

**Université de Reims Champagne-Ardenne**

Ecole Doctorale Sciences Technologies Santé

Spécialité : Biochimie et biologie structurale

*Biochimie et biologie structurale  
appliquées à l'œnologie :  
Etude des protéines de raisin  
Thaumatococcus-like et Chitinase*

Thèse présentée pour obtenir le grade de Docteur  
de l'Université de Reims Champagne-Ardenne

par

***Doriane Le Bourse***

*Soutenue le 9 Décembre 2011 devant le jury composé de :*

Dr. Régis Gougeon	Université de Bourgogne	Rapporteur
Pr. Vinh Tran	Université de Nantes	Rapporteur
Pr. Manuel Dauchez	Université de Reims Champagne-Ardenne	Examineur
Dr. Marc-André Delsuc	Université de Strasbourg	Examineur
Dr. Jean-Marc Nuzillard	Université de Reims Champagne-Ardenne	Examineur
Pr. Christophe Clément	Université de Reims Champagne-Ardenne	Invité
Pr. Philippe Jeandet	Université de Reims Champagne-Ardenne	Directeur de thèse



**Université de Reims Champagne-Ardenne**

Ecole Doctorale Sciences Technologies Santé

Spécialité : Biochimie et biologie structurale

*Biochimie et biologie structurale  
appliquées à l'œnologie :  
Etude des protéines de raisin  
Thaumatococcus-like et Chitinase*

Thèse présentée pour obtenir le grade de Docteur  
de l'Université de Reims Champagne-Ardenne

par

***Doriane Le Bourse***

*Soutenue le 9 Décembre 2011 devant le jury composé de :*

Dr. Régis Gougeon	Université de Bourgogne	Rapporteur
Pr. Vinh Tran	Université de Nantes	Rapporteur
Pr. Manuel Dauchez	Université de Reims Champagne-Ardenne	Examineur
Dr. Marc-André Delsuc	Université de Strasbourg	Examineur
Dr. Jean-Marc Nuzillard	Université de Reims Champagne-Ardenne	Examineur
Pr. Christophe Clément	Université de Reims Champagne-Ardenne	Invité
Pr. Philippe Jeandet	Université de Reims Champagne-Ardenne	Directeur de thèse



*I think that's what it's like with all our dreams and our nightmares,  
we've got to keep feeding them for them to stay alive.*

*John F. Nash, prix Nobel d'économie 1994  
Extrait du film "A beautiful mind"*



# Remerciements

J'ai le souvenir d'avoir commencé ma thèse avec l'envie de changer le monde et la volonté de faire avancer la science. Trois ans plus tard, à l'heure du bilan de mes tribulations et de mes rencontres au gré des bulles de Champagne, force est de constater que ce qui a le plus évolué, c'est ma *vision* du monde et de la science. Le temps est donc venu de remercier comme il se doit toutes les personnes qui ont contribué de près ou de loin à faire de ce travail ce qu'il est et de moi une personne grandie.

Je remercie en premier lieu le Docteur Régis Gougeon et le Professeur Vinh Tran pour l'honneur qu'ils me font en jugeant ce travail. Je remercie également les Docteurs Marc-André Delsuc et Jean-Marc Nuzillard, ainsi que les Professeurs Christophe Clément et Manuel Dauchez d'avoir accepté de participer au jury.

J'exprime ma gratitude envers le Professeur Philippe Jeandet, qui a supervisé ma thèse en m'accordant toute sa confiance et qui m'a laissé la liberté de mettre en place et développer les différents projets qui m'ont menée au-delà des frontières de l'œnologie. Je remercie le Docteur Sandrine Jégou, à qui je dois d'avoir été recrutée. J'adresse toute ma reconnaissance au Professeur Christophe Clément, qui m'a parfois apporté un soutien précieux.

J'ai une pensée pour les membres du Laboratoire d'Oenologie et Chimie Appliquée et de l'Unité de Recherche Vigne et Vins de Champagne. Bonne continuation.

J'ai eu la chance de faire deux rencontres formidables dès le début de ma thèse : Alexandra Conreux et Sandra Villaume. Vous avez toujours été à mes côtés, en recadrant et en canalisant mes idées souvent fougueuses (pour ne pas dire sautillantes), et en m'apportant une aide considérable sur les manip purif/quantif. Votre soutien s'est poursuivi dans la période post-œno, ce dont je vous suis extrêmement reconnaissante. Alex, ton investissement dans ma thèse et ton amitié sont les deux points qui m'ont permis de surmonter les difficultés scientifiques et les épreuves morales tout au long de ces trois ans... Sandra, tu m'as appris à prendre un recul indispensable lors de chaque petit tracas et tu as été la garante de ma survie physique : manger et dormir, c'est la clé... Pour cela et pour tout le reste, je vous remercie infiniment !

J'en arrive au « bâtiment 18 », qui m'a permis de passer les meilleurs moments de ma thèse que ce soit au fin fond de la passerelle du 2<sup>ème</sup> étage – avec ou sans bureau – ou dans l'aile des chimistes.

Mes plus profonds remerciements vont ainsi aux membres du laboratoire ICMR. J'exprime ma plus sincère reconnaissance au Docteur Jean-Marc Nuzillard, pour avoir montré un réel enthousiasme vis-à-vis de mes projets dès leur genèse et pour m'avoir toujours encouragée... votre humanité et votre passion pour la recherche sont des modèles à suivre ! J'ai eu également la chance et l'immense plaisir de travailler avec le Docteur Pedro Lameiras... merci, merci, merci de t'être autant investi dans mes projets et de m'avoir fait partager tes compétences en RMN (surtout la nuit et le weekend !). J'espère sincèrement qu'on aura l'occasion de collaborer à nouveau ensemble. Je remercie également tous ceux que j'ai pu rencontrer là-bas : Bertrand, toujours disponible pour discuter de tout et de rien ; Nicolas, toujours prêt à donner un coup de main ; Arnaud, toujours pétillant ; et tous les autres...

Last but not least, je remercie très chaleureusement le Professeur Manuel Dauchez du laboratoire SiRMa, qui m'a fait plonger avec entrain dans le monde merveilleux de la bioinformatique. En plus de votre soutien scientifique, j'ai grandement apprécié votre soutien moral et vos conseils très avisés tout au long de ma thèse... le tout dans la joie et la bonne humeur ! Merci également au Docteur Nicolas Belloy pour son aide lors de la mise en place des simulations de dynamique moléculaire.

Merci à Patricia et Aziz du laboratoire PPDD, avec qui j'ai passé de très agréables moments autour d'un café sans café.

Merci aux membres du laboratoire GEGENA<sup>2</sup> qui m'ont grand ouvert la porte de leur micro-ondes pendant toute ma dernière année de thèse. Vous m'avez accueillie chaleureusement en m'offrant des moments de détente et de convivialité qui ont grandement égayé mon quotidien ! En particulier je tiens à remercier Béatrice, dont le soutien, l'écoute et la compréhension ont souvent été salutaires... je t'en suis profondément reconnaissante. Une spéciale dédicace à Damien, dont les sous-vêtements sont une éternelle source de débats surréalistes.

Merci à Sandrine pour ton amitié précieuse... Que de chemin parcouru depuis les années prépa, mais on est toujours les mêmes ! Merci à Sébastien l'expatrié, my best supportive friend. Merci également à Johann et Rémy pour m'avoir fait découvrir le birlou.

Merci à Matthieu, fidèle acolyte depuis l'école, pour avoir admiré ma vue sur la cathédrale (oui, oui, je t'ai piqué l'idée) et pour nos – interminables – conversations à refaire la biologie, la recherche académique et les relations humaines... ainsi que pour notre collaboration très productive cher Docteur Tanty ! Un immense merci également à l'inséparable G de D&G... j'espère qu'on aura à nouveau l'occasion de boire des mangolicious sous les cocotiers, aloha ! L'époque formidable de l'A7 team et du bureau-dont-on-ne-doit-pas-prononcer-le-nom (prez' un jour, prez' toujours) se résume par un simple « c'était mieux avant » tout à fait de circonstance !

Merci à Sarah qui me rappelle le bon vieux temps, à Grégoire qui a sauvé la recherche, à Kelly et Kevin qui resteront à jamais sur un bateau, à Tiphaine&Co qui m'ont apporté un soutien nutritionnel inimitable.

Pour terminer, que mes parents trouvent ici le témoignage de ma plus profonde reconnaissance pour leur soutien sans plafond et pour leurs encouragements sur le chemin de la performance. A ma mère, n'oublions pas que tu as posé le premier jalon de ma vie étudiante à l'époque du CP. A mon père, puissais-je te ressembler un peu plus chaque jour.

Ainsi se clôturent ces trois années de thèse. Que de grands moments partagés, que de projets réalisés, que de désillusions surmontées... A tous ceux qui ont fait de ma vie sociale et scientifique une aventure ubuesque mais formatrice et à tous ceux grâce à qui j'ai survécu... *Merci !*



# Résumé

Les protéines de raisin thaumatin-like et chitinase sont d'un intérêt majeur, tant pour la recherche viticole de par leur implication dans les mécanismes de défense de la vigne, que pour la recherche en œnologie en raison de leur présence en quantité majoritaire dans le pool protéique d'un jus de raisin. La mise au point d'un protocole de quantification par chromatographie liquide et dosage immuno-enzymatique utilisant des gammes de protéines purifiées à partir de jus de raisin a permis de caractériser la diminution de la concentration de ces deux protéines au cours du procédé de vinification. Les propriétés tensioactives des deux protéines thaumatin-like et chitinase purifiées ont été évaluées, révélant que ni l'une ni l'autre ne pouvait à elle seule expliquer la formation et la stabilisation des bulles et de la mousse d'un Champagne. Une étude structurale de la protéine thaumatin-like VVTL1 a ensuite été menée dans l'optique de mieux comprendre les mécanismes chimiques, biologiques et physiques dans lesquels elle peut être impliquée. Une structure de VVTL1 a été modélisée par homologie et l'analyse de ses modes normaux a permis de révéler un mécanisme de type mâchoire autour d'une cavité acide, site potentiel de l'activité enzymatique de la protéine. Un feuillet  $\beta$  en épingle à cheveux isolé dans la structure s'est révélé être très conservé et absolument spécifique à la superfamille des protéines thaumatin-like, ouvrant peut-être la voie vers l'élucidation complète du rôle biologique de ces protéines. Dans une seconde approche, la détermination de la structure d'un peptide sélectionné dans la séquence de VVTL1 par modélisation sous contraintes RMN a posé les bases d'une étude modèle de l'adsorption des protéines à la surface de la bentonite, argile utilisée pour la clarification des vins.

# Abstract

Grape proteins thaumatin-like and chitinase are of major interest, as much by the vine defense mechanisms they are involved in as by their dominance over the grape juice protein pool. Liquid chromatography and immunoassays allowed both proteins to be quantified in grape juice and Champagne, showing that their concentration decreases through winemaking. The involvement of these proteins in gas/liquid interfaces was also studied on the purified fractions from grape juice previously made for quantification standards. Results clearly indicated that neither thaumatin-like nor chitinase could alone explain bubble formation and foam stabilization in Champagne. A first study of the three-dimensional structure of the main thaumatin-like protein VVTL1 using homology molecular modeling was then achieved and normal modes analysis was performed on the VVTL1 model. It revealed a jaw-like mechanism opening and closing an acidic cleft assessed to be the enzymatic binding site. An isolated  $\beta$ -hairpin turned out to be highly conserved and specific to the thaumatin-like superfamily. This domain could provide a first clue to unravel the mystery of the protein biological activity in the field of plant-pathogen interactions. A second approach was set up for the structure determination of a VVTL1 peptide using molecular modeling under NMR restraints. It could lead to a model study of protein adsorption on bentonite, a clay used for wine clarification.



# Valorisation des travaux

## Publications dans des revues internationales

**Le Bourse D.**, Jégou S., Conreux A., Villaume S. et Jeandet P. Review of preparative and analytical procedures for the study of proteins in grape juice and wine. *Analytica Chimica Acta*, 2010, 667, 33-42.

**Le Bourse D.**, Conreux A., Villaume S. Lameiras P., Nuzillard J.M. et Jeandet P. Quantification of chitinase and thaumatin-like proteins in grape juices and wines. *Analytical and Bioanalytical Chemistry*, 2011, 401, 1545-1553.

**Le Bourse D.**, Tanty M., Belloy N., Conreux A., Villaume S., Lameiras P., Nuzillard J.M., Delsuc M.A., Jeandet P. et Dauchez M. Insights into structure and function of a pathogenesis-related thaumatin-like protein from grape. *Proteins: Structure, Function, and Bioinformatics*, 2011, soumis pour publication.

## Communications orales

**Le Bourse D.**, Conreux A., Villaume S., Lameiras P., Nuzillard J.M. et Jeandet P. Purification and quantification of chitinase and thaumatin-like proteins in grape juice and wine by liquid chromatography. *12th International Meeting and Workshop of the Society for Biochromatography and Nanoseparations*, Lyon, 2010.

Jeandet P., Cilindre C., Marchal R., Jégou S., **Le Bourse D.**, Younes B., Parmentier M., Vasserot Y., Lameiras P., Nuzillard J.M., Dauchez M., Gougeon R., Schmitt-Kopplin P., Polidori G. et Liger-Belair G. Focus sur la physique des bulles et la chimie de la mousse du champagne. *La Vigne: de la Molécule au Paysage*, ENSAIA, Nancy, 2011.

## Posters

**Le Bourse D.**, Jégou S. et Jeandet P. Analytical techniques for the study of proteins and peptides in grapes, musts and wines: A review. *6th International Symposium In Vino Analytica Scientia*, Angers, 2009.

**Le Bourse D.**, Conreux A., Villaume S., Lameiras P., Nuzillard J.M. et Jeandet P. Quantification of chitinase and thaumatin-like pathogenesis-related proteins in grape juice and wine by high performance liquid chromatography. *3rd International Symposium on Macromolecules and Secondary Metabolites of Grapevine and Wine*, Turin, Italie, 2010.

**Le Bourse D.**, Belloy N., Jeandet P. et Dauchez M. Three-dimensional structural behavior of pathogenesis-related thaumatin-like protein from grapes : Toward a better understanding of its biological activity and its properties in wine. *7th International Symposium In Vino Analytica Scientia*, Graz, Autriche, 2011.

**Le Bourse D.**, Lameiras P., Tanty M., Gougeon R., Conreux A., Delsuc M.A., Jeandet P. et Nuzillard J.M. Détermination de la structure d'un peptide par modélisation sous contraintes RMN. Journée d'Inauguration RMN 600 MHz de l'Université de Reims Champagne-Ardenne, Reims, 2011.



# Abréviations

<b>CNS</b>	Crystallography and NMR System
<b>COSY</b>	COrrrelation SpectroscopY
<b>CSI</b>	Chemical Shift Index
<b>DC</b>	Dichroïsme Circulaire
<b>ELISA</b>	Enzyme-Linked ImmunoSorbent Assay
<b>FPLC</b>	Fast Protein Liquid Chromatography
<b>HIC</b>	Hydrophobic Interaction Chromatography
<b>HMBC</b>	Heteronuclear Multiple Bond Correlation spectroscopy
<b>HPLC</b>	High Performance Liquid Chromatography
<b>HSQC</b>	Heteronuclear Single Quantum Correlation spectroscopy
<b>NOE</b>	Nuclear Overhauser Effect
<b>NOESY</b>	Nuclear Overhauser Enhancement SpectroscopY
<b>PR</b>	Pathogenesis-Related
<b>RMN</b>	Résonance Magnétique Nucléaire
<b>RMSD</b>	Root Mean Square Deviation
<b>RP</b>	Reversed-Phase
<b>SDS-PAGE</b>	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
<b>TFE</b>	TriFluoroEthanol
<b>TL</b>	Thaumatococcus-like
<b>TOCSY</b>	TOtal Correlated SpectroscopY
<b>UHPLC</b>	Ultra-HPLC
<b>VVTL1</b>	<i>Vitis Vinifera</i> Thaumatococcus-like 1



# Sommaire

# Sommaire

---

<b>AVANT-PROPOS .....</b>	<b>2</b>
<b>PARTIE I : ETUDE DES PROTEINES DE RAISIN THAUMATIN-LIKE ET CHITINASE.....</b>	<b>5</b>
1 INTRODUCTION.....	7
1.1 <i>Les protéines dans les jus de raisin et les vins de Champagne</i> .....	7
1.1.1 Pool protéique.....	7
1.1.2 Protéines PR .....	8
1.1.3 Propriétés des protéines .....	8
1.2 <i>Techniques préparatives et analytiques pour l'étude des protéines en œnologie</i> .....	8
2 MATERIELS ET METHODES.....	19
2.1 <i>Chromatographie liquide</i> .....	19
2.2 <i>Electrophorèse des protéines</i> .....	24
2.2.1 Principe de l'électrophorèse SDS-PAGE.....	24
2.2.2 Détection des bandes.....	25
2.3 <i>RMN</i> .....	26
2.4 <i>ELISA</i> .....	26
2.5 <i>Tensiométrie à lame de Wilhelmy</i> .....	27
2.5.1 Notion de pression de surface.....	27
2.5.2 Protocole expérimental.....	28
3 RESULTATS ET DISCUSSION .....	29
3.1 <i>Purification et quantification des protéines thaumatin-like et chitinase du jus de raisin au Champagne</i> .....	29
3.2 <i>Propriétés tensioactives des protéines thaumatin-like et chitinase</i> .....	41
3.3 <i>Conclusion</i> .....	42
<b>PARTIE II : ETUDE DE LA STRUCTURE D'UNE PROTEINE THAUMATIN-LIKE ET D'UN PEPTIDE .....</b>	<b>43</b>
1 INTRODUCTION.....	45
1.1 <i>Structure des protéines</i> .....	45
1.1.1 Structure primaire : chaîne d'acides aminés .....	45
1.1.2 Structure secondaire : repliement local .....	46
1.1.3 Structure tertiaire : agencement tridimensionnel.....	49
1.1.4 Structure quaternaire : assemblage de plusieurs chaînes .....	49
1.2 <i>Superfamille des protéines thaumatin-like : « such a sweet world »</i> .....	50
1.2.1 Structure.....	50
1.2.2 Fonction.....	51
2 MATERIELS ET METHODES.....	53
2.1 <i>Etude de la structure de la protéine VVTL1</i> .....	53
2.1.1 Structure primaire : séquence déposée .....	53
2.1.2 Structure secondaire : dichroïsme circulaire.....	54
2.1.3 Structure tertiaire : méthodes <i>in silico</i> .....	55



