:** The interactions between the B3 (catechin-4 $\alpha$ ,8-catechin) red wine tannin and the human salivary protein fragment IB7₁₄ (SPPGKPGQGGPPQGG) were monitored by ¹H magic angle spinning NMR, circular dichroism, electrospray ionization mass spectrometry, and molecular modeling. It is found that the secondary structure of IB7₁₄ is made of a type II helix (collagen helix) and random coil. The central glycine 8 appears to act as a flexible rotula separating two helix II regions. Three tannin molecules tightly complex the peptide, without modifying its secondary structure, but seem to reduce its conformational dynamics. The binding dissociation constant is in the millimolar range. B3 tannins with a "tweezers" conformation bind to the hydrophilic side of the saliva peptide, suggesting that the principal driving forces toward association are governed by hydrogen bonding between the carbonyl functions of proline residues and both the phenol and catechol OH groups. These findings are further discussed in the frame of an astringency phenomenon.

Astringency is an important mouth-feel character determining the quality of red wine. This sensation, earlier considered by Aristotle (384–322 B.C.) as a taste, was described in physicochemical terms by Joslyn and Golstein in 1964 (1) as the result of a strong interaction between tannins and salivary proteins. The formed tannin–protein complex aggregates, thus reducing the lubricating property of saliva (2, 3). A dry, rough, and pucker sensation is then perceived as a diffuse stimulus in the entire mouth (4–6).

Tannins responsible for wine astringency are mostly flavan-3-ol polymers commonly referred to as proanthocyanidins or condensed tannins (7); they are extracted from grape skins and seeds during the maceration phase of wine-making and may reach a concentration up to 4 g/L in Bordeaux red wine (8).

About 70% of the human salivary proteins secreted by the parotid gland are proline-rich proteins (PRPs)¹ (9, 10). They are made up of three types: acidic, glycosylated, and

basic (11). The specific role of each type is not clear yet, but recent studies suggest multiple protective functions such as virus binding (12). Acidic PRPs could maintain calcium homeostasis (9, 13), while glycosylated ones could play a role in the oral cavity lubrication (14) and prevent bacterial agglutination (15). Nevertheless, the only known biological function of basic PRPs is to bind polyphenols (16–18). All of these PRPs share a high degree of homology, because they are encoded by a small number of genes (19). Their unique composition, where proline, glycine, and glutamine account for 70–80% of the total amino acid content (20), confers to PRPs an extended and opened structure (21) that may be favorable to tannin binding. Also they are made in repeated patterns of five aa (PQGPP) and nine aa (PPGK-PQGPP). Among them, the IB7 protein appears to be a generic protein whose sequence is almost entirely found in many other PRPs.

Interactions occurring between tannins and proteins were largely studied this past decade. An understanding of this phenomenon is of great interest in order to characterize the gustative sensation called astringency in foods and beverages (22). It is also of interest to control the production of leather or to define the role of tannins in medicine, a fashionable theme after the "French paradox" hypothesis (23). The above studies were carried out by using models such as the proline amino acid alone, the tetrapeptide Gly-Pro-Gly-Gly (24), the nonapeptide bradykinin (25–27), angiotensins I and II (27), PRP typical repeat units (28–30), and even an entire PRP (31–33). For the tannin counterpart, catechin, epicatechin, procyanidin B3, or hydrolyzable tannins such as tri-, tetra- and pentagalloylglucose were used. Various experiments were used in these contributions to highlight protein–tannin interactions, including solution NMR, molecular modeling,

[†] This work is supported by grants from the Conseil Interprofessionnel du Vin de Bordeaux (CIVB).

* Address correspondence to this author. Tel.: +33556846448. Fax: +33556842623. E-mail: ipianet@cesamo.u-bordeaux.fr.

[‡] Institut Européen de Chimie et Biologie.

[§] Laboratoire de Chimie Organique et Organométallique.

¹ Abbreviations: aa, amino acids; CD, circular dichroism; CID, collision-induced dissociation; DQF-COSY, double-quantum filtered correlation spectroscopy; DIPSI, decoupling in the presence of scalar interactions; ESI, electrospray ionization; HOMOHA, homonuclear Hartmann–Hahn; HRMAS, high-resolution magic angle spinning; MALDI, matrix-assisted laser desorption ionization; MS, mass spectrometry; m/z, mass over charge ratio; NMR, nuclear magnetic resonance; PRPs, proline-rich proteins of saliva; RMSD, root mean square deviation; ROESY, rotating frame enhancement spectroscopy; Th, Thomson (unit for m/z ratio); WATERGATE, WATER suppression by gradient-tailored excitation.

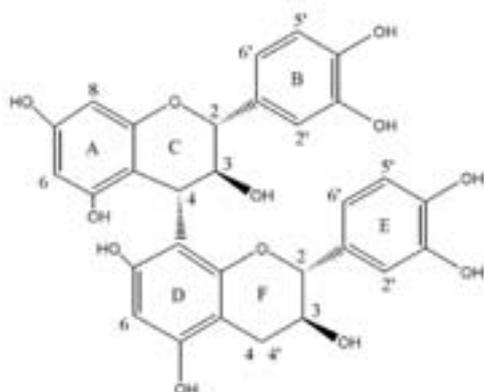


FIGURE 1: Schematic representation of the B3 procyanidin dimer, catechin-4 $\alpha$ ,8-catechin.

mass spectrometry, and gel electrophoresis. Hydrophobic effects have been considered in various papers as the main driving forces toward association (26, 27), probably enhanced by hydrogen bonding between the phenolic groups and the protein, mainly through the carbonyl groups of prolines (34). Nonetheless, it is commonly thought that a specific mode of binding between polyphenols and a PRP is highly improbable. The nature of the interaction probably depends on the nature of the polyphenol, its size, its stereochemistry, the nature of the protein, and the medium in which the interaction takes place. Moreover, one of the difficulties encountered in these studies is the formation of a colloidal insoluble complex in water, requiring to work in a solvent mixture containing DMSO (10–20%), that is far from representing the wine and mouth medium.

In the present work, we describe the interactions of the B3 procyanidin (catechin-4 $\alpha$ ,8-catechin) (Figure 1) with a 14-residue fragment of IB7. Unlike others, we decided to study the association in a medium as close as possible to reality, i.e., water–ethanol (88:12 v/v), pH 3.5. Noninvasive methods such as solid-state NMR by HRMAS, mass spectrometry in MALDI and electrospray modes, circular dichroism, and molecular mechanics have been used to determine the 3D structure, the stoichiometry in the complex, the dissociation constant, and the dynamics of the aggregates.

## MATERIALS AND METHODS

**Chemicals.** Amino acids, solid-phase support, HOBt, and HBTU for solid-phase synthesis were purchased from Novabiochem (Läufelfingen, Switzerland). Catechin was obtained from Sigma (Sigma-Chimie, France), and other solvents for synthesis and HPLC were from SDS (Peypin, France).

**Peptide Synthesis and Purification.** The chemical synthesis of the IB7 peptide (IB7₁₄: SPPGKPPQGG) was performed on an Applied Biosystems peptide synthesizer 433A (PE Biosystem, Courtaboeuf, France) by the FastMoc method on a solid state, using the method that we have described previously (21). After purification by reversed-phase liquid chromatography over a C18 column (Alliance chromatographic system, Waters, Saint-Quentin-en-Yvelines, France), the product was analyzed by MALDI mass spectrometry.

**B3 Synthesis.** The synthesis of the catechin-4 $\alpha$ ,8-catechin dimer, procyanidin B3 (Figure 1), was achieved with the method described by Tückmantel et al. (35), with slight modifications (36, 37).

**Circular Dichroism.** Far-ultraviolet (190–370 nm) circular dichroism spectra were recorded on a Mark VI Jobin-Yvon dichrograph (Longjumeau, France), calibrated with isoandrosterone (Roussel-Uclaf, France) in dioxane and camphor-sulfonic acid (Sigma-Chimie, France) as previously described (38). Measurements were carried out at room temperature with 1 or 0.1 mm path length cells, in water or in a water–ethanol mixture (88:12 v/v), pH 3.5, obtained by dropwise addition of HCl. Titrations were performed as follows: Procyanidin B3 solution (0.025–1 mM final concentration) was progressively added to a 0.1 mM peptide sample, or, at the opposite, the B3 concentration was fixed to 0.1 mM and the peptide concentration was varied from 0.025 mM to 1 mM. Both procedures gave the same result.

**Mass Spectrometry.** Mass analysis of the IB7₁₄ peptide was performed on a MALDI mass spectrometer in the reflectron mode (Bruker Reflex III) using  $\alpha$ -cyano-4-hydroxycinnamic acid as a matrix.