# Sommaire

2.2	<i>Etude d'un peptide par modélisation moléculaire sous contraintes RMN</i>	57
2.2.1	Sélection et synthèse du peptide	57
2.2.2	Modélisation <i>de novo</i>	57
2.2.3	Dichroïsme circulaire	58
2.2.4	Expériences RMN	59
2.2.5	Modélisation moléculaire sous contraintes RMN	69
3	RESULTATS ET DISCUSSION	72
3.1	<i>Détermination et comportement dynamique de la structure de VVTL1</i>	72
3.2	<i>Détermination de la structure d'un peptide issu de VVTL1 par RMN</i>	86
3.2.1	Choix du solvant	86
3.2.2	Attribution des déplacements chimiques	87
3.2.3	Indices de déplacements chimiques	88
3.2.4	Détermination des contraintes structurales	89
3.2.5	Modélisation moléculaire sous contraintes RMN	90
3.2.6	Poursuite de l'étude	92
3.3	<i>Conclusion</i>	93
	<b>CONCLUSION GENERALE ET PERSPECTIVES</b>	<b>95</b>
	<b>ANNEXES</b>	<b>99</b>
1	DONNEES SUPPLEMENTAIRES RELATIVES A L'ARTICLE « QUANTIFICATION OF CHITINASE AND THAUMATIN-LIKE PROTEINS IN GRAPE JUICES AND WINES »	101
2	MODELISATION ET SIMULATIONS MOLECULAIRES POUR L'ETUDE DE LA STRUCTURE DE PROTEINES ET DE PEPTIDES	103
2.1	<i>Principe général de la mécanique moléculaire</i>	103
2.1.1	Champ de forces	103
2.1.2	Minimisation de l'énergie	104
2.1.3	Dynamique moléculaire	105
2.2	<i>Modélisation d'une structure</i>	106
2.2.1	Modélisation sous contraintes RMN	106
2.2.2	Modélisation par homologie	110
2.2.3	Modélisation <i>de novo</i>	111
2.3	<i>Evolution dynamique d'une structure dans une boîte d'eau</i>	113
2.3.1	Traitement de l'eau	114
2.3.2	Protocole de simulation	114
2.4	<i>Analyse des modes normaux</i>	115
3	FICHIERS DE CONTRAINTES RMN	117
3.1	<i>Contraintes de distances</i>	117
3.2	<i>Contraintes d'angles dièdres</i>	119
	<b>BIBLIOGRAPHIE</b>	<b>121</b>
	<b>LE MOT DE LA FIN</b>	<b>133</b>



# Avant-propos

*Rêve de grandes choses : cela te permettra d'en faire au moins de toutes petites.*

*Jules Renard, Journal (1887-1910)*



Un délicat *plop*, une mousse pétillante qui s'envole puis qui s'efface en ouvrant la voie vers l'ivresse des sens... le Champagne est une promesse de célébration et d'instant magiques ! Pendant que certains trinquent – avec modération –, d'autres se plongent dans les insoupçonnables mystères qu'il renferme et s'étourdissent de l'incroyable et passionnant objet d'étude qu'il représente.

Pour un consommateur, le Champagne se définit comme un vin effervescent – à servir frais. Ces deux aspects, vin et effervescence, sous-tendent la recherche en Œnologie à Reims : la composition du jus de raisin et son évolution au cours des procédés de fermentation d'une part, les phénomènes de bulles et de mousse d'autre part. Cette démarche s'inscrit dans l'idée que la compréhension des mécanismes liés à la culture de la vigne, à la vinification et à l'effervescence conditionne l'amélioration des techniques de culture et de production. C'est la qualité finale du Champagne qui en dépend !

Un Champagne est composé principalement d'eau, de dioxyde de carbone, de petites molécules organiques comme l'éthanol et les arômes et de macromolécules, terme regroupant les polysaccharides et les protéines. Ces protéines proviennent du raisin et des micro-organismes tels que les champignons, levures ou bactéries associés à la culture des baies ou à la fermentation du jus. Elles sont impliquées à tous les niveaux autour du Champagne : de la protection de la vigne en cas d'attaque par un agent pathogène pour les protéines de défense, au désagréable trouble des vins lors de leur agrégation dans une bouteille en passant par leur rôle potentiel dans la formation et la stabilisation des bulles en vertu de leurs propriétés tensioactives. Vaste sujet d'étude donc, auquel nous avons tenté d'apporter une contribution centrée autour des deux protéines de raisin d'intérêt majeur thaumatin-like et chitinase.

Le choix de ces deux protéines a découlé conjointement de leur présence en quantité majoritaire dans le pool protéique d'un jus de raisin, de leur implication dans les mécanismes de défense de la vigne et de l'absence de leur caractérisation de manière individuelle jusqu'alors.

Avant d'imaginer étudier les propriétés biologiques et physico-chimiques des protéines, il est indispensable d'avoir accès à leur concentration dans un milieu donné. Le premier objectif de la thèse a donc été de proposer un suivi de la quantité de thaumatin-like et chitinase du jus de raisin au Champagne par chromatographie liquide et dosage immuno-enzymatique. Les deux méthodes nécessitant préalablement l'obtention de protéines pures pour la réalisation de gammes standards, le premier travail a, en réalité, consisté en la purification à partir de jus de raisin de chacune des protéines. Ce projet a conduit à la mise en place de la première collaboration avec le laboratoire ICMR (Institut de Chimie Moléculaire de Reims) de l'URCA (Université de Reims Champagne-Ardenne), afin de détecter la présence de contaminants organiques dans les fractions de protéines purifiées. Dans un second temps, les protéines pures ont pu être utilisées afin d'obtenir des données expérimentales sur leur potentiel tensioactif et évaluer ainsi si elles pouvaient être à elles seules responsables de la formation d'une interface gaz/liquide.

A l'issue de ce projet axé biochimie, un constat s'est rapidement imposé : pour aller plus loin dans l'étude de thaumatin-like et chitinase, la biologie structurale était un aspect indispensable pouvant apporter des réponses aussi bien sur l'arrangement des protéines à l'interface gaz/liquide, voire liquide/solide, que sur leur mode d'action en tant qu'outil de défense de la vigne. De là est né un nouveau projet, faisant l'objet de la seconde partie du manuscrit. L'étude a été focalisée sur la

thaumatin-like VVTL1, plus petite donc à priori de structure plus accessible que la chitinase. La collaboration avec le laboratoire SiRMA (Signalisation et Récepteurs Matriciels) de l'URCA a permis en premier lieu de modéliser la structure de VVTL1 par homologie puis d'effectuer des simulations de dynamique moléculaire en boîte d'eau sur la plateforme Romeo2 et une analyse de la flexibilité conformationnelle par modes normaux. Des intérêts communs ont également mené à la mise en place d'une collaboration avec l'IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire) à Illkirch pour valider et analyser la structure de VVTL1. En parallèle de cette étude axée entièrement bioinformatique structurale, nous avons utilisé une approche expérimentale par RMN pour déterminer la structure d'un peptide issu de la séquence de VVTL1, de nouveau en partenariat avec l'ICMR. Dans le cadre d'une collaboration initiée par l'IUVV (Institut Universitaire de la Vigne et du Vin) à Dijon, ce peptide peut ainsi maintenant servir de modèle pour ouvrir la voie vers l'étude de l'adsorption des protéines sur la bentonite, argile utilisée pour la clarification des vins. L'objectif de l'ensemble de cette partie « biologie structurale » ne se définit pas comme la réponse définitive à une problématique initiale, mais plutôt comme l'exploration de nouveaux axes de recherche jusqu'alors non-exploités dans le domaine de l'œnologie. Ces deux premières approches de bioinformatique structurale et de RMN ont donc pour vocation de démontrer qu'elles peuvent naturellement s'appliquer à l'étude des protéines de raisin et permettre d'acquérir des informations capitales sur les mécanismes dans lesquels elles sont impliquées dans la vigne et le Champagne.

# Partie I : Etude des protéines de raisin thaumatin-like et chitinase

*Remember gentlemen, it's not just France we are fighting for, it's Champagne !*

*Winston Churchill (1874-1965)*





# 1 Introduction

## 1.1 Les protéines dans les jus de raisin et les vins de Champagne

Le Champagne est élaboré à partir de trois cépages : Chardonnay, Pinot Noir et Pinot Meunier. Les raisins vendangés sont pressés pour obtenir le jus, autrement appelé moût. Ce dernier subit une première fermentation en cuve, la fermentation alcoolique, grâce à l'introduction de levures qui vont transformer le sucre du raisin en alcool en produisant du gaz carbonique évacué dans l'atmosphère. Une fermentation malolactique est éventuellement induite par des bactéries permettant de transformer l'acide malique en acide lactique. Après clarification (collage, filtration), le vin de base Champenois, non-effervescent, est obtenu. Plusieurs étapes sont ensuite nécessaires avant de parvenir au Champagne : l'assemblage de différents vins de base (combinaison de cépages, de parcelles ou d'années) et le tirage, ou mise en bouteille, avec l'ajout d'une liqueur contenant du sucre et des levures, puis une deuxième fermentation alcoolique ou prise de mousse (le CO<sub>2</sub> reste emprisonné dans la bouteille, le vin devient pétillant), pour finir par un temps de maturation pour permettre aux arômes de se développer et une étape de remuage et de dégorgement pour éliminer le dépôt de levures. Ce procédé constitue la méthode traditionnelle, ou champenoise, définie par une appellation d'origine contrôlée (Figure 1).

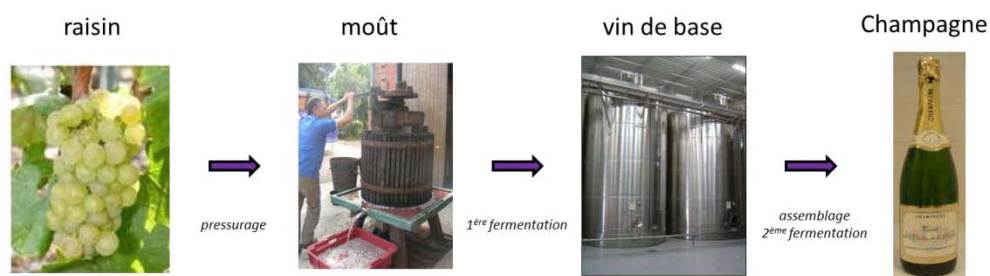


Figure 1 : Principales étapes de l'élaboration d'un Champagne.

### 1.1.1 Pool protéique

Les protéines, de 9 à 65 kDa [1-3], font partie des principales macromolécules des jus de raisin et des vins, avec les polysaccharides. Les protéines de défense (PR pour Pathogenesis-Related) [4] et l'invertase de raisin [5] sont présentes de manière majoritaire dans le pool protéique. Bien qu'elles soient présentes à une faible concentration, quelques dizaines à quelques centaines de mg/L [6-9], elles jouent un rôle majeur dans le domaine de l'œnologie.

La composition du pool protéique d'un vin dépend d'un grand nombre de facteurs : cépage, origine géographique et millésime (« effet terroir »), pratiques culturales et maturité des baies et procédé de vinification [7-15]. Les protéines d'un vin proviennent essentiellement des baies de raisin et éventuellement de micro-organismes [16, 17]. Les levures utilisées au cours de la fermentation alcoolique peuvent en effet influencer la composition du pool protéique de deux manières : par le transfert de protéines dans le vin au cours de l'autolyse et/ou par la sécrétion d'enzymes protéolytiques pouvant dégrader les protéines de raisin [18]. Les organismes pathogènes qui attaquent les baies de raisin, comme le champignon *Botrytis cinerea*, peuvent également modifier la composition finale d'un vin [19, 20].

### 1.1.2 Protéines PR

Le concept de protéines PR a été introduit en 1980 pour désigner les protéines codées par un organisme hôte (plante) en réponse à une interaction avec un agent pathogène [21]. L'induction de protéines PR sous l'action d'un stress biotique suggère, mais ne prouve pas, leur implication dans les mécanismes de défense de la plante [22]. Indépendamment de cette induction, certaines protéines PR sont exprimées de manière constitutive dans les baies de raisin, simplement régulées par le développement de la vigne [23-28]. Dix-sept familles de protéines PR ont jusqu'à présent été référencées, de PR1 à PR17, dans l'ordre de leur découverte [29, 30]. Quatre de ces familles ont été jusqu'à présent identifiées dans un vin : PR2 ( $\beta$ -1,3-glucanase), PR3 (chitinase type I, II, IV, V, VI, VII), PR5 (thaumatin-like) et PR14 (lipid-transfer protein).

Les protéines PR de raisin chitinases et thaumatin-like (TL) sont reportées comme les protéines solubles majoritaires des baies de raisin [7, 26, 27, 31, 32]. Les chitinases (EC 3.2.1.14) sont retrouvées dans une grande diversité d'organismes, contenant ou non de la chitine : virus, bactéries, champignons, plantes et animaux [33-35]. Ces protéines sont capables d'hydrolyser la liaison  $\beta$ -1,4-glucosidique présente dans les biopolymères de N-acétyl-D-glucosamine, notamment la chitine [36]. L'activité antifongique de la plupart des chitinases pourrait être ainsi liée à leur capacité à dégrader la chitine, conduisant à un affaiblissement de la paroi cellulaire des champignons et à la lyse de ces cellules. Cependant, les mécanismes par lesquels les chitinases inhibent le développement de champignons pourraient découler simplement de la présence d'un domaine pouvant lier la chitine plutôt que d'une réelle activité chitinolytique [37]. Les protéines TL peuvent également présenter une activité antifongique ; elles sont décrites dans la Partie 2 section 1.2.

### 1.1.3 Propriétés des protéines

Les protéines sont impliquées dans un grand nombre de phénomènes influençant dans le bon ou le mauvais sens la qualité d'un vin ou d'un Champagne. Entre autres, la formation d'un trouble dans des bouteilles de vins blancs serait ainsi due à l'agrégation de protéines PR chitinases et TL [27, 32, 38]. Un traitement par la bentonite, une argile montmorillonite, peut être mis en place au cours du procédé de vinification pour diminuer la concentration en protéines du vin et ainsi augmenter sa limpidité [39-41]. D'autre part, les protéines contribuent aux propriétés organoleptiques d'un vin par leur interaction avec des arômes [31, 42-44] et par leur rôle tensioactif potentiel dans la formation des bulles et la stabilisation de la mousse d'un Champagne [45-48]. En effet, les protéines présentant une tête hydrophile et une queue hydrophobe pourraient s'adsorber aux interfaces gaz/liquide et faire diminuer la tension de surface. Elles constitueraient ainsi, avec les polysaccharides, la couche d'adsorption autour des bulles de Champagne [49, 50].

## 1.2 Techniques préparatives et analytiques pour l'étude des protéines en œnologie

L'article suivant, "Review of preparative and analytical procedures for the study of proteins in grape juice and wine", a été publié dans le journal *Analytica Chimica Acta* en 2010 [51]. Il présente l'ensemble des techniques appliquées pour obtenir, purifier, quantifier et analyser les protéines dans le domaine de l'œnologie. Les méthodes de chromatographie liquide, d'électrophorèse et immunologiques y sont développées en particulier ; pour un rappel de leurs principes généraux se référer préalablement aux sections 2.1, 2.2 et 2.4.



## Review

## Review of preparative and analytical procedures for the study of proteins in grape juice and wine

D. Le Bourse\*, S. Jégou, A. Conreux, S. Villaume, P. Jeandet

Laboratoire d'Oenologie et Chimie Appliquée, Université de Reims Champagne-Ardenne, URVVC-SE UPRES EA 2069, BP 1039, 51687 Reims Cedex 2, France

## ARTICLE INFO

## Article history:

Received 20 January 2010

Received in revised form 29 March 2010

Accepted 30 March 2010

Available online 7 April 2010

## Keywords:

Protein

Grape juice

Wine

Chromatography

Electrophoresis

Mass spectrometry

## ABSTRACT

Proteins have a great influence on wine quality as they exhibit a various range of properties. In fact, they are involved among others in white wine turbidity, organoleptic characteristics and foam formation in sparkling wines. These compounds could also be of major interest for varietal differentiation, regarding wine authentication and traceability issues.

To provide a better understanding of the role played by these biomolecules in wine processing and explore their potential applications, there is a manifest need for the quantification and characterization of each individual one in terms of sequence, structure and intrinsic and functional properties.

We thus present an overview of preparative and analytical methods for the study of proteins in grape juices and wines, from routine techniques to dedicated methodologies. They include sample preparation with chromatographic methods for the purification and identification of proteins, quantification protocols and characterization procedures such as electrophoretic techniques, immunological methods, sequencing, mass spectrometry, physico-chemical and structural analyses, and so on. We expose advantages and limits of each technique and focus on the different but complementary information they can provide.

Despite the past years advances in the field proteins identification, the elucidation of the full protein profile for grape juices and wines remains strenuous. Their interactions with other wine compounds make the challenge even harder. We therefore emphasize the requirement of the techniques to be refined and suggest the developments to be expected.

© 2010 Elsevier B.V. All rights reserved.

## Contents

1. Introduction.....	34
2. Obtaining protein fractions from grape juice and wine.....	34
3. Methods for the study of proteins from grape juice and wine.....	34
3.1. Quantification.....	34
3.2. Preparative techniques for fractionation and purification of grape juice and wine proteins.....	35
3.2.1. Liquid chromatography.....	35
3.2.2. Electrophoresis.....	36
3.3. Analytical techniques for grape juice and wine proteins.....	37
3.3.1. Liquid chromatography.....	37
3.3.2. Electrophoresis.....	37

**Abbreviations:** PR, pathogenesis-related; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; RP-HPLC, reversed phase high performance liquid chromatography; ELISA, enzyme-linked immunosorbent assay; FPLC, Fast Protein Liquid Chromatography; HPSEC, high performance size exclusion chromatography; HIC, hydrophobic interaction chromatography; ConA, concanavalin A; EPE, electroendosmotic preparative electrophoresis; CE, capillary electrophoresis; CGE, capillary gel electrophoresis; MS, mass spectrometry; FFE, free-flow electrophoresis; IEF, isoelectric focusing; ESI-FTICR-MS, electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry; 2D-E, two-dimensional electrophoresis; MS/MS, tandem mass spectrometry; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; 2D-IF-LDS, two-dimensional isoelectric focusing–lithium dodecyl sulfate; PAS, periodic acid-Schiff; PTA, plate trapped antigen; PVDF, polyvinylidene fluoride; NMR, nuclear magnetic resonance; TOCSY, total correlation spectroscopy; HMQC, heteronuclear multiple quantum coherence; EST, expressed sequence tag; iTRAQ, isobaric tag for relative and absolute quantitation.

\* Corresponding author. Tel.: +33 326 91 34 39; fax: +33 326 33 40.

E-mail address: [doriane.le-bourse@univ-reims.fr](mailto:doriane.le-bourse@univ-reims.fr) (D. Le Bourse).

3.4. Structural analysis of proteins from grape juice and wine .....	38
4. Conclusion.....	40
Acknowledgments .....	40
References .....	40

## 1. Introduction

Although winemaking predates recorded History, this biochemical process appears to be a constant changing area. When talking about winemaking, a number of issues arise, linked to grape berry development and wine making and storage. The main concern for winemakers is the wine quality, which is dependent upon its color, clarity and organoleptic properties [1,2]; and in case of sparkling wine, the elegance of its foam and effervescence.

Proteins, ranging from 9 to 65 kDa [3–5] are one of the main grape juice and wine macromolecules along with polysaccharides and polyphenols. The most abundant proteins in grape juice are pathogenesis-related (PR) proteins, including chitinases and thaumatin-like proteins [6], along with invertase [7]. The same proteins are found in wines [8], in addition to mannoproteins coming from yeast during the juice fermentation [1]. They play a leading part in the wine industry concern, despite their relatively low concentration, around ten to hundreds milligrams per litre [9–12]. They are indeed involved in a number of aspects that can impair the acceptance of the product by consumers, such as the haze formation in white wine attributed to the aggregation of some grape proteins, especially PR proteins [13,14] during bottle storage. Accordingly, treatments during the winemaking process, such as bentonite fining, have been used to lower the protein content in wine thereby enhancing wine clarity and stability [1,15,16]. Another negative property of grape proteins comes from their involvement in some rare cases of grape and wine allergy [17–21]. Proteins can however exhibit positive effects such as the stabilization of foam in sparkling wines [22–24]; their interaction with aroma compounds [25–28]; and the protection of wine against tartaric salt precipitation [29–31]; altogether with the interactions they may be involved in with other wine compounds such as ethanol, polyphenols and polysaccharides [32].

The phenomena the protein could be involved in are thus of major interest but their understanding is not that straightforward. The latter implies a full elucidation of grape juice and wine proteins and peptides profile, and the characterization of each individual one.

For that purpose, considerable efforts have been made to develop analytical methods, from routine techniques for the study of proteins to dedicated methodologies for the analysis of grape juice or wine samples. Since the previous reviews by Moreno-Arribas et al. [33] and Curioni et al. [34], improvements have been achieved to obtain protein extracts, from sample preparation to fractionation and purification techniques, and in the one of their characterization, from identification to quantification and structural analysis.

The scope of the present article is to provide an overview of the current trends and recent contributions in grape juice and wine protein sample preparation, purification and characterization fields, giving references to interesting advances as well as future trends. Advantages and limits of each technique are presented, focusing on the different but complementary information they can provide.

## 2. Obtaining protein fractions from grape juice and wine

The main issue connected to the sample preparation when working on proteins in grape juices and wines follows from

their small concentration and the presence of various compounds that may interfere during analyses. However, once the proteins are extracted from their original environment and processed, they may undergo irreversible denaturation, which would impair the following accuracy of their characterization and alter their functional properties. Researchers must then be aware of the impact of the techniques they set up for the study of proteins.

As described by Moreno-Arribas et al. [33] and Curioni et al. [34], different methods allow obtaining concentrated protein extracts from crude grape juice or wine samples and the removal of other macromolecules and potential contaminants: polyphenolic compounds, carbohydrates, salts, and so on.

The most common ways to isolate proteins often involve precipitation or dialysis and ultrafiltration.

Precipitation is a simple and efficient way for obtaining protein fractions, requiring no specific apparatus but it should be considered with caution as it may denature proteins. Considering proteins are less soluble at high salt concentrations, the addition of large amounts of ammonium sulfate can induce the protein precipitation out of the solution [35]. A potassium dodecyl sulfate (KDS) method was developed especially for wine proteins by Vincenzi et al. [36] and currently finds some applications [37]. A following step of centrifugation is generally required to recover the proteins pellets. Other solvents or acids [12,38] can be used to achieve protein precipitation, but they do not seem to be widely used currently.

Dialysis against water is a well-known non-denaturing method for the removal of low molecular mass compounds used by several authors [7,12,39,40].

Ultrafiltration offers the same advantages as simple dialysis, but also allows the sample concentration in a single step. Actually, this technique has been widely used to obtain proteins and peptides extracts from grape juice and wines thanks to its large adaptability (membrane material, filtration surface, cut-off size) [7,41,42]. Its main limitation could arise from its up to medium yield, the poor selectivity of some membranes and their potential clogging when processing unclear samples.

The Table 1 summarizes two different approaches to obtain protein fractions from must and wine, according respectively to the recent works of Jégou et al. [7] and Marangon et al. [35].

Preparative chromatographic methods, such as molecular exclusion or ion exchange chromatography, are now emerging, as a pre-fractionation step of crude grape juice or wine. They could become a very efficient way to isolate the proteins or peptides of interest from a complex starting matrix, provided the chromatographic buffers do not involve protein denaturation and the overall yield is at least equivalent to the other techniques [43].

## 3. Methods for the study of proteins from grape juice and wine

Proteins exhibit a various range of properties which allows the use of different methods on laboratory scale for their study.

### 3.1. Quantification

A wide range of techniques are available for protein quantification in general, but three major drawbacks when dealing with wine

**Table 1**  
Obtaining protein fractions.

Starting material	Chardonnay must [7]		Semillon grape juice and wine [31]	
Sample preparation	Centrifugation (8000 × g, 10 min)	Removal of biggest aggregates to avoid the clog of filtration membranes Rid of heaviest compounds from the supernatant Concentration and dialysis against water Protein stability	Ammonium sulfate precipitation (80% saturation, 16h, 4 °C)	Modification of protein solubility
	Filtration (0.45 μm membrane)			
	Ultrafiltration (cut-off 10 kDa)			
	Freeze-drying			
Protein fraction	Lyophilized ultraconcentrate		Pellet	Recovery of proteins in the precipitate

and grape juice samples impair the use of some of them. Proteins are typically present at very low concentrations in these media, which leads to the use of techniques with very low detection limits. The presence of contaminants, such as phenolic compounds and ethanol can distort the quantification [44]. Moreover, the absence of standard grape or wine proteins does not allow their direct quantification in a sample.

In the literature, the most common technique for grape juice and wine proteins and polypeptides with molecular mass superior to 3000 Da quantification is the Bradford method [33,34]. This colorimetric assay is based on an absorbance shift in Coomassie Blue G-250. It is a simple and fast method, thus widely used, but which can hardly provide an accurate result, considering the response variation of the reagent to different proteins [45] and the absence of standard grape or wine proteins which does not allow to perform an accurate calibration. Marchal et al. [44] considered the interferences which may falsify the estimation of direct measurement of proteins with the Bradford method in Champagne Pinot Noir and Chardonnay wines and established that ethanol and exogenous and endogenous phenolic compounds seriously impair the quantification of the wine protein content. Other colorimetric methods, such as Lowry, Biuret or Smith (also called bicinchoninic acid method assay) [46–48] can provide interesting results despite their potential interferences with other compounds [37,49,50]. Notably, Vincenzi et al. [36] developed a new procedure for protein recovery and quantification in wine, comparing three colorimetric assays (Bradford, Lowry and Smith). They concluded that the Smith assays give the most accurate results, in accordance to the ones obtained by densitometric quantification of protein bands from sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

In complement to the Bradford method, Esteruelas et al. [51] performed an organic elemental analysis to assess the proportion of nitrogen in a wine precipitate. Total protein quantification can also be achieved by amino acid analysis, as described by Hayasaka et al. [52]. Nitrogen compounds analysis linked to amino-acid analysis can be performed as recently shown by Villamiel et al. [53], using the Kjeldahl method and reversed phase high performance liquid chromatography (RP-HPLC).

In addition to methods that allow total protein quantification in a sample, two other techniques can lead to the quantification of an individual protein in crude or pure sample. The most powerful method is the determination of protein concentration and purity by RP-HPLC, as described by Marangon et al. [35]. The quantification of individual proteins was achieved through comparison to the peak area of two standard proteins (cytochrome c and bovine serum albumin) but one could wonder whereas this approach is truly accurate as these proteins may not give the same HPLC response. High performance size exclusion chromatography (HPSEC) can also be carried out for quantification of wine proteins, as suggested by Nordestgaard et al. [54]. For grape juice and wine proteins having

an enzymatic activity, quantification can be based on their reaction with their substrate. For instance, the active chitinase concentration in a fraction can be determined by measuring its enzymatic activity upon chitin [55].

Other techniques, such as densitometric measurement of bands from SDS-PAGE gels (after Coomassie Brilliant Blue or other stains, or antibody immunostaining) as described by Marchal et al. [56] and Hsu and Heatherbell [3], and immunological assays (enzyme-linked immunosorbent assay (ELISA) method) can give reliable results.

### 3.2. Preparative techniques for fractionation and purification of grape juice and wine proteins

Following isolation and concentration procedures, proteins can undergo separation and characterization steps. This approach can allow the study of individual purified proteins, in terms of structure and functional properties.

#### 3.2.1. Liquid chromatography

With the development of dedicated methodologies, such as fast protein liquid chromatography (FPLC) and the number of media available with different selectivity and efficiency, chromatography became a key tool for the study of proteins and is currently involved in the most part of purification protocols.

HPLC, especially size exclusion (HPSEC), is also used by some authors to study proteins and peptides in berries, grape juice and wine [25,35,38,57–61]. Pocock et al. [62] used for instance a semipreparative C18 HPLC column for the purification and characterization of different grape juice and wine proteins.

FPLC can be used with a wide range of media, all developed by a number of authors: ion exchange, hydrophobic interaction, affinity, gel filtration/size exclusion and chromatofocusing. The form of the sample loaded on the chromatography column varies according to each authors, as well as the conditions of elution and detection techniques. The chromatography step allows to collect purified fractions ready to undergo following characterization steps or can be a powerful tool to compare full protein profiles of different samples. Researchers can set up a one step chromatography technique or use a combination of two or more methods to take advantages of each one and thus achieve an efficient purification.

Different chromatographic techniques have been used for fractionation and purification of grape juice and wine proteins depending on the proteins characteristics they focus on for separation.

A wide number of authors used ion exchange chromatography for grape juice or wine proteins first fractionation [40,49,63–65]. It is based on the reversible interaction between a charged protein and an oppositely charged chromatographic medium, giving a high resolution separation with high sample loading capacity. Proteins bind as they are loaded onto the column, and are then usually eluted

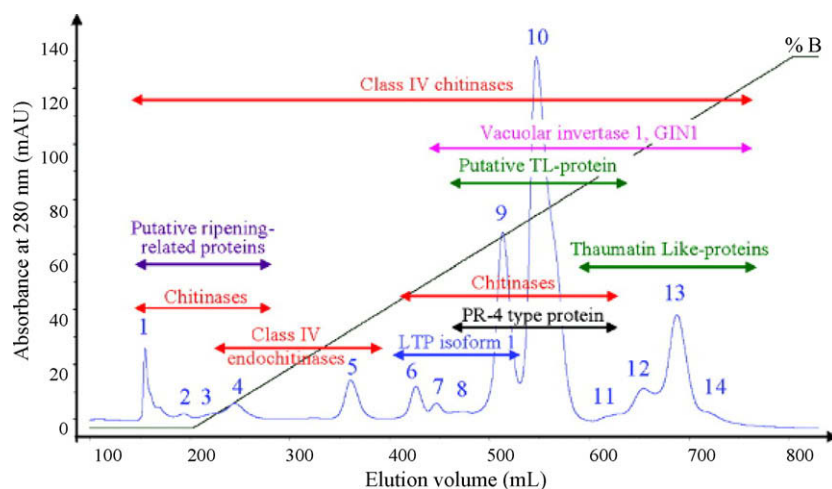


Fig. 1. General scheme of the nature and distribution of grape juice proteins fractionated during HIC. Reprinted from [35], with permission.

by a continuous or stepwise salt gradient (NaCl). Both strong and weak cation and anion exchange techniques can be used, considering major grape juice and wine proteins isoelectric points range from 4.1 and 5.7 [9,66] and are stable and remain soluble over a large pH range.

For instance, Jégou et al. [7] used anion exchange FPLC to purify the grape vacuolar invertase, involving the use of a 7.5 pH elution buffer. At this pH, proteins are positively charged and are selectively eluted in decreasing order of pI by an increasing salt gradient. On the other hand, Van Sluyter et al. [43] used cation exchange FPLC for bulk grape juice fractionation. Cation exchange medium allows the use of an elution buffer pH close to original wine pH of around 3.0 [1]. This team chose a polyacrylamide based column media to allow higher linear flow rates and potentially reduce non-specific interactions of grape proteins with polysaccharide-based gels. As a second step, they developed a method to purify thaumatin-like proteins and chitinases by hydrophobic interaction chromatography (HIC) (Fig. 1), based on previous work [35]. HIC separates proteins with differences in hydrophobicity. The interaction between a protein and the hydrophobic surface of the medium is enhanced by ionic strength buffers, so this technique is often used in relation with ammonium sulfate protein precipitation.

Affinity chromatography is also a common second step after ion exchange chromatography. It can be used with a suitable ligand for the protein of interest. Isoforms of invertase were thus detected in grape berries by Ishikawa et al. [67] using a concanavalin A (ConA) column after an anion exchange chromatography. In the same way, Waters et al. [68] separated a haze protective factor, a macromolecule fraction made up of a polysaccharide and a protein component, from other wine macromolecules by a combination of ConA and anion and cation exchange chromatography. ConA chromatography in combination with preparative isoelectric focusing was also used by Marchal et al. [69] for the purification of glycoproteins of a Champenois wine. Berthier et al. [59] further isolated isolectins from grape juice using ConA and gel filtration after a prepurification step by ion-exchange chromatography, as well as Vanrell et al. [70] who used these four techniques to study the influence of the use of bentonite as riddling agent.

Gel filtration chromatography is also dedicated to separate proteins with different molecular size. It was set up as a single step by Esteruelas et al. [51] to isolate the natural haze protein in white wine or as a second purification step after ion-exchange chromatography to isolate an arabinogalactan-protein from grape berries by Saulnier et al. [71]. This team followed their work by

the characterization of arabinogalactan-proteins from wine [72], adding an affinity chromatography step.

There are only few references in the literature reporting authors resorting to a chromatofocusing protocol, where proteins are separated according to their pIs. However, we could notice the work of Paetzold et al. [73] who fractionated glycoproteins from must over the pH range 4.0–7.0. The limited use of this technique could be linked to the high price of this type of media but an interesting methodology that has been recently developed by Shen et al. [74], allowing chromatofocusing separation with common ion exchange column by using specific buffers, could bring new interest in this approach.

In every chromatography protocols, protein elution is monitored by an UV detector set up at 280 nm (the maximum absorbance of two amino acids, tyrosine and tryptophan) or 210 nm (peptidic bond) [35].

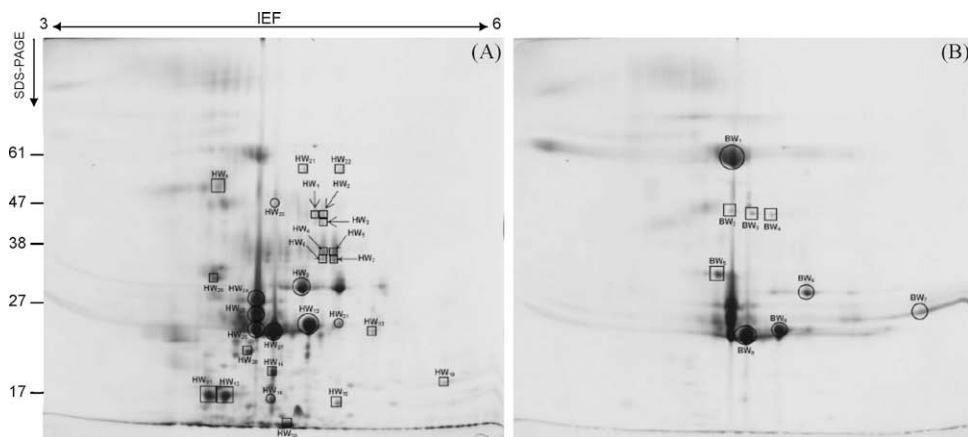
After chromatographic techniques, proteins fractions are usually dialyzed against water to get rid of salts or other contaminants and freeze-dried for long time storage.

### 3.2.2. Electrophoresis

In addition to chromatographic techniques, electrophoresis can be used on preparative scale to study proteins. For instance, Dorrestein et al. [75] used FPLC and PAGE to compare and analyze the soluble proteins of four white wines, as well as Yokotsuka et al. [76] who contributed to the characterization of soluble glycoproteins in red wine by fractionation on gel filtration chromatography and by preparative PAGE.

Vincenzi et al. [77] proposed a one-step purification of nearly all the proteins from wine by electroendosmotic preparative electrophoresis (EPE) using native (N-PAGE) and SDS-PAGE systems. In this method, the proteins of a wine sample, after being electrophoretically separated in a cylindrical gel, are eluted from its bottom through a capillary in which the electroendosmotic buffer flow generated during electrophoresis is channeled. These authors suggest the application of EPE for the production of antibodies with high specificity, considering the high purity and the amount of the protein fractions they obtain.