An ion trap mass spectrometer equipped with a nanospray interface (LCQ Deca XP, Thermo Finnigan) was used for peptide–tannin interaction studies. Premixed peptide tannin solutions in water–ethanol–acetic acid (89:10:1 v/v/v), pH 3.5, were infused with needles having 2  $\mu$ m orifices (New Objective) at a 10  $\mu$ M IB7₁₄ peptide concentration. The temperature of the heated capillary in the interface was set to 150 °C, and the spraying voltage was 1.2 kV. Doubly charged species corresponding to peptide–tannin complexes were isolated in the ion trap with a  $\pm 3$  Da isolation width and disrupted using a collision energy of 18–23% on the instrument's scale.

**NMR Spectroscopy.** The samples were prepared as follows: the lyophilized synthetic peptide was dissolved in a H₂O–D₂O mixture (90:10) with or without 12% deuterated ethanol to 1–5 mM concentration. When needed, procyanidin B3 was added in order to obtain the desired concentrations (between 0 and 13 mM). The pH of the solution mixtures was then adjusted to 3.5. Eighty microliters of each sample was then poured into a 4 mm HRMAS rotor. The 1D and 2D NMR spectra were recorded on a Bruker Avance DSX 500 spectrometer using a dual HRMAS probe with ²H lock and equipped with a z*-gradient. Spectra were recorded with a spinning rate of 5000 Hz at 300 K. The water resonance was suppressed using the watergate sequence (39). The 1D proton spectra were recorded for titration using standard conditions (single pulse sequence, 256 scans, recycling delay 3 s). DQF-COSY, HOHAHA, and ROESY spectra were acquired in the phase-sensitive mode with time-proportional phase incrementation (TPPI) of the first pulse (40). The proton 90° pulse was 5.9  $\mu$ s, and the spectral width was set to 10 ppm (5000 Hz) in both dimensions. HOHAHA spectra were recorded using the MLEV pulse scheme with 100 and 150 ms isotropic mixing periods. ROESY spectra were recorded with a 300 ms spin-lock time (0.3 W, 1.5 kHz). The data size was usually 2048 complex points in the t₂ dimension, and 512 t₁ increments were performed. A total of 128 scans were acquired for the HOHAHA spectrum and 256 scans for COSY and ROESY experiments. GIFA software was used to process the NMR data (41, 42). After

zero filling to 2K in both dimensions, the two-dimensional data matrices were multiplied by a shifted sine-bell window function ( $\sin = 0.2$ ) and Fourier transformed. Distance constraints were obtained from cross-peak intensities of ROESY experiments, as determined with the XEASY package (43). Preliminary structure calculations were performed with the aid of the program DYANA (44).

For titration experiments, chemical shift variations of some peptide protons were analyzed as a function of procyanidin B3 concentration using the equation previously described by Charlton et al. (29) for a multisite model:

$$\Delta\delta_i = 0.5\Delta\delta_{\max}(1 + K_d/n[P_i] + [T_i]/n[P_i]) - \{(1 + K_d/n[P_i] + [T_i]/n[P_i])^2 - 4[T_i]/n[P_i]\}^{1/2} \quad (1)$$

where  $\Delta\delta_i$  = change in chemical shift (ppm),  $\Delta\delta_{\max}$  = maximum change in chemical shift (ppm),  $K_d$  = dissociation constant (M),  $[P_i]$  = total concentration of peptide (M),  $n$  = number of polyphenol binding sites, and  $[T_i]$  = total concentration of polyphenol able to fix the peptide. The  $[T_i]$  concentration is measured using the relationships described by Charlton and Baxter (29, 45). Calculations were performed using the Microsoft Excel software. For the curve fitting using eq 1,  $K_d$ ,  $\Delta\delta_{\max}$ , and  $n$  were set as adjustable parameters.

**Molecular Mechanics. (A) Peptide Alone.** Molecular modeling calculations were performed on a SGI Octane R10K workstation running MacroModel (46) version 6.5 (Schrödinger Inc.). Conformational minima were found using the modified AMBER* (1991 parameters) force field as implemented and completed in the MacroModel program. Built structures were minimized to a final RMSD gradient  $\leq 0.005 \text{ kJ}\cdot\text{Å}^{-1}\cdot\text{mol}^{-1}$  via the truncated Newton conjugate gradient (TNCG) method (1000 cycles). Calculations were performed with the GB/SA continuum solvation model (47). The solvent chosen was water. In all cases the extended cutoff option was used throughout (VdW = 8 Å, electrostatic = 20 Å, and H-bond = 4 Å). It must be mentioned that tannins are very simple molecules from a chemical point of view, and hence the corresponding parameters in actual force fields are well-defined and of high quality. This provides accurate 3D models (48–50).

Stochastic dynamic simulations were accomplished using the variant of molecular dynamics that is implemented in MacroModel. The forces from the force field were augmented by frictional and random forces that simulate some properties of a solvent medium (63). The chosen temperatures were 300, 500, and 1000 K, the time step was 1.5 fs for  $T = 300 \text{ K}$  and 1 fs in the other cases, the total simulation time in each case was 2 ns, and 200 snapshots were saved for each run. All saved conformers were fully minimized and ranked by ascending energy (TNCG, 1000 steps).

Minimization and molecular dynamics runs were performed under constraints coming from ROESY experiments. The following distance constraints were chosen according to the NOE correlation intensities: strong,  $2.2 \pm 0.4 \text{ Å}$ ; medium,  $3.5 \pm 0.9 \text{ Å}$ ; and weak,  $5.0 \pm 0.5 \text{ Å}$ . These distance constraints are chosen in order to ensure three contiguous classes of distances without overlap but also without any "forbidden" intervals.

**(B) Peptide-Tannin Systems.** Calculations were performed with and without the intermolecular NMR constraints but

always starting from the previously found peptide conformation, determined using intramolecular NMR constraints.

For calculations performed with intermolecular NMR constraints, the B3 molecule previously built with calculated NMR coupling constants corresponding to the experimental data was used throughout (36) (results to be published elsewhere). Because there are very few modifications in NOE and  $J$  values upon tannin addition, the lowest energy conformer of peptide IB7₁₄ found at 1000 K was used. The intermolecular NOE constraints were added, and the calculation ran for 2 ns.

Calculations were also performed by removing the intermolecular NMR constraints. In this case, after full minimization, two systems were studied with different IB7₁₄/B3 molar ratios: system A with a 1/1 ratio and system B with a 1/4 ratio. GB/SA continuum solvation was systematically used with water as solvent and with the extended cutoff option. Five different 2 ns runs were performed with system A and two different 5 ns runs with system B, with tannins randomly positioned around the peptide. In each case a snapshot was stored every 10 ps.

Following the constrained molecular dynamics runs, a cluster analysis was performed with Xcluster 1.1 (51) as implemented within MacroModel. For this purpose, we used a distance criterion selection as the RMS difference between corresponding all heavy atoms in pairs of structures. This approach leads to a set of clusters, each of them being a family of conformers. In Xcluster, a conformer belongs to a cluster if it lies within the threshold distance of any component of this cluster and at more than this threshold distance of all components of all other clusters.

**(C) Calculation of Amphiphilic Surfaces.** Molecular lipophilicity potentials (MLP) were calculated at the atomic level with an homemade program originating from an idea of Audry et al. (52) and using an exponential function (53). The fragmental atomic constants used were those of Broto et al. (54), and MLP maps were calculated via the MLPP program (55).

## RESULTS

Our systems were analyzed by physicochemical techniques, CD, NMR, MS, and molecular modeling. Results are presented accordingly. All CD, NMR, and mass spectrometry experiments were performed in water and water-ethanol (88:12) at low pH (3.5) and room temperature.

**Circular Dichroism of IB7₁₄ and IB7₁₄-B3.** CD spectra of IB7₁₄ in the absence and presence of tannins (1/0, 1/1, 1/2, 1/3 molar ratios) are shown in the Figure 2. The CD contribution of B3 was removed by recording independently the spectrum of the tannin alone in solution. All curves superimpose except that for the 1/3 ratio, where a slight deviation is observed. However, it must be mentioned that, in the latter case, the sample was very diluted due to the successive addition of B3 solution. So the spectrum has a poorer signal-to-noise ratio that may explain the slight deviation from the other curves. It appears that the addition of B3 does not markedly modify the peptide folding. All spectra were deconvoluted according to a procedure already reported (21), and all showed a combination of 34% of type II helix and 66% of an extended and random coil conformation, within the experimental error.