To date, there are only few references in the literature above the application of capillary electrophoresis (CE) techniques to grape juice or wine proteins [25,40,78–81]. CE can separate proteins according to their isoelectric point, molecular mass, or charge/mass ratio depending on the CE conditions. The CE method of Luguera and co-workers [40,78] using an uncoated silica capillary at alkaline pH, was applied to the study of proteins during sparkling winemak-



**Fig. 2.** Two-dimensional-PAGE of proteins from a healthy (A) and botrytized (B) Champagne base wine. Reprinted from [8], with permission.

ing according to the traditional method with up to 24 months of aging with yeast. The results were compared with those obtained by PAGE and the authors concluded that the use of CE made it possible to present the protein profiles of sparkling wines, which would not have been possible using other electrophoretic techniques. No changes in the protein profiles of sparkling wines were observed during the first 18 months of aging with yeast. A fast analysis of proteins in wines by capillary gel electrophoresis (CGE) has been performed by Rodríguez-Delgado et al. [79] for the first time. The separation in CGE is based on a molecular sieving mechanism with large molecules being retarded. CGE is thus based on the same separation principle than SDS-PAGE, but displays certain advantages over it: rapid analysis time, low sample volume consumption, as well as automatic evaluation and quantification of the separated protein peaks [82].

### 3.3. Analytical techniques for grape juice and wine proteins

The fractions collected after the purification step can be assessed by different means.

#### 3.3.1. Liquid chromatography

Waters and co-workers [35,43,67,83] used RP-HPLC, based on hydrophobic interactions, to identify grape and wine proteins by the determination of retention time of individual proteins. Proteins are eluted in function of increasing hydrophobicity, using an elution gradient with decreasing polarity. This technique allows the determination of the purity of the fraction, its identity, as

well as its quantification, as described by Waters and co-workers [35,43].

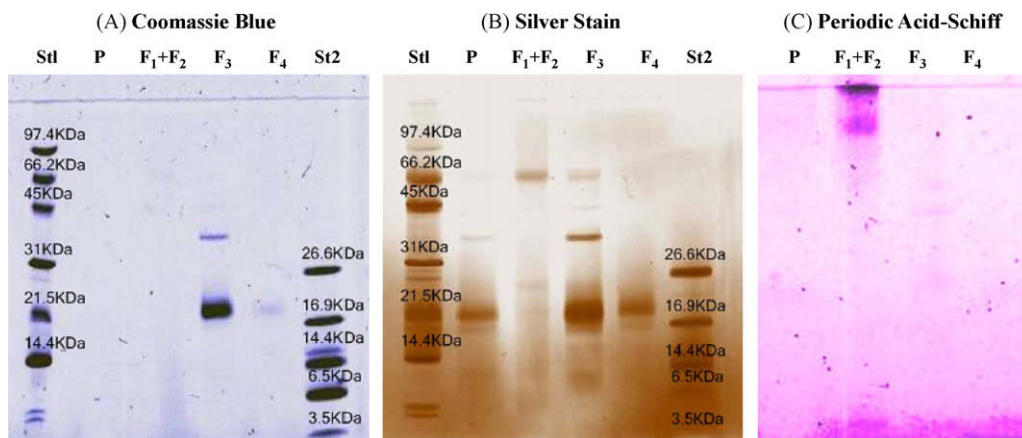
#### 3.3.2. Electrophoresis

Electrophoretic methods are of best interest in the field of analysis of beverages, for the determination of various compounds, as described by Sádecká and Polonský [84]. Electrophoresis can assess fraction purity and estimate proteins molecular masses (accuracy of  $\pm 10\%$  [85]) and pls. Most of the studies on grape and wine proteins have been carried out using the conventional electrophoretic methods of native and SDS-PAGE and IEF.

Two different methods, ion-exchange FPLC and PAGE, were utilized to compare, analyze and fractionate the soluble proteins of four white wines by Dorrestein et al. [75]. Both methods are sensitive, but electrophoresis gave a higher resolution than FPLC.

Qualitative effects of *Botrytis cinerea* infection on a must protein fraction were studied by comparing the electrophoretic patterns of musts obtained from healthy of botrytized grapes [86]. It was found that proteins secreted by *B. cinerea* can degrade grape proteins, which can be linked to the following work of Marchal and co-workers [8,86–88]. Actually, Cilindre et al. [8] interestingly developed the use of two-dimensional electrophoresis (2D-E) and nano-LC-MS/MS and set up the first steps of proteomic approach to study wine proteins, in relation with *B. cinerea* infection (Fig. 2).

The results obtained by Moreno-Arribas et al. [39] confirmed that electrophoretic analysis of proteins is a simple technique that can be used routinely. It also provides complementary infor-



**Fig. 3.** SDS-PAGE of wine proteins fractions using three different stains. Reprinted from [51], with permission.

mation to morphological analysis for varietal characterization of grapevines, and, in most cases, makes it possible to ascertain the grape variety from which musts originate.

Monteiro et al. [89] performed an electrophoretic analysis of the polypeptide composition during grape berry development and revealed some alteration in the protein composition and the presence of a very large number of polypeptides with similar molecular masses but with subtle differences in electric charges.

Electrophoretic methods can be used for the study of all grape and wine proteins, but special attention should be paid for instance when dealing with grape chitinase. Vincenzi and Curioni [90] revealed an anomalous electrophoretic behavior of a chitinase isoform in glycol chitin-containing SDS-PAGE gels. A progressive shift of the relative molecular mass of this enzyme (from 30,500 up to 57,700 Da) with increasing glycol chitin concentration in the gels up to 0.1% was revealed.

A separation of wine proteins by 2D-E and their identification by nano-LC-MS/MS or matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry, as described by Cilindre et al. [8] and Sauvage et al. [91], or by N-terminal amino acid sequences analysis, as described by Okuda et al. [92], opens the way towards further proteomic developments. Since Hsu et al. [3,41] introduced two-dimensional isoelectric focusing-lithium dodecyl sulfate (2D-IF-LDS), recent developments will lead to further developments of this technique.

The detection of proteins after native, SDS-PAGE and IEF is usually carried out by means of staining with dyes. Coomassie Blue G-250, according to Winter et al. [93], is the most commonly used. For samples with low protein concentrations, staining with silver nitrate, following the method of Blum et al. [94] is more suitable, because of its greater sensitivity (up to 100 times) [33]. Other specific methods, such as periodic acid-Schiff (PAS) staining method for glycoproteins can be used [95]. Esteruelas et al. [51] thus used these three different stains to study the same sample and compare the complementary information they provide (Fig. 3).

A special attention should be paid to the quantification of protein bands by computer-assisted densitometry after staining, considering different proteins may not give identical response to dyes [35,96] as developed in Section 3.1.

Proteins separated by gel electrophoresis can be electrophoretically transferred onto a membrane to carry out immunodetection assays, such as western blot. Grape or wine proteins (bulk or purified after preparative technique) are utilized as antigens to immunize rabbits, by methods that ensure the production of highly specific polyclonal antibodies.

Monteiro et al. [50,63,64,97] and Ferreira et al. [98] proposed a study of wine proteins by immunological methods. After the production of highly specific antibodies, this team was allowed to detect the presence of identical or similar proteins in wines prepared from different grape varieties. They raised polyclonal antibodies in rabbits against the total proteins from a Portuguese wine as well as against specific proteins isolated and purified by FPLC cation exchange chromatography in combination with denaturing electrophoresis. They measured the titre of the antisera by the ELISA technique and assessed the specificity of the antibodies by immunoblotting. The use of monoclonal antibodies for grape or wine proteins has never been described. However, Ruiz and Ruffner [99] demonstrated that highly specific polyclonal antibodies can be obtained using a careful protocol for the study of invertase in grapes infected with *B. cinerea*. They used the antibodies they prepared for immunoblotting as described by Towbin et al. [100] and for ELISA. Dambrouck et al. [101] used immunodetection technique to analyze the origin of proteins of a Chardonnay wine. They raised polyclonal antibodies against must, yeast and bacteria proteins and showed that most of the wine proteins come from grapes with many of them being glycosylated. Some

wine proteins are high molecular mass mannoproteins coming from the yeast and are released during the alcoholic fermentation, whereas no bacteria proteins were detected in the wine. Dambrouck et al. [102] followed their work and evaluated the effect of fining on total protein and grape invertase contents of Champenois base wines in relation to their foaming properties. That study can be linked to the results obtained by Manteau et al. [103], who studied the changes in chitinase and thaumatin-like proteins of grape berries during the winemaking process. Immunodetection using plate trapped antigen-ELISA (PTA-ELISA) allowed Dambrouck et al. [102] to assess the grape invertase content while treating a wine with bentonite fining. They thus proved that fining leads in most cases to a decrease in invertase content, along with an alteration of foaming properties of the treated wines in terms of foam height and foam stability. Obanor et al. [104] also used PTA-ELISA detection and quantification of *B. cinerea* infection in grapes.

In another extend, immunology can be used to localize proteins in grapes. Famiani et al. [105] developed in the same approach an immunohistochemical study of the compartmentation of sugar, organic acid and amino acid metabolism during the development of grape berries. Dupin et al. [106] showed that a specific *Saccharomyces cerevisiae* mannoprotein exhibiting a wine protein haze-protective activity was located primarily on the outermost and innermost layers of the yeast cell wall using immunoelectron microscopy with immuno-gold labeling.

New developments in the field of immunological methods, such as improvements of the preparation of antibodies or of the experiments protocols will open a way towards new discovery. Cilindre et al. [8] undertook a thorough study using 2D-E and immunodetection, of the protein content of two Chardonnay base wines prepared with healthy and botrytized grapes (Fig. 4).

Monteiro et al. [49,50] also used both techniques for the study of grapes protein profiles from consecutive years to study the influence of environmental conditions.

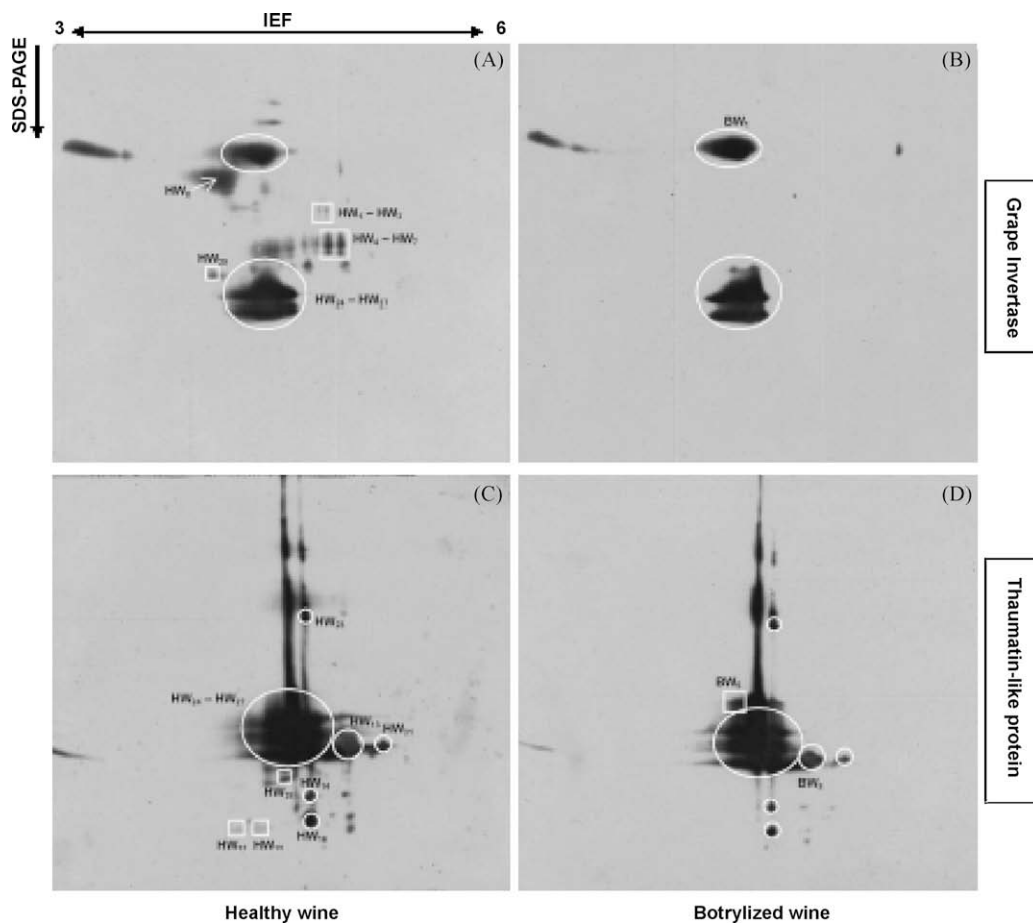
Activity assays can be performed when studying enzymes from grapes and can give useful information for the characterization of these enzymes [67,103,107]. The work of Vincenzi and Curioni [90] is the first report of the study of grape and wine proteins by SDS-PAGE with detection of chitinolytic activity on gels. Ruiz and Ruffner [99] developed invertase assays to study the impact of *B. cinerea* infection on the host assimilate metabolism. Other teams [55,102,108] used enzyme activity in liquid media or on gels.

#### 3.4. Structural analysis of proteins from grape juice and wine

Grape juice and wine proteins exhibit a wide range of properties of enological interest. To provide a better understanding of these macroscopic properties, there is thus a need for the study of protein structure.

The recent years developments in mass spectrometry induced a great improvement in the field of protein characterization, as presented by Flamini and De Rosso [109] in their very interesting review. First of all, MS can provide complementary information to SDS-PAGE by assessing molecular masses [50,62], but also provide structural information (identification of proteins). Different MS techniques have been successfully applied to study the grape and wine proteins. Liquid chromatography electrospray ionization (LC-ESI) MS and nano-LC/MS have been used to identify grape and wine proteins and peptides [7,35,43,96,110–113]. Protein and peptide fingerprinting was achieved by different authors using MALDI-TOF and surface-enhanced laser desorption/ionization TOF [91,114–117]. In addition to the work of Flamini and De Rosso [109], interested readers could appreciate the following articles: [34,110,118–121]. A recent review, published in 2008, dealing with





**Fig. 4.** Two-dimensional immunoblots of healthy and botrytized wine proteins probed with invertase and thaumatin-like protein polyclonal antibodies. Reprinted from [8], with permission.

the current role of mass spectrometry for proteome research by Chen [122] can also be of major interest.

MS is a powerful tool for amino acid sequence determination, in link with HPLC for amino acid analysis [42,55,60,65,69,123]. For the latter, purified proteins are subjected to SDS-PAGE (1D or 2D) and then electroblotted to a polyvinylidene fluoride (PVDF) membrane. Protein bands are visualized with stain and excised from the membrane. They then undergo Edman degradation (protein sequencer on line with HPLC analyzer). Saulnier and Brillouet [71] induced in 1989 a first study of the structure of a grape arabinogalactan-protein, revealing the part of neutral sugars and uronic acids in that molecule, as well as a contribution to the study of amino acids involved in carbohydrate-protein linkages, thanks to enzymic degradation, methylation analysis and reductive alkaline degradation approach. Waters and coworkers also offered a further contribution to the amino acid composition of that protein material [68,72,124]. Gomez-Alonso et al. [125] brought improvement for the HPLC analysis of amino acids from wine and beer sample. Waters et al. [111] presented a sequence analysis of grape berry chitinases.

MS allows the determination of protein molecular mass and amino acids sequence, but it may find other application. Grapevine identification and wine varietal differentiation have been raising concern. Although ampelography is a powerful tool when dealing with grapevine, the development of analytical methods allowed the apparition of efficient techniques dedicated to grape juice and wine varietal differentiation. For instance, the use of DNA fingerprinting techniques can be applied for the study of juice [126]. It has been proven that PR proteins (thaumatin-like proteins and chitinases)

exist under a number of isoforms and their molecular mass might differ slightly across varietal boundaries. The grape proteins might constitute a molecular tag for identification of the grape varieties that have been employed to produce a given wine, as suggested by the work of Hayasaka et al. [52]. This team proposed a base varietal differentiation of grapes on the PR polypeptide profile present in mature grape berries using electrospray mass spectrometry. In contrast, the work of Monteiro et al. [50] raises serious questions about the applicability of this method. Varietal differentiation could also be achieved by MALDI-MS analysis of the grape seed proteins [127]. These studies confirm the increasing interest to establish a reliable tool to identify the origin of wine and grape products, in a consumer protection concern [128].

A main tool for the study of the three dimensional structure of a protein is the determination of its crystal structure. Many works have been reported presenting the resolution of the crystal structure of thaumatin-like protein [129–133] and chitinase [134–136] from other species, but only a few concerning grape proteins. Van Sluyter et al. [43] recently managed to produce four crystals of thaumatin-like protein isoforms and found one of them was very similar to the structure of a banana thaumatin-like protein [133]. These crystals are to be subjected to X-rays for structural determinations [43].

Protein structure can also be determined by other techniques. Gonçalves et al. [61] examined the structure of a white wine mannoprotein by  $^1\text{H}$  and  $^{13}\text{C}$  nuclear magnetic resonance (NMR) spectroscopic techniques such as 1-D or 2-D total correlation spectroscopy (TOCSY) and 2-D heteronuclear multiple quantum coherence (HMQC). Although NMR allows the determination of

three-dimensional structure in a liquid form, assessed to be more accurate than the one determined by crystallography, no reports of the use of this technique were found in the literature for grape proteins.

Recently, Falconer et al. [137] published an innovating work, which provides a great contribution to the study of the secondary structure study of wine proteins. They used differential scanning calorimetry and circular dichroism spectrometry to study the role of chitinase in wine protein haze and the kinetics of its unfolding. Circular dichroism spectrometry has also been recently used by Marzban et al. [138] to study conformational changes of a thaumatin-like protein from apple.

The past years developments in the field of bioinformatics methods are expected to allow breakthrough in the field of the structure of proteins, as suggested by Shi et al. [139], and of the improvement of protein database, considering the recent generation of a predicted protein database from expressed sequence tag (EST) data and application to isobaric tag for relative and absolute quantitation (iTRAQ) analyses in grape berries at ripening initiation [140].

#### 4. Conclusion

The study of grape juice and wines proteins is of major interest as they may affect wine quality although they can be found in low concentrations.