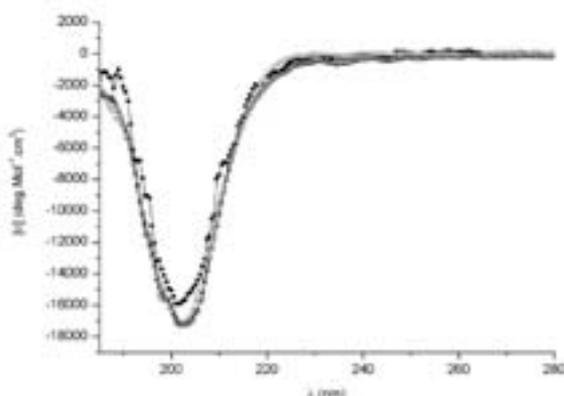


FIGURE 2: Circular dichroism spectra of 200  $\mu\text{M}$  IB7₁₄ in the presence and absence of B3 tannin: (open circle) no tannin; (star, triangle, square) 1/1, 1/2, 1/3 peptide/procyanidin B3 molar ratios, respectively.  $T$  = room temperature. Spectra are normalized according to the peptide concentration. The procyanidin B3 spectral contribution was removed using data obtained from control experiments with B3 alone in solution.

*Nuclear Magnetic Resonance of IB7₁₄ and IB7₁₄-B3.* NMR samples were prepared in water solution. Peptide concentrations ranged from 1 to 5 mM whereas those of B3 varied from 0 to 13 mM. In these conditions, no precipitate could be detected by visual inspection. However, NMR experiments were recorded using the HRMAS technology to take care of the colloidal behavior of these complexes that would otherwise lead to poorly resolved spectra (56).

Samples containing the peptide alone were submitted to the classical battery of 2D techniques for structure determination, as described in Materials and Methods. It must be mentioned that DIPS1 or HOHAHA were used, without noticeable changes in results. ROESY was used instead of

NOESY because of the lack of cross-peak intensities in the latter. No peptide spectral change was observed on the 1–5 mM concentration range. Resonance assignment was performed using the usual procedures, and the chemical shifts, coupling constants, and NOE effects were tabulated (49 NOE and 12  3J ). This information is available as Supporting Information. The IB7₁₄ peptide structure was calculated using the DYANA software and further refined using molecular modeling (*vide infra*).

Experiments were repeated upon addition of procyanidin B3. NMR parameters were recorded for 10 different peptide/procyanidin B3 molar ratios (1/1 to 1/7 molar ratios). Results are also available as Supporting Information. Several comments can be made. Upon tannin addition, one notices gradual changes in some peptide chemical shifts. They will be discussed in the titration section (*vide infra*). No supplementary intramolecular NOE was evidenced in the peptide. Very interestingly, intermolecular NOE were detected between the peptide and the procyanidin dimer (Figure 3 and Figure A in Supporting Information). It must be mentioned that no significant variation was observed for the tannin chemical shifts. Among all of the protons of the peptide, four groups of residues appear to be more influenced by an increasing polyphenol concentration: Pro2, Pro6, Pro9-Pro10, and Gly13-Gly14. Figure 4 displays the Pro9-H $\alpha$  chemical shift variation as a function of the procyanidin/peptide molar ratio. The chemical shift difference values were fitted using eq 1 after the estimation of the procyanidin B3 association constant ( $K_a$ ) using the method previously described by Baxter et al. (45) ( $K_a = 1.8 \text{ M}^{-1}$ ). By working with a procedure that allows  $K_a$ ,  $n$ , and  $\Delta\delta_{\text{max}}$  to vary freely, one finds  $K_a = 1.9 \times 10^{-4} \text{ M}$ ,  $\Delta\delta_{\text{max}} = 0.111$ , and  $n = 3.2$ . The fitting was repeated for peptide protons of Pro2-H $\alpha$ , Pro2-H $\beta$ , Pro6-H $\alpha$ , Pro6-H $\gamma$ , Pro9-H $\alpha$ , and Gly13-NH.

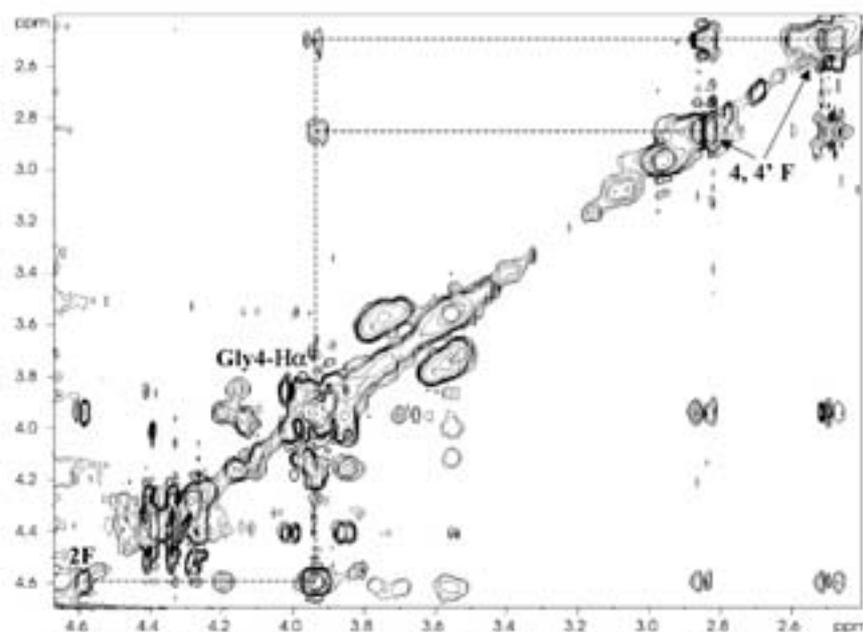


FIGURE 3: Contour plot of the partial ROESY spectrum recorded at 500 MHz, 300 K, and pH 3.5 using a 300 ms spin-lock time of a sample containing 2 mM IB7₁₄ and 6 mM B3 (1/3 IB7₁₄/B3 molar ratio). Expansion is made in the H $\alpha$ -H $\alpha$  region. Only intermolecular NOE effects between Gly4-H $\alpha$  and 2F, 4F, and 4'F are shown.

## Saliva Protein–Tannin Binding

Biochemistry, Vol. 42, No. 35, 2003 10389

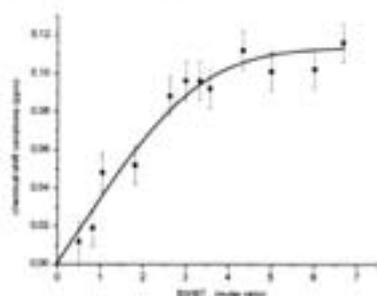


FIGURE 4: Observed (■) and fitted (—) chemical shift differences ( $\Delta\delta$ ) of Pro9-H $\alpha$  as a function of the B3:IB7₁₄ molar ratio.  $\Delta\delta$  is calculated with respect to the chemical shift observed in the absence of tannin. Fitting is performed according to eq 1 of the text. Accuracy is 10% (error bars).

Table 1: Thermodynamic Data for the IB7₁₄-B3 Complex^a

residue	$K_d$ (mM)	$\Delta\delta_{\text{max}}$	$n$	$\chi^2$
Pro2-H $\alpha$	8.15	0.228	3.0	0.028
Pro2-H $\beta$	3.00	0.085	3.5	0.015
Pro6-H $\alpha$	0.28	0.067	2.6	0.003
Pro9-H $\alpha$	0.19	0.111	3.2	0.013
Gly13-NH	2.46	0.186	3.2	0.018

^a  $K_d$ ,  $n$ , and  $\Delta\delta_{\text{max}}$  are obtained from NMR chemical shift variations by a least-squares fit of eq 1.

Results are given in Table 1. It is noticed that  $K_d$  varies by a factor of 50 around the millimolar value whereas  $n$  is about 3. It must be mentioned that the measure of chemical shifts is bound to an intrinsic inaccuracy. Because there is an important overlap of resonances in the regions of interest (particularly for the prolines), one is forced to measure

chemical shifts in the 2D maps, which are much less accurate than in 1D spectra. Accuracy then drops down to 0.01 ppm.

**Mass Spectrometry of IB7₁₄ and IB7₁₄-B3.** Electrospray ionization, which is known to allow the direct observation of supramolecular assemblies in the gas phase (57, 58), has been used under static sample infusion with nanospray needles. While ensuring low sample consumption, this introduction mode greatly facilitates the recording of CID spectra for complexes having various stoichiometries.

When analyzed alone under nanospray ionization, the isolated IB7₁₄ peptide shows a doubly charged ion at  $m/z$ : 650.9 as the major species (Figure 5A; calculated value 650.83 Th). Premixed peptide-tannin solutions in water-ethanol-acetic acid (89:10:1) were infused in the same way, resulting in the observation in the gas phase of noncovalent IB7₁₄-B3 complexes. When infusing a mixture of IB7₁₄-B3 premixed at a 1:4 ratio, doubly charged species corresponding to peptide-tannins complexes with 1:1, 1:2, and 1:3 stoichiometries could be isolated in the ion trap, and CID spectra were recorded (Figure 5B-D). A faint signal possibly corresponding to the 1:4 complex was also observed in the full scan mode, but no CID spectrum could be obtained to confirm its identity. The doubly charged species attributed to the 1:1 IB7₁₄-B3 complex was observed at  $m/z$ : 940.0 (calculated value 939.91 Th). Once fragmented in the ion trap, it leads to the doubly charged species of IB7₁₄ at  $m/z$ : 650.9 (Figure 5B). The signal corresponding to the singly charged species of B3 was not observed here, this tannin being less efficiently ionized than IB7₁₄ under our source conditions. This 1:1 IB7₁₄-B3 complex started to disrupt at 15% collision energy (instrument scale; as an indication, a collision energy higher than 28% is required to fragment the

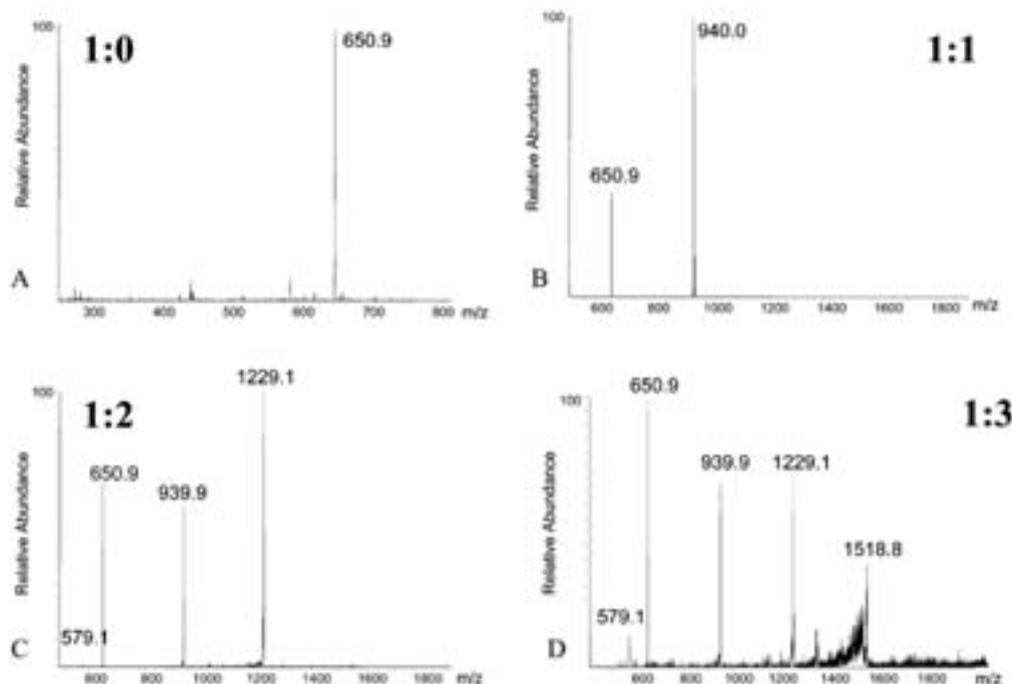


FIGURE 5: ESI mass spectra obtained with a nanospray source for IB7₁₄ alone and complexed with one, two, or three procyanidin B3 molecules. Panels: (A) IB7₁₄ alone; (B) CID spectrum of a 1:1 IB7₁₄:B3 complex; (C) CID spectrum of a 1:2 IB7₁₄:B3 complex; (D) CID spectrum of a 1:3 IB7₁₄:B3 complex.

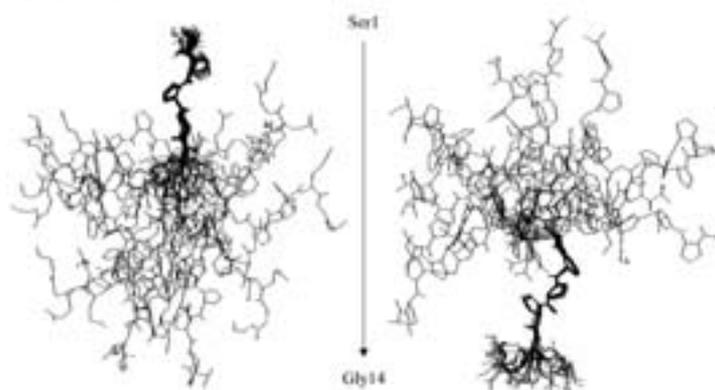


FIGURE 6: Superimposition of the 20 best conformations of IB7₁₄ obtained after a molecular dynamic run (2 ns) under NMR constraints. Two independent superimpositions are shown: Pro2 to Pro6 (RMSD = 0.436) (left); Pro9 to Gln12 (RMSD = 0.348) (right).

doubly charged ion of IB7₁₄ in the ion trap). At 19% collision energy, the 1:1 complex was almost completely disrupted. As shown on Figure 5C, the 1:2 complex was observed at  $m/z$  1229.1 (calculated value 1228.98 Th). It splits into a doubly charged species having a 1:1 IB7₁₄:B3 stoichiometry, while ions corresponding to [IB7₁₄ + 2H]²⁺ and [B3 + H]⁺ are also observed. The behavior of the 1:3 complex under CID conditions is similar, as it splits into 1:2 and 1:1 doubly charged species of these complexes and [IB7₁₄ + 2H]²⁺ and [B3 + H]⁺ species (Figure 5D). However, IB7₁₄-B3 complexes having 1:1 and 1:2 stoichiometries started to dissociate at energy values of 15% and 18%, respectively (instrument scale), whereas the disruption of the 1:3 complex was only observed above a 22% value.

**Molecular Modeling.** The peptide structure was calculated from the constraint files obtained by NMR using the DYANA software. The coordinate file was further refined using MacroModel. We performed a continuous long run (2 ns) using stochastic dynamics at various temperatures with continuum solvation (water was used as solvent). In all cases the set of constraints deriving from ROESY measurements was used throughout. Two hundred snapshots were regularly sampled and fully minimized always under constraints. The temperatures chosen were 300, 500, and 1000 K, respectively. As could be expected, the calculation carried out at 300 K proved to be largely inefficient to sample the conformational space during the 2 ns trajectory run. Interestingly, the two experiments conducted at 500 and 1000 K led to almost identical results: the final number of unique conformers is roughly the same (176 and 180, respectively) and the lowest energy conformers were found at  $-1780.09$  kJ·mol⁻¹ at 500 K and  $-1781.84$  kJ·mol⁻¹ at 1000 K. This indicates that the "melting" at 500 K was sufficient to allow the sampling of the peptide conformational space. The retained solutions are the 20 lowest energy conformers resulting from these two experiments. Figure 6 shows the superimposition of these structures. Superimposition is made by favoring coincidence of residues 2–6 on one hand and residues 9–12 on the other. It can readily be seen that the peptide can be divided into two moieties linked through Gly8, which plays the role of a flexible rotula. These results confirm the extended conformation of the peptide, the type II helix being exclusively represented by the consecutive presence of two (Pro2-Pro3) or three (Pro9-Pro10-Pro11)

prolines. The pitch of the type II helix is found to be about 8 Å. The number of residues in the helix of type II is approximately 7, thus giving an amount of secondary structure of  $50 \pm 10\%$ .

The above procedure was repeated in the case of peptide-tannin complexes. Because there is little variation in the NMR constraints of the peptide in the presence of B3, some of the structures obtained in the absence of B3 were taken together with the B3 structure (36), and the five NOE contacts between the Gly4-Hα and Lys5-NH peptide residues and the tannin 2F, 4F, and 4'F protons were added to the list of constraints. We performed a structure calculation of the peptide under constraints with an IB7₁₄:B3 molar ratio of 1:1 by the procedure already described. The molecular dynamics run was performed at 500 K and after minimization under constraints; only 40 conformers were retained in an energy range of 100 kJ·mol⁻¹. A cluster analysis was then performed leading to five families. Three of them contained only 1 conformer, and the remaining two contained 11 and 26 conformers, respectively. Figure 7A displays the 26 minimized structures of the principal family superimposed from Pro2 to Pro6. Upon comparison of the superimposition obtained with (Figure 7A) or without B3 (Figure 6), it appears that tannin binding reduces the conformational disorder of the peptide.

Because we have only a few intermolecular constraints from NMR, we decided to perform calculations without these contacts, while keeping intramolecular constraints for both the peptide and the tannin. The first set of calculations was performed by considering a 1:1 complex. The initial peptide structure was taken from the above calculations. The tannin was positioned along the peptide structure starting from residue 1 and going down to residue 14. A minimization step was performed to dock the tannin onto the peptide. It ended up with five different positions along the structure. Five different molecular dynamics calculations were thus performed. At the end of three runs, the B3 molecule was found tightly bound to the peptide. Two sites were evidenced, one in the vicinity of residues Pro2-Pro6 and the other close to Gly8-Pro9, Pro11. In two other runs the B3 dimer was ejected from the peptide contact after 1 ns. Of interest is the finding that one of the sites found by molecular dynamics is also found in the NMR data (vide supra).