The availability of single proteins in highly purified form is a major requirement for their biochemical characterization and for understanding the physico-chemical mechanisms of their effects on wine properties. Many procedures are used to fractionate grape juice and wine proteins, mainly chromatography (HPLC and FPLC) and electrophoresis (CE, PAGE, IEF) techniques. The key to successful and efficient protein purification is to select the most appropriate techniques, optimize their performances to suit the requirements and combine them in a logical way to maximize yield and minimize the number of steps required.

Over the past decades, immunology has been an expanding area of biological research, with the development of western blot and ELISA techniques. Immunohistochemistry could also open the way towards the study of the localization of proteins in grape berries.

The primary structure knowledge of grape and wine proteins of a wide molecular mass range will be continuously enlarged by valuable structural information provided by LC-MS, ESI-MS, MALDI-TOF and tandem MS approaches. According to Flamini and De Rosso [109], FTICR mass spectrometers would be able to achieve a higher mass accuracy than modern TOF instruments in future years. The use of 2D-E linked with MS started to provide interesting results in a proteomic approach, and there is no doubt this technique will show a great growth in the future.

The recent advances in the study of proteins also open the way towards their potential application in varietal differentiation of wines, considering authenticity and traceability issues emerge as a main concern for all the enology partners. In addition, future developments in the field of plant-pathogen interactions by the mean of PR proteins would allow improvements in protecting vineyards against devastating diseases, such as *B. cinerea* or powdery mildew infections. The improvement of grapevine defenses against pathogenic agents represents a key issue for enology partners. Van Sluyter et al. [43] suggest the future use of purified PR proteins in pathogen bioassays, as substrates in protease assays, for protein precipitation kinetics studies and for protein melting experiments. Monteiro et al. [63] proposed a putative general defense mechanism of osmotin and thaumatin-like proteins against pathogenic fungi using antifungal assays. Given the physiological role attributed to PR proteins, genetically modified grapevines overexpressing the PR proteins would probably become more resistant to the attack of fungal pathogens.

Another expanding field for the study of proteins from grape is linked with the improvement of grape genes knowledge and expression [141,142] and the development of heterologous proteins production by recombinant organism. Let's bet this research area will open the way towards great breakthrough in the world of grape juice and wine.

#### Acknowledgments

We thank the Champagne Ardenne Area, the General Council of Marne, the Ministry of Research, the Ville de Reims, the A.R.O.C.U. Association, and the Champagne Moët & Chandon for their support.

#### References

- [1] R.B. Ferreira, M.A. Piçarra-Pereira, S. Monteiro, V.B. Loureiro, A.R. Teixeira, *Trends Food Sci. Tech.* 12 (2001) 230–239.
- [2] E.J. Waters, C.B. Colby, in: M.V. Moreno-Arribas, M.C. Polo (Eds.), *Wine Chemistry and Biochemistry*, Springer, New York, 2009, pp. 213–230.
- [3] J.C. Hsu, D.A. Heatherbell, *Am. J. Enol. Vitic.* 38 (1987) 6–10.
- [4] F. Brissonnet, A. Maujean, *Am. J. Enol. Vitic.* 44 (1993) 297–301.
- [5] O. Lamikanra, I.D. Inyang, *Am. J. Enol. Vitic.* 39 (1988) 113–116.
- [6] S.P. Robinson, C. Davies, *Aust. J. Grape Wine Res.* 6 (2000) 175–188.
- [7] S. Jégou, A. Conreux, S. Villaume, A. Hovasse, C. Schaeffer, C. Cilindre, A. Van Dorsseleer, P. Jeandet, *Anal. Chim. Acta* 638 (2009) 75–78.
- [8] C. Cilindre, S. Jégou, A. Hovasse, C. Schaeffer, A.J. Castro, C. Clément, A. Van Dorsseleer, P. Jeandet, R. Marchal, *J. Proteome Res.* 7 (2008) 1199–1208.
- [9] T.H. Lee, *Physical Stability of Wine*, 1986, pp. 23–39.
- [10] K.F. Pocock, E.J. Waters, *Aust. J. Grape Wine Res.* 4 (1998) 136–139.
- [11] J.C. Hsu, D.A. Heatherbell, *Am. J. Enol. Vitic.* 38 (1987) 11–16.
- [12] F.C. Bayly, H.W. Berg, *Am. J. Enol. Vitic.* 18 (1967) 18–32.
- [13] K.F. Pocock, G.M. Alexander, Y. Hayasaka, P.R. Jones, E.J. Waters, *J. Agric. Food Chem.* 55 (2007) 1799–1807.
- [14] K.F. Pocock, E.J. Waters, *Aust. J. Grape Wine Res.* 12 (2006) 212–220.
- [15] W.H. Blade, R. Boulton, *Am. J. Enol. Vitic.* 39 (1988) 193–199.
- [16] P.B. Hoj, D.B. Tattersall, K. Adams, K.F. Pocock, Y. Hayasaka, R. Van Heeswijck, E.J. Waters, *Am. J. Enol. Vitic.* 51 (2001) 147–154.
- [17] E.A. Pastorello, L. Farioli, V. Pravettoni, C. Ortolani, D. Fortunato, M. Gabriella Giuffrida, L. Perono Garoffo, A. Marianna Calamari, O. Brenna, A. Conti, *J. Allergy Clin. Immun.* 111 (2003) 350–359.
- [18] M. Sbornik, J. Rakoski, M. Mempel, M. Ollert, J. Ring, *Allergy* 62 (2007) 1339–1340.
- [19] S.G. Schäd, J. Trcka, S. Vieths, S. Scheurer, A. Conti, E.-B. Bröcker, A. Trautmann, *Int. Arch. Allergy Appl. Immunol.* 136 (2005) 159–164.
- [20] E. Vassilopoulou, L. Zuidmeer, J. Akkerdaas, I. Tassios, N.R. Rigby, E.N.C. Mills, R. Van Ree, P. Saxoni-Papageorgiou, N.G. Papadopoulos, *Int. Arch. Allergy Appl. Immunol.* 143 (2007) 92–102.
- [21] P. Rougé, J.-P. Borges, R. Culierrier, C. Brulé, A. Didier, A. Barre, Les protéines de transfert des lipides: des allergènes importants des fruits, *Rev. Fr. Allergol.* 49 (2009) 58–61.
- [22] T. Girbau-Solà, E. López-Tamames, J. Buján, S. Buxaderas, *J. Agric. Food Chem.* 50 (2002) 5596–5599.
- [23] G. Liger-Belair, G. Polidori, P. Jeandet, *Chem. Soc. Rev.* 37 (2008) 2490–2511.
- [24] J. Senée, B. Robillard, M. Vignes-Adler, *Food Hydrocolloids* 13 (1999) 15–26.
- [25] C. Desportes, M. Charpentier, B. Duteurtre, A. Maujean, F. Duchiron, *Am. J. Enol. Vitic.* 52 (2001) 376–380.
- [26] S. Lubbers, A. Voilley, M. Feuillat, C. Charpentier, *LWT-Food Sci. Technol.* 27 (1994) 108–114.
- [27] P.R. Jones, R. Gawel, I.L. Francis, E.J. Waters, *Food Qual. Prefer.* 19 (2008) 596–607.
- [28] Z. Peng, K.F. Pocock, E.J. Waters, I.L. Francis, P.J. Williams, *J. Agric. Food Chem.* 45 (1997) 4639–4643.
- [29] V. Gerbaud, N. Gabas, J. Blouin, P. Pellerin, M. Moutonnet, *J. Int. Sci. Vigne Vin.* 31 (1997) 65–83.
- [30] S. Lubbers, C. Charpentier, M. Feuillat, A. Voilley, *J. Int. Sci. Vigne Vin.* 27 (1993) 13–22.
- [31] V. Moine-Ledoux, A. Perrin, I. Paladin, D. Dubourdieu, *J. Int. Sci. Vigne Vin.* 31 (1997) 23–31.
- [32] V. Aguié-Béghin, Y. Adriaensen, N. Péron, M. Valade, P. Rouxhet, R. Douillard, *J. Agric. Food Chem.* 57 (2009) 10399–10407.
- [33] M.V. Moreno-Arribas, E. Pueyo, M.C. Polo, *Anal. Chim. Acta* 458 (2002) 63–75.
- [34] A. Curioni, S. Vincenzi, R. Flamini, in: R. Flamini (Ed.), *Hyphenated Techniques in Grape and Wine Chemistry*, 2008, pp. 249–287.
- [35] M. Marangon, S.C. Van Sluyter, P.A. Haynes, E.J. Waters, *J. Agric. Food Chem.* 57 (2009) 4415–4425.
- [36] S. Vincenzi, S. Mosconi, G. Zoccatelli, C.D. Pellegrina, G. Veneri, R. Chignola, A. Peruffo, A. Curioni, C. Rizzi, *Am. J. Enol. Vitic.* 56 (2005) 182–187.
- [37] M. Fusi, F. Mainente, C. Rizzi, G. Zoccatelli, B. Simonato, *Food Control* 21 (2009) 830–834.
- [38] M.V. Moreno-Arribas, E. Pueyo, M.C. Polo, *J. Agric. Food Chem.* 44 (1996) 3783–3788.

- [39] M.V. Moreno-Arribas, F. Cabello, M.C. Polo, P.J. Martín-Alvarez, E. Pueyo, J. Agric. Food Chem. 47 (1999) 114–120.
- [40] C. Luguera, M.V. Moreno-Arribas, E. Pueyo, B. Bartolome, M.C. Polo, Food Chem. 63 (1998) 465–471.
- [41] J.C. Hsu, D.A. Heatherbell, J.H. Flores, B.T. Watson, Am. J. Enol. Vitic. 38 (1987) 17–22.
- [42] E.J. Waters, W. Wallace, P.J. Williams, J. Agric. Food Chem. 40 (1992) 1514–1519.
- [43] S.C. Van Sluyter, M. Marangon, S.D. Stranks, K.A. Neilson, Y. Hayasaka, P.A. Haynes, R. Ian Menz, E.J. Waters, J. Agric. Food Chem. 57 (2009) 11376–11380.
- [44] R. Marchal, V. Seguin, A. Maujean, Am. J. Enol. Vitic. 48 (1997) 303–309.
- [45] M. Tal, A. Silberstein, E. Nusser, J. Biol. Chem. 260 (1985) 9976–9980.
- [46] O.H. Lowry, N.J. Rosebrough, A.L. Farr, R.J. Randall, J. Biol. Chem. 193 (1951) 139–149.
- [47] A. Bensadoun, D. Weinstein, Anal. Biochem. 92 (1976) 497–500.
- [48] A. Bensadoun, D. Weinstein, Anal. Biochem. 70 (1976) 241–252.
- [49] S. Monteiro, M.A. Piçarra-Pereira, V.B. Loureiro, A.R. Teixeira, R.B. Ferreira, Phytochemistry 68 (2007) 416–425.
- [50] S. Monteiro, M.A. Piçarra-Pereira, A.R. Teixeira, V.B. Loureiro, R.B. Ferreira, J. Agric. Food Chem. 51 (2003) 4046–4053.
- [51] M. Esteruelas, P. Poinssaut, N. Siczekowski, S. Manteau, M.F. Fort, J.M. Canals, F. Zamora, Food Chem. 113 (2009) 28–35.
- [52] Y. Hayasaka, K.S. Adams, K.F. Pocock, G.A. Baldock, E.J. Waters, P.B. Høj, J. Agric. Food Chem. 49 (2001) 1830–1839.
- [53] M. Villamiel, M.C. Polo, M.V. Moreno-Arribas, LWT-Food Sci. Technol. 41 (2008) 1842–1846.
- [54] S.J. Nordestgaard, C.B. Colby, B.K. O'Neill, E.J. Waters, in: M. Rhodes (Ed.), *Chemeca 2007: Academia and Industry: Strengthening the Profession*, 2007, pp. 1154–1164.
- [55] R.A. Salzman, I. Tikhonova, B.P. Bordelon, P.M. Hasegawa, R.A. Bressan, Plant Physiol. 117 (1998) 465–472.
- [56] R. Marchal, D. Chaboche, L. Marchal-Delahaut, C. Gerland, J.P. Gandon, P. Jeandet, J. Agric. Food Chem. 48 (2000) 3225–3231.
- [57] K. Yokotsuka, V.L. Singleton, Am. J. Enol. Vitic. 48 (1997) 100–114.
- [58] M.V. Moreno-Arribas, B. Bartolomé, E. Pueyo, M.C. Polo, J. Agric. Food Chem. 46 (1998) 3422–3425.
- [59] L. Berthier, R. Marchal, H. Debray, E. Bonnet, P. Jeandet, A. Maujean, J. Agric. Food Chem. 47 (1999) 2193–2197.
- [60] C. Desportes, M. Charpentier, B. Duteurtre, A. Maujean, F. Duchiron, J. Chrom. A 893 (2000) 281–291.
- [61] F. Gonçalves, A. Heyraud, M.N. De Pinho, M. Rinaudo, J. Agric. Food Chem. 50 (2002) 6097–6101.
- [62] K.F. Pocock, Y. Hayasaka, M.G. McCarthy, E.J. Waters, J. Agric. Food Chem. 48 (2000) 1637–1643.
- [63] S. Monteiro, M. Barakat, M.A. Piçarra-Pereira, A.R. Teixeira, R.B. Ferreira, Phytopathology 93 (2003) 1505–1512.
- [64] S. Monteiro, M.A. Piçarra-Pereira, P.R. Mesquita, V.B. Loureiro, A. Teixeira, R.B. Ferreira, J. Agric. Food Chem. 49 (2001) 3999–4010.
- [65] R.A. Muhlack, E.J. Waters, A. Lim, B.K. O'Neill, C.B. Colby, Asia-Pac. J. Chem. Eng. 2 (2007) 70–74.
- [66] H. Dawes, S. Boyes, J. Keene, D. Heatherbell, Am. J. Enol. Vitic. 45 (1994) 319–326.
- [67] N. Ishikawa, H. Nakagawa, N. Ogura, Agric. Biol. Chem. 53 (1989) 837–838.
- [68] E.J. Waters, W. Wallace, M.E. Tate, P.J. Williams, J. Agric. Food Chem. 41 (1993) 724–730.
- [69] R. Marchal, S. Bouquelet, A. Maujean, J. Agric. Food Chem. 44 (1996) 1716–1722.
- [70] G. Vanrell, R. Canals, M. Esteruelas, F. Fort, J.M. Canals, F. Zamora, Food Chem. 104 (2007) 148–155.
- [71] L. Saulnier, J.-M. Brillouet, Carbohydr. Res. 188 (1989) 137–144.
- [72] P. Pellerin, E. Waters, J.-M. Brillouet, Carbohydr. Polym. 22 (1993) 187–192.
- [73] M. Paetzold, L. Dulau, D. Dubourdieu, J. Int. Sci. Vigne Vin. 24 (1990) 13–28.
- [74] H. Shen, X. Li, C.J. Bieberich, D.D. Frey, Methods Molec. Biol. (Clifton, NJ) 424 (2008) 187–203.
- [75] E. Dorrestein, R.B. Ferreira, O. Laureano, A.R. Teixeira, Am. J. Enol. Vitic. 46 (1995) 235–242.
- [76] K. Yokotsuka, K. Nozaki, T. Takayanagi, Am. J. Enol. Vitic. 45 (1994) 410–416.
- [77] S. Vincenzi, G. Zoccatelli, C. Rizzi, A. Curioni, in: A. Lonvaud, G. De Revel, P. Darriet (Eds.), *Oenologie 2003, 7e Symposium International d'Oenologie, Tec&Doc, Paris, 2003*, pp. 538–541.
- [78] C. Luguera, M.V. Moreno-Arribas, E. Pueyo, M.C. Polo, J. Agric. Food Chem. 45 (1997) 3766–3770.
- [79] M.A. Rodríguez-Delgado, S. Malovaná, F.J. Montelongo, A. Cifuentes, Eur. Food Res. Technol. 214 (2002) 536–540.
- [80] D. Chabreyrie, S. Chauvet, F. Guyon, M.-H. Salagoity, J.-F. Antinelli, B. Medina, J. Agric. Food Chem. 56 (2008) 6785–6790.
- [81] M. Dizzy, L.F. Bisson, Am. J. Enol. Vitic. 50 (1999) 120–127.
- [82] A. Guttman, Electrophoresis 17 (1996) 1333–1341.
- [83] T. Girbau, B.E. Stummer, K.F. Pocock, G.A. Baldock, E.S. Scott, E.J. Waters, Aust. J. Grape Wine Res. 10 (2004) 125–133.
- [84] J. Sádecká, J. Polonský, J. Chrom. A 880 (2000) 243–279.
- [85] R.T. Swank, K.D. Munkres, Anal. Biochem. 39 (1971) 462–477.
- [86] R. Marchal, L. Berthier, L. Legendre, L. Marchal-Delahaut, P. Jeandet, A. Maujean, J. Agric. Food Chem. 46 (1998) 4945–4949.
- [87] R. Marchal, M. Warchol, C. Cilindre, P. Jeandet, J. Agric. Food Chem. 54 (2006) 5157–5165.
- [88] C. Cilindre, A.J. Castro, C. Clément, P. Jeandet, R. Marchal, Food Chem. 103 (2007) 139–149.
- [89] S. Monteiro, M.A. Piçarra-Pereira, L.M. Batista, V.B. Loureiro, A.R. Teixeira, R.B. Ferreira, Vitis 45 (2006) 149–150.
- [90] S. Vincenzi, A. Curioni, Electrophoresis 26 (2005) 60–63.
- [91] F.-X. Sauvage, B. Bach, M. Moutounet, A. Vernhet, Food Chem. 118 (2010) 26–34.
- [92] T. Okuda, M. Fukui, T. Takayanagi, K. Yokotsuka, Food Sci. Technol. Res. 12 (2006) 131–136.
- [93] A. Winter, K. Ek, V.B. Anderson, LKB Appl. Note (1977) 250.
- [94] H. Blum, H. Beier, H.J. Gross, Electrophoresis 8 (1987) 93–99.
- [95] A.E. Eckhardt, C.E. Hayes, I.J. Goldstein, Anal. Biochem. 73 (1976) 192–197.
- [96] S.W. Kwon, J. Agric. Food Chem. 52 (2004) 7258–7263.
- [97] S. Monteiro, M.A. Piçarra-Pereira, M.C. Tanganho, J.P. Rente, V.B. Loureiro, A.R. Teixeira, R.B. Ferreira, J. Sci. Food Agric. 79 (1999) 772–778.
- [98] R.B. Ferreira, S. Monteiro, M.A. Piçarra-Pereira, M.C. Tanganho, V.B. Loureiro, A.R. Teixeira, Am. J. Enol. Vitic. 51 (2000) 22–28.
- [99] E. Ruiz, H.P. Ruffner, J. Phytopathol. 150 (2002) 76–85.
- [100] H. Towbin, T. Staehelin, J. Gordon, Proc. Natl. Acad. Sci. U.S.A. 76 (1979) 4350–4354.
- [101] T. Dambrouck, R. Marchal, L. Marchal-Delahaut, M. Parmentier, A. Maujean, P. Jeandet, J. Agric. Food Chem. 51 (2003) 2727–2732.
- [102] T. Dambrouck, R. Marchal, C. Cilindre, M. Parmentier, P. Jeandet, J. Agric. Food Chem. 53 (2005) 8782–8789.
- [103] S. Manteau, B. Lambert, P. Jeandet, L. Legendre, Am. J. Enol. Vitic. 54 (2003) 267–272.
- [104] F.O. Obanon, K. Williamson, D.C. Mundy, P.N. Wood, M. Walter, New Plant Prot. 57 (2004) 130–137.
- [105] F. Fiamani, R.P. Walker, L. Técsi, Z.-H. Chen, P. Proietti, R.C. Leegood, J. Exp. Bot. 51 (2000) 675–683.
- [106] I.V.S. Dupin, V.J. Stockdale, P.J. Williams, G.P. Jones, A.J. Markides, E.J. Waters, J. Agric. Food Chem. 48 (2000) 1086–1095.
- [107] J. Trudel, A. Asselin, Anal. Biochem. 105 (1989) 673–678.
- [108] S. Van Sluyter, M.J. Durako, C.J. Halkides, Am. J. Enol. Vitic. 56 (2005) 81–85.
- [109] R. Flamini, M. De Rosso, Expert Rev. Proteomics 3 (2006) 321–331.
- [110] P. Wigand, S. Tenzer, H. Schild, H. Decker, J. Agric. Food Chem. 57 (2009) 4328–4333.
- [111] E.J. Waters, Y. Hayasaka, D.B. Tattersall, K.S. Adams, P.J. Williams, J. Agric. Food Chem. 46 (1998) 4950–4957.
- [112] M. De Person, A. Sevestre, P. Chaimbault, L. Perrot, F. Duchiron, C. Elfakir, Anal. Chim. Acta 520 (2004) 149–158.
- [113] R.R. Catharine, I.B.S. Cunha, A.O. Fogaça, E.M.P. Facco, H.T. Godoy, C.E. Daudt, M.N. Eberlin, A.C.H.F. Sawaya, J. Mass Spectrom. 41 (2006) 185–190.
- [114] A. Chambery, G. Del Monaco, A. Di Maro, A. Parente, Food Chem. 113 (2009) 1283–1289.
- [115] A. Carpentieri, G. Marino, A. Amoresano, Anal. Bioanal. Chem. 389 (2007) 969–982.
- [116] K.C. Weiss, T.-T. Yip, T.W. Hutchens, L.F. Bisson, Am. J. Enol. Vitic. 49 (1998) 231–239.
- [117] Z. Szilágyi, G. Vas, G. Mády, K. Vékey, Rapid Commun. Mass Spectrom. 10 (1996) 1141–1143.
- [118] N.D. Padliya, W.M. Garrett, K.B. Campbell, D.L. Tabb, B. Cooper, Proteomics 7 (2007) 3932–3942.
- [119] D.N. Perkins, D.J.C. Pappin, D.M. Creasy, J.S. Cottrell, Electrophoresis 20 (1999) 3551–3567.
- [120] J.X. Yan, R. Wait, T. Berkelman, R.A. Harry, J.A. Westbrook, C.H. Wheeler, M.J. Dunn, Electrophoresis 21 (2000) 3666–3672.
- [121] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Anal. Chem. 68 (1996) 850–858.
- [122] C.-H.W. Chen, Anal. Chim. Acta 624 (2008) 16–36.
- [123] E.J. Waters, N.J. Shirley, P.J. Williams, J. Agric. Food Chem. 44 (1996) 3–5.
- [124] E.J. Waters, P. Pellerin, J.M. Brillouet, Biosci. Biotechnol. Biochem. 58 (1994) 43–48.
- [125] S. Gómez-Alonso, I. Hermosín-Gutiérrez, E. García-Romero, J. Agric. Food Chem. 55 (2007) 608–613.
- [126] R. Siret, O. Gigaud, J.P. Rosec, P. This, J. Agric. Food Chem. 50 (2002) 3822–3827.
- [127] I.C. Pesavento, A. Bertazzo, R. Flamini, A.D. Vedova, M. De Rosso, R. Seraglia, P. Traldi, J. Mass Spectrom. 43 (2008) 234–241.
- [128] R. Flamini, A. Panighel, Mass Spectrom. Rev. 25 (2006) 741–774.
- [129] H. Koiba, H. Kato, T. Nakatsu, J. Oda, Y. Yamada, F. Sato, J. Mol. Biol. 286 (1999) 1137–1145.
- [130] V.S.M. Ho, J.H. Wong, T.B. Ng, Peptides 28 (2007) 760–766.
- [131] R. Ghosh, C. Chakrabarti, Planta 228 (2008) 883–890.
- [132] Y. Looze, P. Boussard, J. Huet, G. Vandenbussche, V. Raussens, R. Wintjens, Phytochemistry 70 (2009) 970–978.
- [133] P. Leone, L. Menu-Bouaouiche, W.J. Peumans, F. Payan, A. Barre, A. Rousset, E.J.M. Van Damme, P. Rougé, Biochimie 88 (2006) 45–52.
- [134] W. Ubhayasekera, R. Rawat, S.W.T. Ho, M. Wiweger, S. Von Arnold, M.-L. Chye, S.L. Mowbray, Plant Mol. Biol. 71 (2009) 277–289.
- [135] W. Ubhayasekera, C.M. Tang, S.W.T. Ho, G. Berglund, T. Bergfors, M.-L. Chye, S.L. Mowbray, FEBS J. 274 (2007) 3695–3703.
- [136] J. Huet, J. Wyckmans, R. Wintjens, P. Boussard, V. Raussens, G. Vandenbussche, J.M. Ruyschaert, M. Azarkan, Y. Looze, Cell. Mol. Life Sci. 63 (2006) 3042–3054.
- [137] R.J. Falconer, M. Marangon, S.C. Van Sluyter, K.A. Neilson, C. Chan, E.J. Waters, J. Agric. Food Chem. (2009).

- [138] G. Marzban, A. Herndl, S. Pietrozotto, S. Banerjee, C. Obinger, F. Maghuly, R. Hahn, D. Boscia, H. Katinger, M. Laimer, *Food Chem.* 112 (2009) 803–811.
- [139] M.-W. Shi, B.-Z. Wang, X.-L. Cheng, J.-H. Wang, Q.-L.L. Yangrui, *Proc. 2nd International Conference on Bioinformatics and Biomedical Engineering (iCBBE)*, 2008, pp. 88–92.
- [140] J. Luckner, M. Laszczak, D. Smith, S. Lund, *BMC Genom.* 10 (2009) 50.
- [141] S.P. Robinson, A.K. Jacobs, I.B. Dry, *Plant Physiol.* 114 (1997) 771–778.
- [142] D.B. Tattersall, R. Van Heeswijck, P.B. Høj, *Plant Physiol.* 114 (1997) 759–769.

## 2 Matériels et méthodes

Cette section présente les différentes techniques utilisées pour la purification des protéines TL et chitinase à partir de jus de raisin et le suivi de leur concentration du jus au Champagne. Les principes des techniques y sont simplement rappelés, le détail des protocoles étant développé dans la section 3.1 [52]. L'analyse par tensiométrie des protéines purifiées est présentée en dernier lieu.

### 2.1 Chromatographie liquide

Dans le cadre de la purification et de l'analyse de protéines, la chromatographie liquide est l'une des méthodes les plus répandues depuis les années 1970. Elle peut se décliner sous trois variantes : HPLC (High Performance Liquid Chromatography), UHPLC (Ultra ~) et FPLC (Fast Protein Liquid Chromatography). Le principe général reste le même pour ces trois techniques (Figure 2) et les différences d'appareillage résultent essentiellement du type de colonne utilisé, selon l'application considérée (Tableau 1).

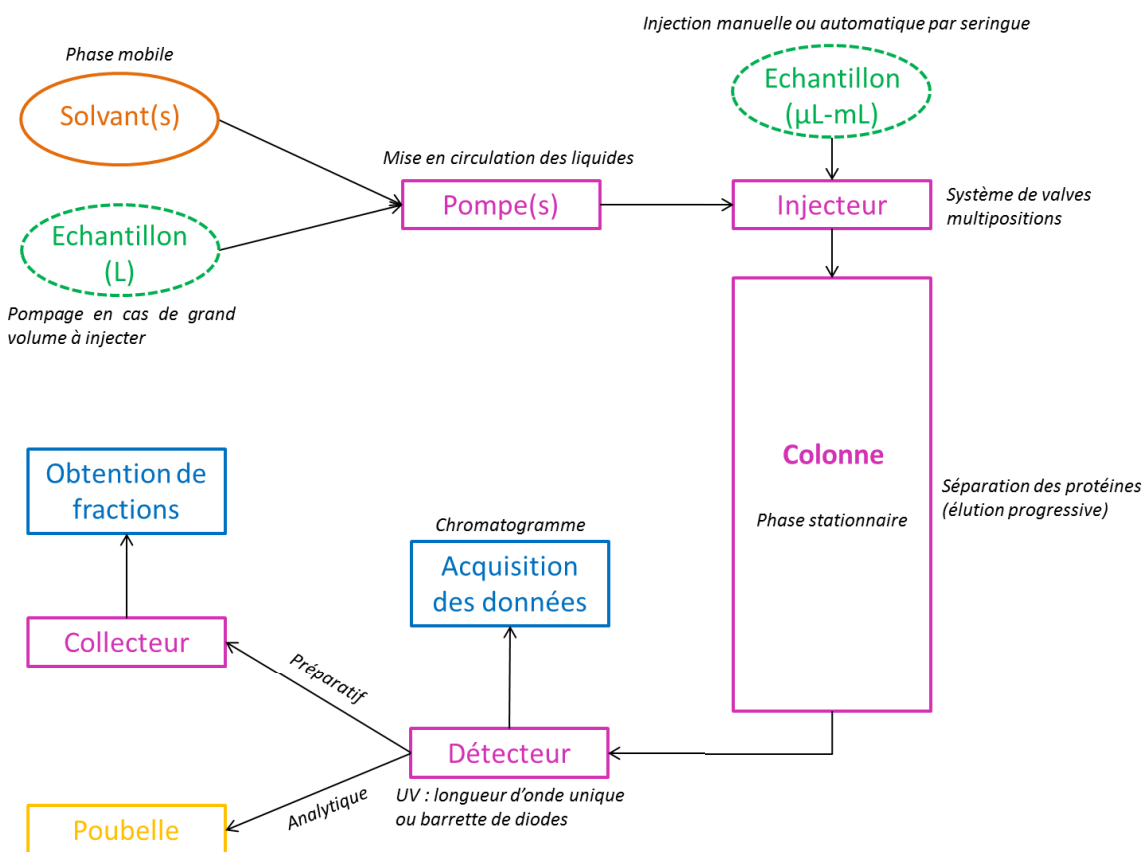


Figure 2 : Représentation schématique du fonctionnement des techniques de chromatographie liquide pour les protéines.

Chromatographie liquide	Premier système commercial	Objectif visé	Colonnes	Implications sur l'appareillage / Spécificités vs HPLC
<b>FPLC</b>	1982 (Pharmacia®, devenu GE Healthcare®)	Purifier des protéines à grande échelle	Très grande variété de gels disponibles. Toutes tailles de colonne (du mm au m). Prépackées ou design à façon.	Basse pression (jusqu'à ~ 10 bars). Détecteur UV pour le suivi des protéines et conductimètre pour le suivi du gradient d'éluion. Possibilité de chargement sur la colonne d'un grand volume d'échantillon (> L) par une pompe (idem solvant ou annexe) vs seringue dans l'injecteur HPLC/UHPLC. Un protocole de purification (injection, éluion de la fraction non-retenue, éluion progressive des protéines et collecte) s'évalue en heures voire en jours vs quelques dizaines de minutes en HPLC et quelques minutes en UHPLC.
<b>UHPLC</b>	2005 (Waters®)	Séparer des composés le plus rapidement possible	Granulométrie < 2 µm (vs ~ 5 µm HPLC) Longueur de 3 à 15 cm (vs ~ 25 cm HPLC) Diamètre semblable ou inférieur (de 2 à 4.6 mm)	Augmentation de la tenue en pression du système (pompe, tubulure et injecteur) : pression applicable jusqu'à ~ 1000 bars (vs 300-400 bars en HPLC). Pas de pression après la colonne donc détecteur UV similaire (mais fréquence d'acquisition plus grande pour les analyses ultra-rapides). Débit liquide légèrement supérieur en vitesse linéaire. Economie de solvant car temps d'analyse fortement diminué. Augmentation de la résolution des pics chromatographiques.

Tableau 1 : Caractéristiques des techniques FPLC et UHPLC en comparaison avec l'HPLC.

La chromatographie liquide, qu'elle soit utilisée dans un but préparatif ou analytique FPLC ou (U)HPLC, permet de séparer les protéines d'un mélange selon leurs différentes propriétés tel que résumé dans le Tableau 2.

Propriétés des protéines	Type de chromatographie
Charge	Echange d'ions
Hydrophobicité	Interaction hydrophobe Phase inverse
Taille	Gel filtration
Reconnaissance spécifique d'un ligand	Affinité
Point isoélectrique	Chromatofocusing

Tableau 2 : Propriétés des protéines selon lesquelles la séparation est effectuée par les principales techniques de chromatographie liquide.

Ces différents types de chromatographie ont été testés au laboratoire pour la purification des protéines de moût par FPLC (Äkta Purifier, GE Healthcare®). Une approche croisée échange de cations puis interaction hydrophobe (HIC) adaptée d'un protocole existant [53] a finalement été retenue, suivie d'une purification supplémentaire par phase inverse (RP) HPLC (Tableau 3). L'éluion des protéines a été suivie par un détecteur UV à 280 nm, maximum d'absorption des groupements tryptophanes et tyrosines. Les principes de ces trois techniques sont présentés dans le Tableau 4. La quantification des protéines TL et chitinase dans différents moûts et Champagnes à partir de gammes de protéines purifiées a été réalisée par UHPLC (UltiMate 3000 RLSC Dionex®, appareil arrivé au laboratoire après le travail de purification) sur une colonne similaire à celle utilisée pour la purification par HPLC après transfert de la méthode.


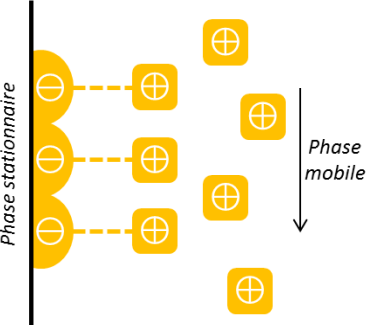
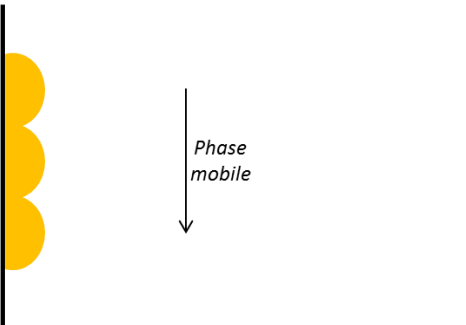
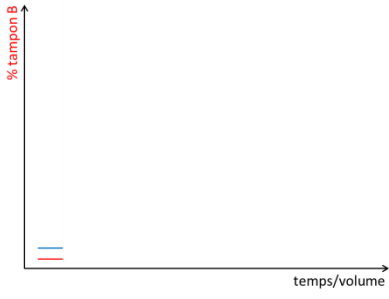
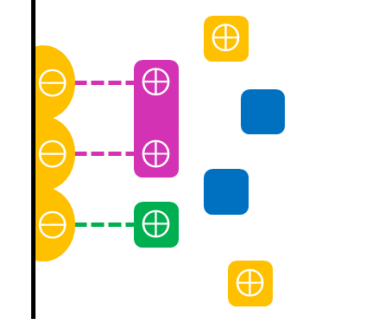
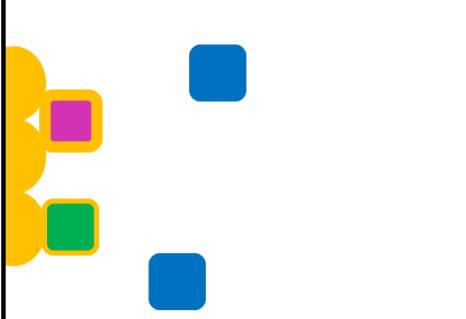
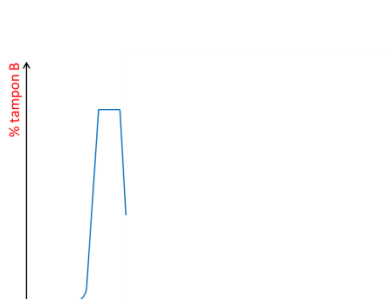
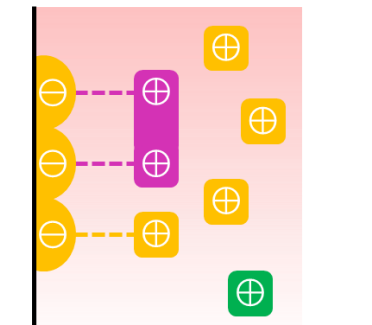
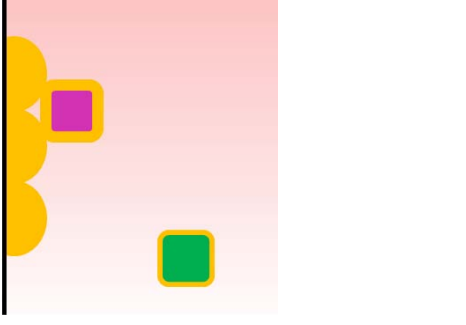
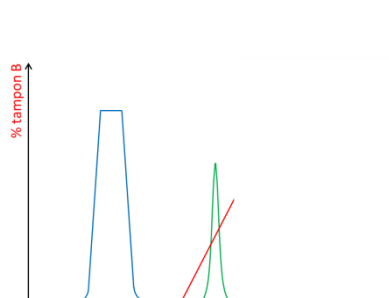
Colonne	Phase stationnaire	Phase mobile
Echange de cation (FPLC #1)	<b>Macro-Prep High S</b> (Bio-rad®) Support : cation fort Ligand fonctionnel : $-\text{SO}_3^-$ Contre-ion : $-\text{Na}^+$ Gel : 240 mL ( $\varnothing$ 50 mm)	<b>Tampon A</b> : 30 mM citrate de sodium, pH 3.0 <b>Tampon B</b> : 30 mM MES ( <i>acide 2-MorpholinoEthaneSulfonique</i> ) / 1 M NaCl, pH 6.0
Interaction hydrophobe (FPLC #2)	<b>Phenyl Sepharose High Performance</b> (GE Healthcare®) Ligand fonctionnel :  Matrice : agarose Gel : 127 mL ( $\varnothing$ 26 mm)	<b>Tampon A</b> : 50 mM citrate de sodium / 1.25 M sulfate d'ammonium, pH 5.0 <b>Tampon B</b> : 50 mM citrate de sodium, pH 5.0
Phase inverse (purification HPLC / quantification UHPLC)	<b>HPLC</b> C <sub>18</sub> Nucleosil 300–5 (4.6×250 mm, 5 $\mu\text{m}$ ) Macherey-Nagel® <b>UHPLC</b> ProSwift RP-1S (4.6×50 mm, 1.1 $\mu\text{m}$ ) Dionex®	<b>Tampon A</b> : 8% acétonitrile, 0.1% acide trifluoroacétique <b>Tampon B</b> : 80% acétonitrile, 0.1% acide trifluoroacétique

Tableau 3 : Conditions opératoires des chromatographies d'échange de cations, d'interaction hydrophobe et en phase inverse utilisées pour la purification et la quantification des protéines thaumatin-like et chitinase. Les colonnes sont équilibrées dans le tampon A puis les protéines sont progressivement éluées avec un gradient croissant de tampon B.