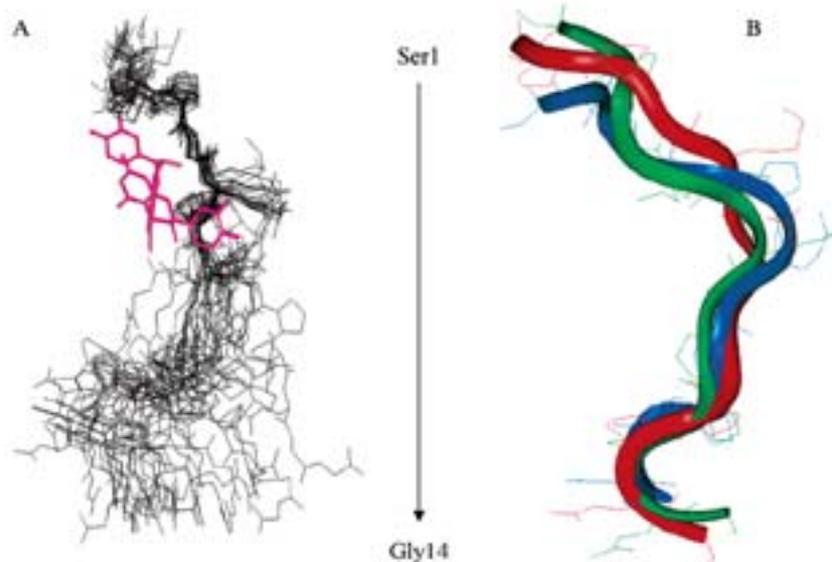


FIGURE 7: (A) Superimposition of the 26 conformations of IB7₁₄/B3 (1/1 molar ratio) obtained after cluster analysis (see text). The molecular dynamics were performed during 2 ns under NMR constraints. For purpose of clarity, only one B3 molecule is shown (purple sticks). (B) Superimposition of IB7₁₄ backbones (ribbon representation) at the end of molecular dynamics without constraints (blue and green) and with NMR constraints (red). Calculations without constraints were performed in the presence of four tannins (not shown).

Because we found a 1:3 peptide:tannin stoichiometry by mass spectrometry and NMR, we performed calculations by considering that more than three tannins could complex the peptide. Molecular dynamics runs were thus performed by considering a system made of one peptide and four tannin molecules. B3 was positioned at both peptide extremities and oriented at random. Minimization was also performed that ended up with two different starting positions. Two independent molecular dynamic runs were performed for 5 ns. During the two calculations, one of the four B3 dimers was always completely ejected after 2 ns. The last three tannins remained tightly bound to the peptide. The locations were found to be quasi identical for the two runs. One procyanidin is linked to the Pro2 and Pro6 residues, the second one to Pro9 and Pro10, and the third to Gly13 and Gly14 (Figure 8). It is noteworthy that tannins are only found on one face of the peptide. Figure 7B displays the superimposition of the peptide conformations at the end of the two dynamics, together with the peptide conformation obtained under NMR constraints. It is remarkable that all structures superimpose very nicely. One must also indicate that, during the two trajectories, the B3 molecules did not exhibit any noticeable conformational change; i.e., they stayed in the more compact conformation (Figure B in Supporting Information), which is the exclusive conformer present in D₂O (26). First, this observation is confirmed by NMR data since no chemical shift change is observed when the peptide:tannin molar ratio increases. Second, two independent molecular dynamics runs were performed during 2 ns in the same conditions as above but on an isolated molecule of B3, starting from two slightly different conformations. They fully confirm that B3 always experienced an identical and very reduced conformational space. As an additional result coming from the calculations, the peptide appears to be stabilized in one conformation probably by hydrogen bonds between the carbonyl of Pro2 and Pro6 and the OH functions of the phenolic part of

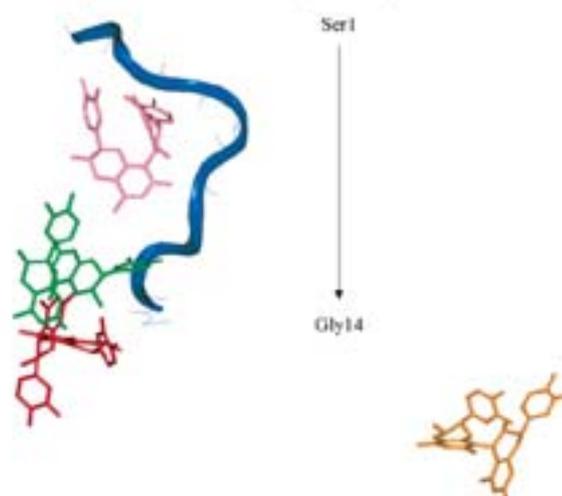
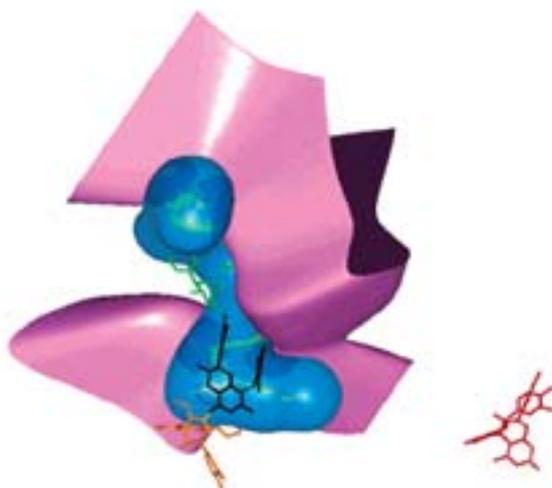


FIGURE 8: Structure of the IB7₁₄/B3 (1/4 molar ratio) complex after completion of a 5 ns molecular dynamics run. The protein is shown as a dark blue ribbon with heavy atoms as thin light gray lines. The three bound tannins are shown as purple, green, and red sticks with H-bonds between B3 molecules and protein. The fourth tannin (yellow) was ejected after 2 ns.

procyanidin. The H-bonding pattern is rather complicated, with a majority of H-bonds between the peptide and the catechol phenolic OH but also with the phenolic OH of the benzopyran moiety. It is also worth noting that one H-bond can be found between a carbonyl group of the peptide and the 3-OH of the pyran cycle.

Because tannins appear to bind to only one side of the peptide, we decided to calculate the peptide surfaces that are hydrophilic and lipophilic at the atomic level, as described in Materials and Methods. The starting structure was that coming from the calculation where three tannins are in



**FIGURE 9:** Molecular lipophilicity potential contours of IB7₁₄. The lowest energy conformation as obtained after a molecular dynamics run (IB7₁₄/B₃, 1/4) has been used. The trace of the protein is shown as a green ribbon. Hydrophilic potentials are shown as a light blue isosurface whereas the lipophilic counterpart is shown as a red isosurface (not visible in this representation). The purple isosurface represents the interface between lipophilic and hydrophilic regions. The four B₃ are shown as sticks; the red one in the right lower part has been ejected after 2 ns.

contact with the peptide. As shown in Figure 9 the peptide exhibits an amphiphilic profile: a hydrophilic part shaded in blue and a lipophilic part shaded in red (not visible in the figure). In Figure 9 the boundary between these two regions is shaded in pink. It is remarkable that the peptide structure is separated in two faces of opposite lipophilicity properties. The most lipophilic part of the peptide includes essentially the alkyl chain of Gln7 and Lys5 and  $\gamma$ - and  $\delta$ -CH₂- of all prolines, the  $\beta$ -CH₂- being always borderline (see Figure C in Supporting Information). It appears that tannins bind to the hydrophilic face that is mainly defined by the peptide backbone (amide, carbonyl, and -CH₂ of glycines and prolines).

## DISCUSSION

This work reports a new and comprehensive approach for studying, in their natural conditions, the interactions between human salivary peptides and condensed tannins of red wines. Using three experimental techniques and molecular modeling, information on 3D structure, dynamics, stoichiometry and stability of the complex, dissociation constants, and minute description of binding phenomenon could be determined. For the first time, magic angle spinning techniques are used to cope with the colloidal nature of complexes and reach structural resolution. Major results are as follows: (i) The IB7₁₄ structure is akin to a helix of type II like polyproline peptides, where glycine 8 seems to play a central role. (ii) Three tannin molecules tightly complex the peptide without modifying its secondary structure, but its conformational disorder appears to be reduced. (iii) Three binding sites are evidenced with a global dissociation constant in the millimolar range. (iv) All tannins bind on the hydrophilic side of the saliva peptide, suggesting that the major interaction forces are governed by hydrogen bonds. All of these points will be

developed below and finally discussed in the general frame of wine astringency.

**Peptide Structures with and without Tannins.** The IB7₁₄ peptide displays an extended conformation due to an almost even repartition of type II helix and random coil. The two portions of helix II are separated by the glycine 8 that appears to play the role of a rotula. The helix of type II structure is more opened than a regular  $\alpha$ -helix: this is demonstrated by measuring the helix pitch. We found  $\sim 8$  Å, which is larger than that of the  $\alpha$ -helix (5–6 Å). Such type II helices are found in collagens (59) and are mainly due to the presence of proline residues. Interestingly, we found almost the same helix II content in the entire IB7 protein (59 aa) (21). It has been proposed by Hagerman and Butler (60) that this opened and flexible structure is a prerequisite for tannin binding.

Adding procyanidin B₃ to IB7₁₄ results in the formation of a complex without affecting the secondary structure of the peptide, as demonstrated by CD, NMR, and molecular dynamics. As observed from molecular dynamics calculation, the two helix II parts on both sides of the Gly8 appear to be retained and fixed by three B₃ molecules, which results in a loss in degrees of freedom of the peptide. Instead of obtaining a cloud of conformations as depicted in Figure 6, only one of them is favored (Figure 7B). When comparing the results obtained from structure calculation under NMR constraints and molecular dynamics without intermolecular constraints, the two structures coming from the pure calculation nicely superimpose with that obtained using intermolecular constraints. On the other hand, the B₃ molecule adopts the more compact form like "tweezers" that offers the maximum of OH group pointing outside, thus favoring H-bonding.

**Peptide–Tannin Complexes.** There is clear evidence that IB7₁₄ and B₃ form a 1:3 complex. First, complexes having 1:1, 1:2, and 1:3 stoichiometries have been observed by means of electrospray mass spectrometry. It must be noticed that even if the ionization conditions are quite mild in the MS source, a lot of energy is given to the complexes. The assemblies are stable enough to stand this energy and travel as noncovalently bound species. Furthermore, these complexes were infused at low pH in the presence of ethanol, conditions that most often disrupt noncovalent assemblies involving proteins. Of interest is the fact that one must increase the voltage applied to the ion trap to dissociate the 1:3 complex, in comparison to the value used for the 1:1 or 1:2 complex. This strongly indicates the increased stability of the 1:3 assembly. Second, NMR titration experiments favor a 1:3 association characterized in a noninvasive manner. It is worth mentioning here that we worked at tannin concentrations that were much lower than the B₃ self-aggregation constant ( $1.8 \text{ M}^{-1}$ ). Fitting of eq 1 indeed leads to a number of binding sites of ca. 3, in all cases. As mentioned in the Results section,  $K_d$  values obtained by proton NMR titration bear some inaccuracy, giving us only 1 order of magnitude around the millimolar value. This is nonetheless in accordance with values ranging between 2 and 33 mM as already published for a similar polyphenol (galloyl epigallocatechin) and a 19-mer peptide (29). The authors also found a stoichiometry of 3 for their protein–tannin system. Although close to our values, one must mention that the above authors always used 10–20% DMSO to cope with aggregation. In this work HRMAS NMR offers

the advantage of working in a natural medium even in the presence of colloidal particles. At this level of the discussion, it is important to note that the model used to fit our data and those of Charlton (33) is not applicable for a cooperative mechanism for B3 fixation (29). Our mass spectrometry data suggest that the binding mechanism may be cooperative; hence, our NMR data could be reinterpreted in the light of a new model, when available. Also of interest is the fact that molecular dynamics of a system initially constituted by one peptide and four tannins ends up in a tight 1:3 complex. So we have very strong experimental evidence for a stoichiometry of 1:3 independently of a model.

*Detailed Structural Analysis of the Complex.* Chemical shift variations of defined peptide protons (H $\alpha$  of Pro2, Pro6, Pro9, Pro10; NH of Gly13-Gly14) upon B3 addition could be assimilated with interaction sites. This highlights the important role of proline residues for the interaction as already described (26, 27, 29, 33, 60). ROESY experiments performed for a 1:3 molar ratio only led to intermolecular NOE effects between B3 and the Pro2-Pro6 site, thus confirming the existence of this site. When the structure of a 1:1 complex obtained by molecular dynamics using intermolecular NMR constraints is compared to those obtained from calculations in the absence of these constraints, two sites are evidenced: one made by the Pro2 and Pro6 residues and the second by Pro9-Pro10-Pro11. Please note that in this calculation the starting position of B3 was varied from one peptide end to the other. Thus, the positioning of the first site is in complete accordance with the ROESY experiments and the second with the chemical shift variations observed for Pro9-Pro10 protons. The molecular dynamics calculations with a 1:4 molar ratio led to the binding of three tannins, the fourth one being expelled from the complex. Again, there is a complete agreement with the NMR data (ROESY plus chemical shift variations), the third site being identified close to Gly13-Gly14. Since we only detect intermolecular NOE effects between tannin and the Pro2, Pro6 site, it can be suggested that the two other sites are much more dynamic than that detected by ROESY experiments. The contact between tannins and peptide does exist, but the exchange could be interpreted as fast (correlation times shorter than nanoseconds), thus precluding detection of correlations by through-space NMR experiments. Also of interest is the fact that the three binding sites that have been experimentally evidenced by NMR and mass spectrometry are only seen by molecular dynamics when enough tannins are placed around the peptide. For the 1/1 molecular dynamic simulation, only two sites were identified. This could suggest that the binding process may be cooperative, which tells us that the model used to determine  $K_d$  (vide supra) is probably not appropriate.

The structure of the complex shows that B3 molecules are fixed always on one side of the peptide; this site has been identified as hydrophilic after our amphiphilicity calculations. This is in agreement with molecular modeling calculations that show the presence of numerous hydrogen bonds between the carbonyl functions of some proline residues and the OH group of the A, D or B, E rings of procyanidins. Conversely, no stacking is observed between the pyrrolidine ring of proline residues and the phenolic rings of procyanidin that would be in favor of a hydrophobic effect. The observation of IB7₁₄-B3 complexes using nanospray

ionization mass spectrometry also highlights interactions occurring through hydrogen bonds. Indeed, supramolecular complexes involving proteins, while losing hydrophobic interactions, most often survive the desolvation-ionization process in the electrospray source because of the occurrence of hydrogen bonds and electrostatic interactions (57, 58, 61). It is also interesting to mention that Hagerman and Butler by performing competitive binding assays in water also described proanthocyanidin-protein interactions as being mainly governed by hydrogen bonding (60).

The nature of the binding has been widely discussed. Haslam (27) described the complexation of polyphenols with proteins as a specific example of molecular recognition, where polyphenols act as multidentate ligands via their aromatic nuclei and phenolic groups, the principal driving force being thought to be the hydrophobic effect. The interaction has been interpreted as an association between the phenol rings and the open, flat, and rigid hydrophobic surface of the pyrrolidine ring of a proline residue (24, 26). These findings appear to be in contradiction with ours. It must, however, be mentioned that their experiments were performed by adding 10–20% DMSO to increase the solubility of their preparations. This may change the hydrophobic/hydrophilic properties of both the tannin and the proteins in this no longer "natural" medium. The delicate balance between the amphiphilic forces may then strongly depend on solvent.

*A Molecular Vision of Astringency.* Protein-tannin interaction is the source of astringency, the mouth dryness sensation due to the complexation of the lubricant salivary proteins. It is often asked why very different feelings are perceived depending on wines. This may be linked to the specific type of interaction between tannins and proteins that may depend on tannins and protein nature (30) and also to the stability and the lifetime of complexes. We have shown herein that the IB7 fragment can complex three B3 tannins and that the interaction occurs in the hydrophilic side of the peptide. The peculiar structure of the peptide can play an important role here. In the absence of tannins it appears very flexible and offers type II helices (proline-rich parts) for binding. Once complexed, the conformational freedom of the peptide, as well as that of the entire complex, appears much reduced. Is this phenomenon related to astringency? It is interesting trying to predict what would be the stoichiometry for B3 interacting with the entire IB7 human saliva protein. Because the IB7₁₄ sequence is almost contained three times in IB7, one could guess that a 1/9 complex could be stabilized. If this happened to be true, PRP proteins of the IB7 family could be called "tannin sponges". Whereas the properties of PRP become unveiled, the specificity of the tannin counterpart is more subtle. It probably depends on its chemical nature (hydrolyzable, condensed, galloylated, etc.). Its molecular weight, 3D structure, propensity to auto-associate, and water solubility may lead to noticeable differences in the protein-tannin complex formation. This explains, perhaps, the gustative differences noticed by enologists about tannin behavior in red wine such as drying, puckering, sour, astringent, bitter, or rough character (62). Of interest is the fact that the tannin binding sites appear to have different physicochemical properties. In our case the sites appear to have different dynamics. This might be linked to their lifetime that could modulate the time during which

the proteins are complexed and therefore no longer available to lubricate the mouth.