Etape	Echange de cations (FPLC #1)	Interaction hydrophobe (FPLC #2) Phase inverse (purification HPLC / quantification UHPLC)	Chromatogramme (UV)
Equilibration			
Injection de l'échantillon	 <p>Les protéines chargées ++ et + sont retenues, les protéines – ou neutres et les autres composés sont élués.</p>	 <p>Les protéines <b>fortement</b> et <b>faiblement</b> hydrophobes sont retenues, les autres protéines et composés sont élués.</p>	
Elution 1	 <p>Faible augmentation de la force ionique du tampon, élution des protéines les moins liées.</p>	 <p>Elution des protéines les moins liées par un gradient décroissant de sel (HIC) ou une diminution de la polarité du solvant (RP).</p>	



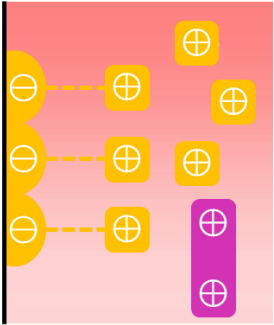
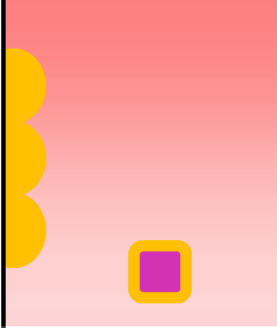
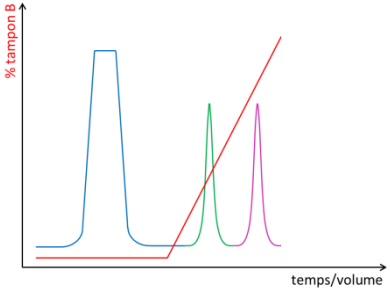
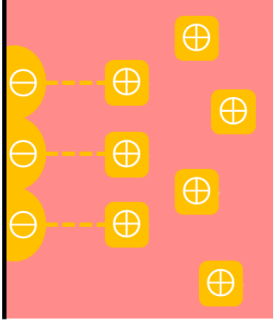
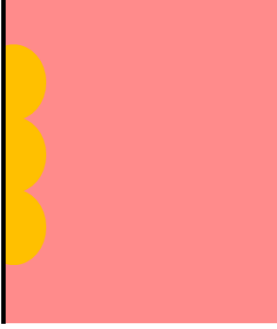
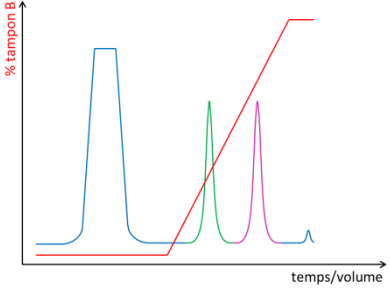
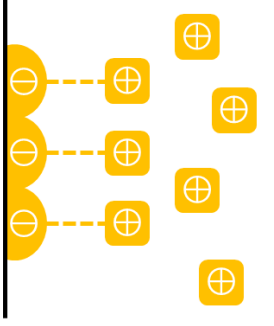
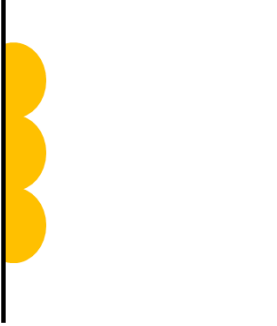
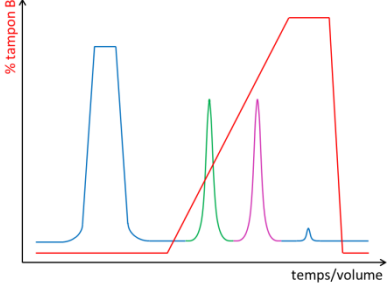
<p><b>Elution 2</b></p>	 <p>Forte augmentation de la force ionique du tampon, élution des protéines les plus liées.</p>	 <p>Elution successive des protéines les plus liées.</p>	
<p><b>Lavage</b></p>			
<p><b>Ré- équilibrage</b></p>			

Tableau 4 : Principe des trois techniques de chromatographie liquide utilisées pour la purification et la quantification des protéines TL et chitinase.

## 2.2 Electrophorèse des protéines

Le principe de l'électrophorèse consiste à soumettre un mélange de molécules chargées à un champ électrique, induisant la migration des molécules vers le pôle opposé à leur charge. La vitesse de cette migration sur un support dépend des caractéristiques de chaque molécule (charge, masse, forme...), ce qui permet d'obtenir une séparation du mélange considéré.

Dans le cas des protéines, l'électrophorèse peut être employée de manière analytique ou plus rarement préparative dans un grand nombre d'applications, de la simple observation de la composition en protéines d'un échantillon jusqu'à l'approche protéomique. Dans le cadre de cette étude, deux contextes de mise en place de l'électrophorèse sont présentés : pour le suivi de purification des protéines TL et chitinase et la visualisation de la répartition en protéines de différents moûts et vins d'une part, son prolongement dans le cadre d'analyses Western blot avec l'utilisation d'anticorps spécifiques anti-TL et anti-chitinase pour la révélation des protéines d'autre part.

### 2.2.1 Principe de l'électrophorèse SDS-PAGE

La technique SDS-PAGE (Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis) permet de réaliser la séparation verticale des protéines d'un mélange en conditions dénaturantes [54]. Le SDS est un détergent anionique (Figure 3), dont la queue hydrophobe s'enroule autour du squelette des protéines. La migration des protéines, alors toutes chargées négativement, dépend ainsi uniquement de leur encombrement moléculaire. Les structures tertiaires et secondaires étant perdues lors de l'addition de SDS (les liaisons hydrogène et les interactions hydrophobes stabilisant les structures sont perdues et les charges négatives enveloppées autour du squelette se repoussent, donc permettent de déplier la protéine), les différences de forme des protéines sont minimisées et leur encombrement moléculaire peut être relié à leur masse moléculaire. Les protéines ainsi dénaturées et chargées négativement sont séparées par un gel réticulé formant un tamis moléculaire. Le gel est obtenu par polymérisation d'acrylamide (création des chaînes) et de bisacrylamide (pontage des chaînes entre elles) (Figure 3). Trivialement, plus le pourcentage d'acrylamide dans la composition du gel est élevé, plus les mailles du réseau sont serrées. Avec l'application d'un courant électrique, les petites protéines chargées négativement migrent plus rapidement vers l'anode (pôle  $\oplus$ ), les grosses étant retenues en haut du gel (Figure 4). Après la séparation des protéines, le gel peut être directement coloré afin d'observer chaque bande, ou transféré sur une membrane dans le cas d'une expérience Western Blot.

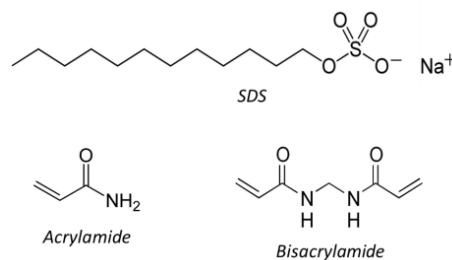


Figure 3 : Formules du SDS, de l'acrylamide et du bisacrylamide utilisés dans le cadre d'une SDS-PAGE.

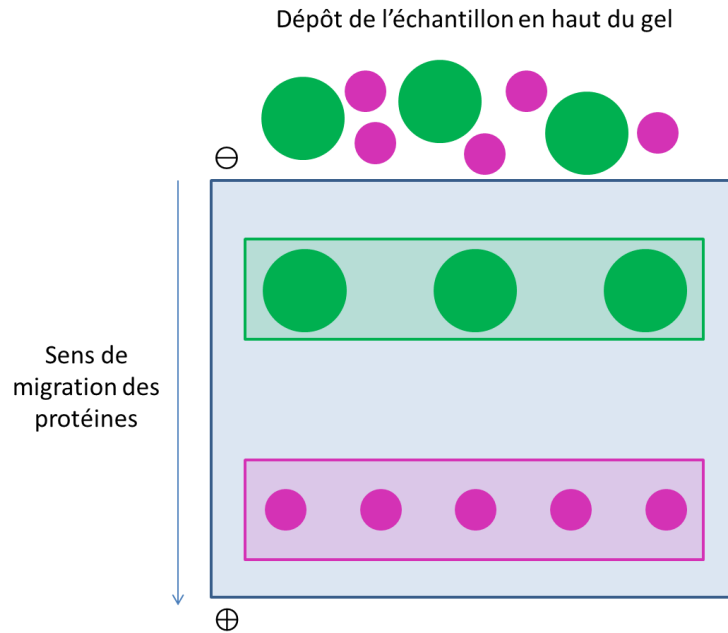


Figure 4 : Principe d'une SDS-PAGE. Un échantillon composé de deux protéines (haut poids moléculaire en vert, faible en rose) est déposé en haut du gel. Les protéines, qui sont toutes chargées négativement par le SDS, migrent vers l'anode et sont séparées selon leur poids moléculaire grâce au réseau d'acrylamide, formant deux bandes distinctes sur le gel.

### 2.2.2 Détection des bandes

Après la migration, les différentes bandes peuvent être directement observées par coloration des protéines sur le gel. Deux types de révélation ont été utilisés dans cette étude : le bleu de Coomassie [55], dont la limite de détection est d'environ 0.1  $\mu\text{g}$  de protéine par bande, et une procédure au nitrate d'argent [56], similaire au développement de photographies argentiques, environ 100 fois plus sensible mais bien plus lourde à mettre en place. Dans les deux cas, toutes les protéines sont colorées et l'intensité des bandes dépend de leur concentration et de leur affinité avec le colorant (Figure 5).

Dans une autre approche, les bandes de protéines peuvent être sélectivement révélées par l'utilisation d'un protocole Western blot (Figure 5) [57]. Après la migration, les protéines séparées sur un gel sont transférées sur une membrane de nitrocellulose. Cette membrane est ensuite successivement incubée avec une solution d'anticorps primaire (qui se fixe de manière spécifique sur la protéine à détecter) puis une solution d'anticorps secondaire (spécifique de l'espèce animale utilisée pour l'obtention de l'anticorps primaire). Dans cette étude, les protéines purifiées à partir de jus de raisin ont été injectées à des lapins pour obtenir des anticorps primaires spécifiques anti-TL et anti-chitinase. Un anticorps secondaire « anti-lapin » lié à une enzyme (peroxydase de raifort) qui réagit à l'addition d'un substrat (4-chloro-1-naphthol) a permis la formation d'une zone colorée visible sur la membrane.

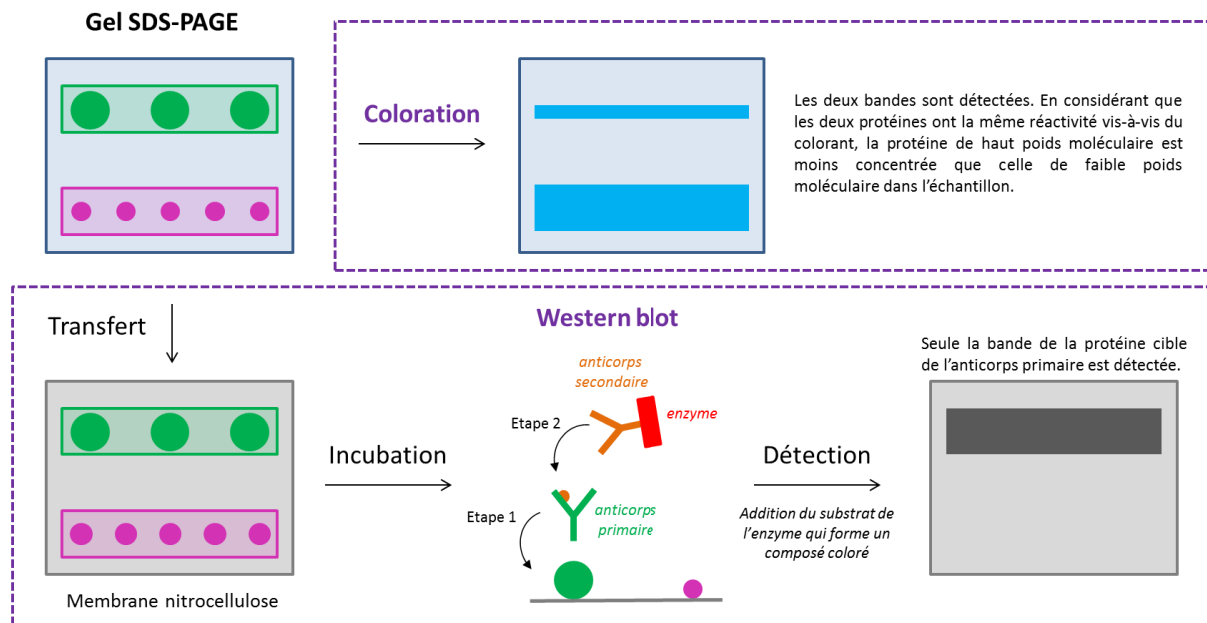


Figure 5 : Représentation schématique des principales étapes après un gel SDS-PAGE : visualisation directe après coloration et immunodétection par Western blot. Toutes les manipulations intermédiaires (fixation, blocage, lavages...) sont omises.

### 2.3 RMN

L'éventuelle présence de contaminants organiques dans les fractions de TL et chitinase en cours de purification a été détectée par spectrométrie de Résonance Magnétique Nucléaire (RMN). Les échantillons ont été préparés par dissolution de chaque fraction lyophilisée dans 600  $\mu\text{L}$  d'un mélange 90 %  $\text{H}_2\text{O}$  / 10%  $\text{D}_2\text{O}$  puis ajustés à pH 6.0. Les spectres de RMN du  $^1\text{H}$  ont été obtenus à 298 K sur un spectromètre Bruker Avance III 500 MHz équipé d'une sonde BBFO+ (BroadBand Fluorine Observation). Le principe de la technique RMN est développé dans le cadre de la Partie 2, section 2.2.4.

### 2.4 ELISA

Un dosage immuno-enzymatique ELISA (Enzyme-Linked ImmunoSorbent Assay) fonctionne sur le même principe qu'une expérience Western blot, avec l'incubation successive par un anticorps primaire spécifique de la protéine à doser et un anticorps secondaire fixé à une enzyme capable de réagir avec un substrat pour former un composé coloré (Figure 6). Cependant, l'échantillon contenant le mélange de protéines n'est ici plus séparé préalablement sur un gel mais directement déposé dans une cuve. Les anticorps sont appliqués après fixation de toutes les protéines sur les parois de la cuve. Le protocole du dosage a été adapté selon des travaux antérieurs sur l'invertase de raisin [58]. Après réalisation de gammes des protéines purifiées, la quantification des protéines TL et chitinase dans différents jus et Champagnes est réalisée par mesure de l'absorbance dans la cuve à 490 nm après incubation par les mêmes anticorps que ceux utilisés pour l'expérience Western Blot. Le substrat de l'enzyme fixée à l'anticorps secondaire a été réalisé selon le mélange suivant : 25 mL tampon phosphate-citrate 0.1 M pH 5.0, 6  $\mu\text{L}$   $\text{H}_2\text{O}_2$ , 10 mg ortho-phényldiamine.