## CONCLUSION

The concomitant use of CD, HR-MAS NMR, electrospray ionization mass spectrometry, molecular modeling, and dynamics afforded a comprehensive physicochemical description of PRP-tannin interactions. Information on stoichiometry, affinity constants, binding site localization, nature, and dynamics could be obtained. Of importance, all experiments were accomplished in a medium very close to the natural environment, i.e., the mouth saliva.

## ACKNOWLEDGMENT

We thank Philippe Picard (UBS, Bordeaux) for help in debugging GIFA, XEASY, and DYANA programs and Katell Bathany and Marc Bonneau (University Bordeaux 2) for help in mass spectrometric measurements.

## SUPPORTING INFORMATION AVAILABLE

Four tables giving proton chemical shifts,  3J  scalar coupling constants, and intra- and intermolecular NOE effects of IB7₁₄ and three figures showing a contour plot of partial ROESY spectra, a superimposition of the 4/1 tannin/IB7₁₄ system, and molecular lipophilicity potentials. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- Joshi, M. A., and Goldstein, J. L. (1964) *Arrangement of fruit and fruit products in relation to phenolic content* (Mack, M. E., and Stewart, G. F., Eds.) Academic Press, New York and Chichester.
- Green, B. G. (1993) *Acta Psychol.* 84, 119.
- Prinz, J. F., and Lucas, P. W. (2000) *J. Oral Rehabil.* 27, 991–994.
- Bate-Smith, E. C. (1954) *Food* 23, 124.
- Kallithraka, S., Bakker, J., and Clifford, M. N. (1998) *J. Sens. Stud.* 13, 29–43.
- Haslam, E. (1998) *Practical Polyphenolics*, Cambridge University Press, Cambridge, U.K.
- Garwel, R. (1998) *Ann. J. Grape Wine Res.* 4, 74–95.
- Ribéreau-Gayon, P. (1972) *Plant phenolics*, Oliver and Boyd, Edinburgh.
- Bennick, A., McLaughlin, A. C., Geery, A. A., and Madapallimattam, G. (1981) *J. Biol. Chem.* 256, 4741–4746.
- Kauffman, D. L., and Keller, P. J. (1979) *Arch. Oral Biol.* 24, 249–256.
- Bennick, A. (1982) *Mol. Cell. Biochem.* 45, 83–99.
- Beeley, J. A. (2001) *Oral Dis.* 7, 69–70.
- Gron, P., and Hay, D. I. (1976) *Arch. Oral Biol.* 21, 201–205.
- Hatton, M. N., Loomis, R. E., Levine, M. J., and Tabak, L. A. (1985) *Biochem. J.* 230, 817–820.
- Bergey, E. J., Levine, M. J., Reddy, M. S., Bradway, S. D., and Al-Hashimi, I. (1986) *Biochem. J.* 234, 43–48.
- Mehansho, H., Clements, S., Shewes, B. T., Smith, S., and Carlson, D. M. (1985) *J. Biol. Chem.* 260, 4418–4423.
- Mehansho, H., Ann, D. K., Butler, L. G., Rogler, J., and Carlson, D. M. (1987) *J. Biol. Chem.* 262, 12344–12350.
- Lu, Y., and Bennick, A. (1998) *Arch. Oral Biochem.* 43, 717–728.
- Kauffman, D. L., Keller, P. J., Bennick, A., and Blum, M. (1993) *Crit. Rev. Oral Biol. Med.* 4, 2287–2292.
- Stubbs, M. J., Chan, J., Kwan, A., So, J., Berczynski, U., Rasoulit-Rashti, M., Robinson, R., and Bennick, A. (1998) *Arch. Oral Biol.* 43, 753–770.
- Simon, C., Panet, I., and Dufoure, E. J. (2002) *J. Pept. Sci.* 9, 125–131.
- Breslin, P. A. S., Gilmore, M. M., Beauchamp, G. K., and Green, B. G. (1993) *Chem. Senses* 18, 405–417.
- de Lorgeril, M., Salen, P., Paillard, F., Laporte, F., Boucher, F., and de Leiris, J. (2002) *Circulation. Res.* 54, 503–515.
- Hatano, T., and Hemingway, R. W. (1996) *J. Chem. Soc., Chem. Commun.* 6, 2537.
- Vergé, S., Richard, T., Moreau, S., Narich, A., Merillon, J. M., Vercateren, J., and Mouth, J. P. (2002) *Biochim. Biophys. Acta* 1571, 89–101.
- Hatano, T., Yoshida, T., and Hemingway, R. W. (1999) in *Plant polyphenol 2: Chemistry, Biology, Pharmacology, Ecology* (Gross, G. G., Hemingway, R. W., and Yoshida, T., Eds.) pp 509–526, Plenum Publishers, New York.
- Haslam, E. (1996) *J. Nat. Prod.* 59, 205–215.
- Murray, N. J., Williamson, M. P., Lilley, T. H., and Haslam, E. (1994) *Eur. J. Biochem.* 219, 923–935.
- Charlton, A. J., Baxter, N. J., Khan, M. L., Moir, A. J. G., Haslam, E., Davies, A. P., and Williamson, M. P. (2002) *J. Agric. Food Chem.* 50, 1593–1601.
- De Freitas, V., and Mateus, N. (2001) *J. Agric. Food Chem.* 49, 940–945.
- Charlton, A. J., Baxter, N. J., Lilley, T. H., Haslam, E., McDonald, C. J., and Williamson, M. P. (1996) *FEBS Lett.* 382, 289–292.
- Saru-Manchado, P., Cheymier, V., and Moutouret, M. (1999) *J. Agric. Food Chem.* 47, 42–47.
- Baxter, N. J., Lilley, T. H., Haslam, E., and Williamson, M. P. (1997) *Biochemistry* 36, 5566–5577.
- Hagerman, A. E., and Butler, L. G. (1980) *J. Agric. Food Chem.* 28, 944–947.
- Tuckmantel, W., Kozimowski, A. P., and Romanczyk, J. (1999) *J. Am. Chem. Soc.* 121, 12073–12081.
- Barathias, K. (2002) Thesis, Université Bordeaux 1, Bordeaux.
- Barathias, K., Fouquet, E., Dufour, E. J., Panet, I., Saucier, C., and Glories, Y. (2002) in *XXth International Conference on Polyphenols* (Hadrami, I. E., Ed.) pp 551–552, Marrakech, Morocco.
- Goetz, M., Carlotti, C., Bontemps, F., and Dufour, E. J. (2001) *Biochemistry* 40, 6534–6540.
- Protto, M., Sudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* 2, 661–666.
- Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 113, 967–974.
- Malliaris, T. E., Pons, J. L., and Delsuc, M. A. (1996) *J. Biomol. NMR* 8, 445–452.
- Malliaris, T. E., Pons, J. L., and Delsuc, M. A. (1998) *Biomol. NMR* 14, 624–631.
- Eccles, C., Güntert, P., Billeter, M., and Wüthrich, K. (1991) *J. Biomol. NMR* 1, 111–130.
- Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol.* 273, 283–298.
- Baxter, N. J., Williamson, M. P., Lilley, T. H., and Haslam, E. (1996) *J. Chem. Soc., Faraday Trans.* 92, 231–234.
- Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., and Still, W. C. (1990) *J. Comput. Chem.* 11, 441.
- Still, W. C., Tempczyk, A., Hawley, R. C., and Hendrickson, T. (1990) *J. Am. Chem. Soc.* 112, 6127.
- Vivas, N., Glories, Y., Panet, I., Barbe, B., and Laguerre, M. (1996) *Tetrahedron Lett.* 37, 2015–2018.
- Saucier, C., Guerra, C., Panet, I., Laguerre, M., and Glories, Y. (1997) *Phytochemistry* 46, 229–234.
- Saucier, C., Panet, I., Laguerre, M., and Glories, Y. (1998) *J. Chim. Phys.* 95, 357–365.
- Shenkin, P. S., and McDonald, D. Q. (1994) *J. Comput. Chem.* 15, 899–916.
- Audry, E., Dubost, J. P., Colletier, J. C., and Dallet, P. (1986) *Eur. J. Med. Chem.* 21, 71–72.
- Fauchère, J. L., Quaremon, P., and Kaestner, L. (1988) *J. Mol. Graphics* 6, 203–206.
- Broto, P., Morona, G., and Vandeyke, C. (1984) *Eur. J. Med. Chem.* 19, 71–78.
- Laguerre, M., Saux, M., Dubost, J. P., and Carpy, A. (1997) *Pharm. Sci. J.* 217–222.
- Mirabel, M., Glories, Y., Panet, I., and Dufoure, E. J. (1999) *J. Chim. Phys.* 96, 1629–1634.

## Saliva Protein-Tannin Binding

57. Robinson, C. V., Chung, E. W., Kragefund, B. B., Knudsen, J., Aplin, R. T., Poulsen, F. M., and Dobson, C. M. (1996) *J. Am. Chem. Soc.* *118*, 8646–8653.
58. Loo, J. A. (1997) *Mass Spectrom. Rev.* *16*, 1–23.
59. Ramachandran, G. N., and Ramakrishnan, C. (1976) in *Biochemistry of collagen* (Ramachandran, G. N., and Reddi, A. H., Eds.) pp 45–84, Plenum Press, New York.
60. Hagerman, A. E., and Butler, L. G. (1981) *J. Biol. Chem.* *256*, 4494–4497.
61. Vergé, S., Richard, T., Moreau, S., Richelme-David, S., Vercauteren, J., Promé, J. C., and Mouti, J. P. (2002) *Tetrahedron Lett.* *43*, 2363–2366.
62. Lee, C. B., and Lawless, H. T. (1991) *Chemical Scent.* *16*, 225–238.
63. Van Gunstem, W. F., and Berendsen, M. J. C. (1988) *Mol. Simul.* *1*, 173.

BI034354P

## HIGHLIGHTS OF THE RECENT LITERATURE

## EDITORS' CHOICE

edited by Gilbert Chin

## BIOCHEMISTRY

## A Dry Year

The unusually high summer temperatures in France have raised winemakers' hopes for a truly memorable vintage. The polyphenolic tannins, a component of red wines, come primarily in two varieties:

the flavan-3-ol polymers (made of catechin and epicatechin monomers) from the grape skins and seeds, and the hydrolyzable tannins (made of ellagic and gallic acids) from the oaken barrels. These compounds are responsible for the astringent taste of red wine and are thought to bind to proline-rich proteins (PRPs) in saliva. Previous attempts to define this interaction have been hampered by the limited



Sequestering tannins (green, black, and orange) on the hydrophilic side of IB7 (blue surface).

solubility of the complexes.

Working in the remarkably true-to-life solvent of water:ethanol (88:12, v/v; pH 3.5), Simon et al. describe the association of three procyanidin B3 (catechin-4a,8-catechin) molecules with a 14-residue fragment of salivary protein IB7. Using circular

dichroism, mass spectrometry, NMR, and molecular dynamics calculations, they find that the peptide adopts a type II helical structure typical of proline-rich sequences. Furthermore, the B3 tannins bind via hydrogen bonds with only millimolar affinity to the hydrophilic side of the peptide, reducing its conformational flexibility. The potential range of binding stoichiometries and dynamics in a mixture of tannins and

PRPs reminds us how infinitely varied wines can be. — GJC

*Biochemistry* 10.1021/BI034254g (2003)

## GEOCHEMISTRY

## A History of Grass

Grasslands are thought to have spread widely during the late Cenozoic, starting about 10 to 15 million years ago, and this expansion has been suggested to have had major effects on ecology and evolution: a radiation of ruminants and decline of other grazers, and eventually the rise of humans. Grasses use a different photosynthetic process ( $C_4$ ) than trees and shrubs ( $C_3$ ); the  $C_4$  pathway is favored at lower atmospheric  $CO_2$  levels and cooler temperatures, conditions that became prevalent later in the Cenozoic. Assessing the expansion of grasses and whether there was an abrupt change in ecosystems is critical for testing these proposed links.

Because the  $C_4$  and  $C_3$  pathways produce different stable carbon isotope signatures in

plants, a fingerprint of the abundance of grasses can be recovered from the geologic record. Fox and Koch examined paleosoils from the Great Plains in the United States (now mostly grassland) spanning the period since about 23 million years ago. Their data imply that grasslands were about 20% of the plant biomass until about 5 million years ago, and then expanded to modern levels by about 2.5 million years ago. This record differs from those suggesting an earlier expansion, and implies a local, not global, climate control of grassland extent. — BH

*Geology* 31, 809 (2003)

## BIOPHYSICS

## A Gentle Tug

Designing a protein is such a complicated undertaking in terms of satisfying multiple operating constraints that it seems miraculous that anything works at all. The order of amino acids must allow the spontaneous formation of secondary and tertiary structure during protein synthesis, and the global fold must be flexible enough in enzymes to support catalytic movements and sturdy enough in structural components to resist deformation. Finally, it has to be feasible to tear down these constructs in order to recycle damaged molecules.

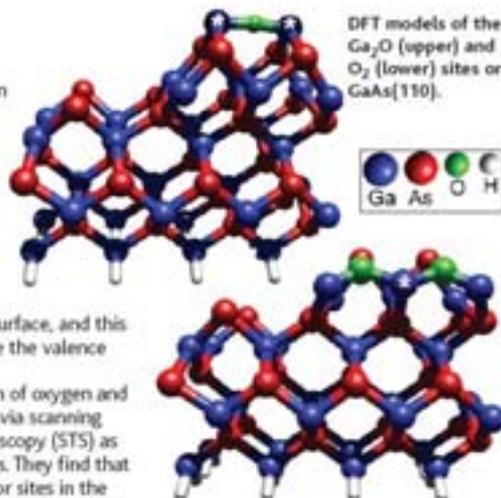
Brockwell et al. pull on the N terminus of a lipoyl domain (from a subunit of pyruvate dehydrogenase) while holding onto either the nearby C terminus or the lipoyl acid, which is covalently attached to a lysine midway between the termini. Carrion-Vasquez et al. pull on polyubiquitin in which the monomers are linked either by an amide bond between the N and C termini or between an interior lysine and the C terminus. In both cases, yanking on the middle leads

## SURFACE SCIENCE

## Unpinning Gallium Arsenide

The metal oxide semiconductor field-effect transistor (MOSFET) is made from silicon rather than GaAs, even though GaAs has a direct band gap and high carrier mobilities. The reason is that the insulating oxide layer that forms on the GaAs surface has a high density of states that "pin" the valence and conduction bands, rather than letting them move with respect to the Fermi level. The classical explanation is that oxidation liberates As atoms from the surface, and this "antisite" indirectly creates states above the valence band maximum.

Hale et al. have studied the adsorption of oxygen and  $Ga_2O$  onto the GaAs(110)- $c(2 \times 8)/(2 \times 4)$  via scanning tunneling microscopy (STM) and spectroscopy (STS) as well as via density functional calculations. They find that oxygen in fact creates donor and acceptor sites in the band gap directly from the Ga atom to which they are bonded. However,  $Ga_2O$  bonds by inserting into As dimer pairs to form a  $c(2 \times 2)$  unpinned surface. Chemisorption of  $Ga_2O$  appears to restore surface Ga and As atoms to nearly their bulk charge states, whereas  $O_2$  creates charged Ga atoms that lead to pinning. — PDS



DFT models of the  $Ga_2O$  (upper) and  $O_2$  (lower) sites on GaAs(110).

*J. Chem. Phys.* 10.1063/1.1601596

CONTINUED ON PAGE 1293

## **Résumé :**

La complexation des tanins avec les protéines de la salive est à l'origine de l'astringence, sensation tactile propre aux vins rouges.

La protéine IB7, son peptide dérivé IB7₁₄ et le tanin B3 ont été synthétisés. Leur mode d'interaction a été étudié par RMN des liquides et des solides (rotation à l'angle magique), spectrométrie de masse, dichroïsme circulaire et dynamique moléculaire. Le peptide IB7₁₄ et la protéine entière sont constitués d'hélice de type II (1/3) et de pelote statistique (2/3); la glycine centrale du peptide joue le rôle de charnière de part et d'autre de régions en hélice. Le tanin, quant à lui, revêt une forme compacte majoritaire en milieu aqueux. L'ensemble des paramètres physico-chimique ont été mesurés: la stœchiométrie du complexe est de 3 molécules B3 par peptide IB7₁₄; la constante de dissociation est de l'ordre du mM; la nature des liaisons assurant la cohésion du complexe est de type liaisons hydrogène; les sites de fixation ainsi que la taille du complexe ont été déterminés. Ces informations permettraient de comprendre l'astringence d'un point de vue physico-chimique.

Mots clés : Protéines de la salive humaine, tanins, structure, interaction, RMN.

## **Abstract:**

The complexation of tannins with salivary proteins is the source of astringency phenomenon, red wine tactile sensation.

IB7 protein, protein fragment IB7₁₄ and B3 tannin were synthesized. The interaction between B3 tannin and protein were monitored by ¹H magic angle spinning NMR, mass spectrometry, circular dichroism and molecular dynamic. Peptide and protein are constituted by type II helix (1/3) and random coil conformation (2/3); the central glycine appears to act as a flexible rotula separating two helices II regions. The tannin adopt a "tweezer" conformation. All the physico-chemical parameters were measured: the complex stoichiometry is of 3 tannins molecules per peptide IB7₁₄; the dissociation constant is of millimolar range; the principal driving forces toward association are governed by hydrogen bonding; the nature of the binding sites and the size of the complex were determined. These findings allow us to understand the physico-chemistry of astringency.

Key words : Human salivary proteins, tannins, structure, interaction, NMR.