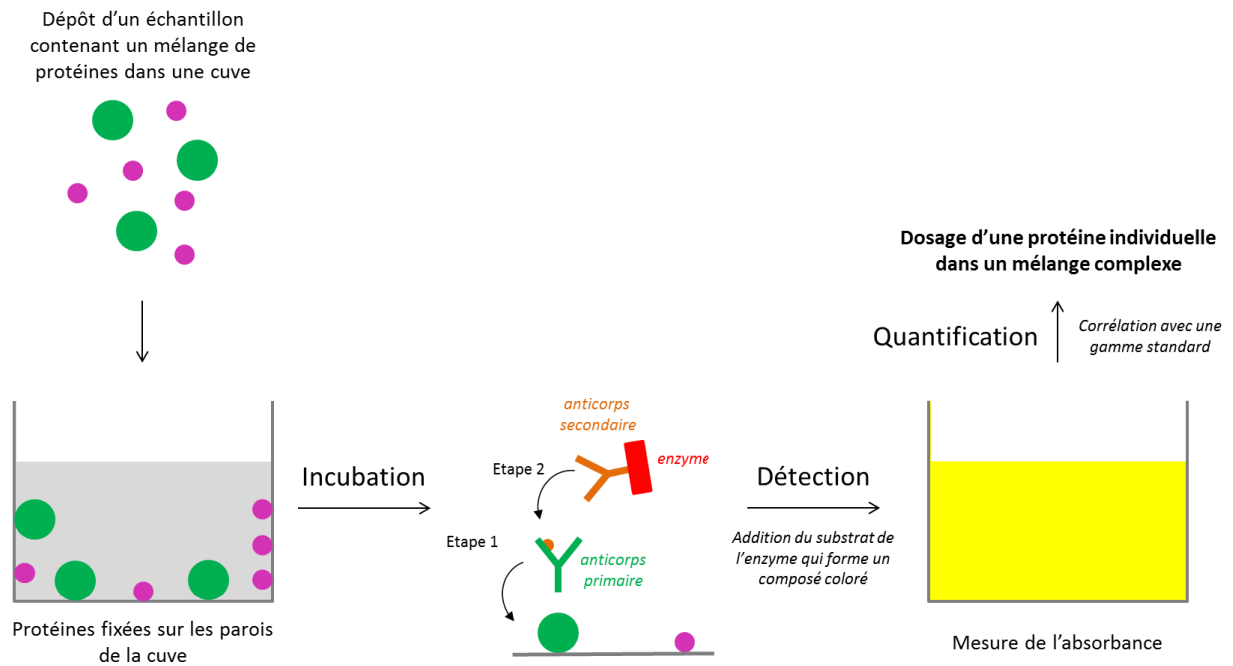


Figure 6 : Principales étapes d'un protocole ELISA. Toutes les manipulations intermédiaires (fixation, blocage, lavages...) sont omises.

## 2.5 Tensiométrie à lame de Wilhelmy

### 2.5.1 Notion de pression de surface

A l'interface entre un liquide et un gaz, il existe une dissymétrie des forces d'interactions attractives et répulsives entre les molécules. Tout accroissement de surface  $\delta A$  de l'aire d'une interface induit une variation d'énergie libre de surface  $\delta F$ , proportionnelle à  $\delta A$ . Par définition, la constante de proportionnalité  $\gamma$  est appelée tension superficielle [59]. Elle peut s'exprimer en  $J.m^{-2}$  ou en  $N.m^{-1}$ .

Lorsque des molécules tensioactives comme les protéines sont introduites dans une solution aqueuse, elles vont s'adsorber à l'interface air/liquide pour réduire le contact entre leur partie hydrophobe et l'eau. En conséquence, certaines molécules d'eau interfaciales vont être remplacées par des molécules de tensioactif induisant un abaissement de la tension superficielle. Il est alors possible de définir la notion de pression de surface  $\Pi$  comme étant la différence des tensions superficielles de la sous-phase seule ( $\gamma_0$ ) et de la sous-phase couverte de tensioactifs ( $\gamma$ ) :

$$\Pi = \gamma_0 - \gamma$$

### 2.5.2 Protocole expérimental

La mesure de la pression de surface  $\Pi$  statique peut être déterminée par la méthode de la lame de Wilhelmy [60]. Lorsqu'une lame mince est plongée dans un liquide, la tension de surface se manifeste par l'apparition d'un ménisque le long de son périmètre. La tension superficielle est déterminée en mesurant la traction au moment où la lame se détache de la surface selon :

$$\Pi = \frac{\Delta W}{2(L + d) \cos \theta}$$

avec  $\Delta W$  la variation du poids de la plaque au moment de son contact avec le liquide,  $L$  et  $d$  les dimensions de la plaque et  $\theta$  l'angle de mouillage. En cas de mouillage parfait ( $\theta = 0^\circ$ ), la variation du poids de la plaque mesurée par le tensiomètre permet donc d'accéder directement à la tension de surface (Figure 7).

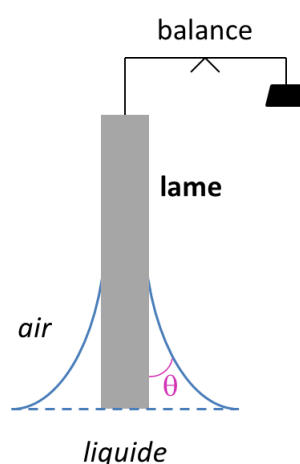


Figure 7 : Mesure de la pression de surface par la méthode de Wilhelmy.

Afin de minimiser la quantité de protéines purifiées utilisée dans le cadre de l'étude des propriétés tensioactives des protéines TL et chitinase, la lame a été remplacée par un cylindre en platine de périmètre 6.283 mm pour permettre l'utilisation d'une mini-cuve sans apparition d'effets de bord [61]. Les mesures ont ainsi été effectuées sur un tensiomètre K11 Krüss® dans des cuves de 4 mL à 20°C. Une mesure de pression de surface a été effectuée toutes les 10 s pendant une durée totale de 300 s. La valeur moyenne retenue a été calculée en prenant en compte uniquement les dix dernières mesures.

## 3 Résultats et discussion

### 3.1 Purification et quantification des protéines thaumatin-like et chitinase du jus de raisin au Champagne

Avant d'évaluer les propriétés des protéines TL et chitinase et d'étudier leur rôle potentiel dans la formation et la stabilisation d'interfaces gaz/liquide, il apparaissait indispensable de déterminer leur concentration dans un Champagne. La difficulté majeure rencontrée lors du dosage d'une protéine individuelle dans un mélange complexe provient de la condition pré-requise d'obtention de ladite protéine pure pouvant servir de standard dans une gamme d'étalonnage. Les isoformes majoritaires de TL et chitinase ont donc été en premier lieu purifiées à partir de jus de raisin (moût) par trois étapes successives de chromatographie liquide, les deux premières ayant été adaptées d'un protocole existant [53]. Lors de la première étape, le jus de raisin Chardonnay (20 L) a été directement chargé sur une colonne d'échange de cations (FPLC) équilibrée à pH 3.0, proche de celui du jus. Le point isoélectrique des protéines TL et chitinase étant situé entre 4 et 5, elles sont donc chargées positivement à ce pH et retenues sur la colonne. La composition de chacune des fractions collectées a ensuite été analysée par SDS-PAGE afin de détecter celles contenant majoritairement les protéines d'intérêt. Une seconde étape FPLC sur une colonne d'interaction hydrophobe a permis une purification croisée et l'élimination des protéines minoritaires. A l'issue de ce protocole, la pureté des fractions de TL et chitinase a été analysée par spectrométrie de masse (prestation de service, Université de Strasbourg) et spectrométrie RMN (collaboration avec le laboratoire ICMR de l'URCA) révélant à la fois la présence de traces d'autres protéines de raisin et la présence d'un contaminant organique, l'acide citrique, provenant certainement des tampons utilisés en FPLC. Une étape supplémentaire de purification par RP-HPLC a alors été mise en place afin d'éliminer tous ces contaminants. Les fractions pures lyophilisées ont enfin permis la réalisation de gammes standards par pesée de masse exacte pour la mise en place des méthodes de quantification par chromatographie liquide et dosage immuno-enzymatique. Dans le cas de l'UHPLC, différents échantillons de jus de raisin, vin de base et Champagnes ont été injectés sur une colonne C18 et la quantification a été effectuée par comparaison des aires des pics de TL et chitinase par rapport aux gammes standards. Les protéines purifiées ont d'autre part permis la réalisation d'anticorps anti-TL et anti-chitinase, dont la spécificité a été vérifiée par Western Blot. Un protocole ELISA a ainsi pu être également appliqué pour la détermination de la concentration des protéines dans les mêmes échantillons. Les deux méthodes, basées sur des principes totalement différents, ont donné des résultats très cohérents (moins de 16% d'écart entre deux valeurs). Le suivi des protéines TL et chitinase au cours du procédé de vinification a permis d'estimer la diminution de leur concentration de l'ordre d'un facteur 100 (autour de 100 mg/L dans un jus, quelques mg/L dans un Champagne), confirmant l'observation grossière réalisée sur l'intensité relative des bandes lors de l'analyse SDS-PAGE des échantillons de jus et de Champagnes.

Les résultats de cette partie sont présentés dans l'article suivant, publié en 2011 dans le journal *Analytical and Bioanalytical Chemistry* sous le titre « Quantification of chitinase and thaumatin-like proteins in grape juices and wines » [52].





# Quantification of chitinase and thaumatin-like proteins in grape juices and wines

D. Le Bourse · A. Conreux · S. Villaume · P. Lameiras · J.-M. Nuzillard · P. Jeandet

Received: 18 February 2011 / Revised: 10 March 2011 / Accepted: 14 March 2011 / Published online: 5 April 2011  
© Springer-Verlag 2011

**Abstract** Chitinases and thaumatin-like proteins are important grape proteins as they have a great influence on wine quality. The quantification of these proteins in grape juices and wines, along with their purification, is therefore crucial to study their intrinsic characteristics and the exact role they play in wines. The main isoforms of these two proteins from Chardonnay grape juice were thus purified by liquid chromatography. Two fast protein liquid chromatography (FPLC) steps allowed the fractionation and purification of the juice proteins, using cation exchange and hydrophobic interaction media. A further high-performance liquid chromatography (HPLC) step was used to achieve higher purity levels. Fraction assessment was achieved by mass spectrometry. Fraction purity was determined by HPLC to detect the presence of protein contaminants, and by nuclear magnetic resonance (NMR) spectroscopy to detect the presence of organic contaminants. Once pure fractions of lyophilized chitinase and thaumatin-like protein were obtained, ultra-HPLC (UHPLC) and enzyme-linked immunosorbent

assay (ELISA) calibration curves were constructed. The quantification of these proteins in different grape juice and wine samples was thus achieved for the first time with both techniques through comparison with the purified protein calibration curve. UHPLC and ELISA showed very consistent results (less than 16% deviation for both proteins) and either could be considered to provide an accurate and reliable quantification of proteins in the oenology field.

**Keywords** Wine · Chromatography · Protein · ELISA · Quantification · Purification

## Abbreviations

BSA	bovine serum albumin
Cyt C	horse heart cytochrome C
FPLC	fast protein liquid chromatography
HIC	hydrophobic interaction
HPLC	high-performance liquid chromatography
MES	2-( <i>N</i> -morpholino)ethane sulfonic acid
NMR	nuclear magnetic resonance
PBS	phosphate-buffered saline
PR	pathogenesis-related
PTA-ELISA	plate-trapping antigen–enzyme-linked immunosorbent assay
RP	reversed-phase
RSD	relative standard deviation
SCX	strong cation exchange
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
TBS	Tris-buffered saline
TFA	trifluoroacetic acid
TL	thaumatin-like
UHPLC	ultra-HPLC

Published in the special issue *Wine Analysis* with Guest Editor Isabelle Pianet.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00216-011-4912-8) contains supplementary material, which is available to authorized users.

D. Le Bourse (✉) · A. Conreux · S. Villaume · P. Jeandet  
Laboratoire d'Oenologie et Chimie Appliquée, Université de Reims Champagne-Ardenne,  
URVVC-SE UPRES EA 2069, BP 1039,  
51687 Reims Cedex 2, France  
e-mail: doriane.le-bourse@univ-reims.fr

P. Lameiras · J.-M. Nuzillard  
CNRS, Institut de Chimie Moléculaire de Reims,  
BP 1039, 51687 Reims Cedex 2, France

## Introduction

Proteins play a leading role in the wine industry concern. They exhibit a wide range of properties and are involved in a number of aspects linked to wine quality and its acceptance by a consumer [1]. Their quantification is thus crucial, but major drawbacks when dealing with grape juice and wine samples can impair the use of available techniques, such as the presence of contaminants which can distort the measurement and the absence of standard grape or wine proteins [1].

Chitinases and thaumatin-like (TL) proteins are some of the most abundant proteins in grapes (*Vitis vinifera*) [2, 3]. They are classified as pathogenesis-related (PR) proteins, although they are expressed at high levels in berries regardless of pathogen attack [3–5]. They have been shown to be resistant to yeast proteolysis and can thus persist in wines [6]. The presence of chitinase and TL protein in white wines can induce the well-known haze formation phenomenon, attributed to the protein aggregation during bottle storage [7–11]. These two protein families are also alleged to have an influence on foaming properties of sparkling wines and Champagne [12, 13] but no evidence of the role they play has yet been demonstrated.

Before elucidating the chitinases' and TL proteins' properties and the phenomena they are involved in, one must purify them and determine their exact concentration in grape juice and wine. Marangon et al. [14] fractionated the whole set of proteins from grape juice, revealing the wide diversity of isoforms of chitinase and TL proteins. In a subsequent study [15], isoforms of each protein family were isolated and purified by fast protein liquid chromatography (FPLC). The quantification of the different chitinases and TL proteins was then achieved by reversed-phase (RP)-high-performance liquid chromatography (HPLC) through comparison of peak areas with those of bovine serum albumin (BSA) and horse heart cytochrome C (Cyt C) used as standards. To our knowledge, no direct quantification using pure grape proteins has yet been developed.

The aim of this work was thus to provide a reliable quantification of the main isoforms of chitinases and TL proteins in grape juices and wines. For that purpose, these two proteins were purified from grape juice using two successive FPLC steps according to the previous work developed by Van Sluyter et al. [15] with modifications, followed by an extra preparative HPLC clean-up step. Pure proteins were then used as standards for quantification in grape juice and wine samples, achieved by ultra-HPLC (UHPLC) and plate-trapping antigen–enzyme-linked immunosorbent assay (PTA-ELISA).

## Materials and methods

### Chemicals

Tris(hydroxymethyl)aminomethane (Tris), chloridric acid, tetramethylethylenediamine, ammonium persulfate, silver nitrate, ammonium sulfate, methanol, BSA, goat anti-rabbit IgG peroxidase-conjugated antibody, acetonitrile, hydrogen peroxide, sodium chloride, and sodium hydroxide were purchased from Sigma–Aldrich (Steinheim, Germany). Formaldehyde, sodium dodecyl sulfate, citric acid trisodium salt, citric acid, 2-(*N*-morpholino)ethane sulfonic acid (MES), and trifluoroacetic acid (TFA) were obtained from Acros Organics (Geel, Belgium). Acetic acid and ethanol were respectively purchased from Fisher Scientific (Leicestershire, UK) and Charbonneaux-Brabant (Reims, France). Acrylamide solution was obtained from Bio-Rad (Hercules, USA), phosphate-buffered saline (PBS) from Biosolve (Valkenswaard, Netherlands), and skim milk powder from Merck (Darmstadt, Germany). Deuterium oxide was purchased from Euriso-Top (Gif-sur-Yvette, France). Ultra-pure water produced from a Milli-Q Water System™ (Millipore, Guyancourt, France) was used for the preparation of the buffers and solvents. The two FPLC chromatography columns were respectively packed with MacroPrep High S resin (Bio-Rad, Hercules, USA) and Phenyl Sepharose HP medium (GE Healthcare, Uppsala, Sweden).

### Plant material and juice fractions

For the protein purification, 20 L of grape juice was collected from the Nogent l'Abesse wine cooperative. Grape berries of *Vitis vinifera* cv. Chardonnay were hand-harvested in the Champagne area (France), in September 2009. Grapes were pressed with a pneumatic press, applying a pressure between 1.5 and 2 bar. Sulfur dioxide ( $7 \text{ g hL}^{-1}$ ) was added to the free run juice and the grape juice was collected after static settling for 24 h and stored at 4 °C until use.

For the protein quantification, five different grape juice and wine samples from Nogent l'Abesse wine cooperative were used. Two grape juices from two different grape varieties (Chardonnay and Pinot Meunier) harvested in 2009 were compared. Both grape juices were obtained and collected in the same way as described for the 20 L of grape juice dedicated for the protein purification. One base wine and two commercial Champagne wines from the Chardonnay variety were also analyzed. The base wine was made from grape juice harvested in 2008. This grape juice was racked and chaptalized with sucrose. It then underwent alcoholic fermentation (achieved by adding the *Saccharomyces cerevisiae* yeast at 18 °C) and was collected. The two

Champagne wines were elaborated from Chardonnay grapes harvested in 2006 and 2008 according to the “méthode traditionnelle,” in which the alcoholic fermentation is followed by a second fermentation taking place in bottles (called “prise de mousse”). The aging period of the Champagne wines was at least 15 months. All the six samples were stored at  $-20\text{ }^{\circ}\text{C}$  to avoid sample denaturation which could occur over time.

#### Electrophoresis of grape juices and wine samples

The five grape juices and wines studied were analyzed with sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE). Vertical SDS-PAGE was performed in a Mini-Protean III electrophoresis cell (Bio-Rad, USA) according to the method of Laemmli [16], using a 4% stacking gel and a 13% resolving gel. A  $6\text{-}\mu\text{L}$  sample was loaded on 1 mm gel without prior treatment. Proteins were separated under a constant voltage of 175 V for 45 min. Standard protein markers (Precision Plus Protein™ Standards, Bio-Rad, Hercules, USA) were simultaneously loaded at each run. After electrophoresis, gels were silver-stained [17].

#### Macromolecule isolation and protein purification

Proteins from Chardonnay grape juice were isolated and purified using the methods reported by Marangon et al. [14] and Van Sluyter et al. [15] with modifications. The determination of the protein composition in the different purified fractions was achieved by SDS-PAGE and western blots. Their purity was assessed by HPLC, mass spectrometry, and nuclear magnetic resonance (NMR) spectroscopy. Two dialyzed and freeze-dried fractions, containing mostly TL protein for one and chitinase for the other, were selected to undergo a further purification step.

This supplementary clean-up step was then performed using preparative HPLC. The two selected dried fractions were redissolved at 5 mg/mL in ultrapure water. These fractions were loaded as successive  $100\text{-}\mu\text{L}$  samples on a C18 column ( $4.6\times 250\text{ mm}$ , Nucleosil 300–5, Macherey-Nagel) at 1 mL/min equilibrated in a mixture of 10% (v/v) solvent A [80% acetonitrile, 0.1% (v/v) TFA] and 90% (v/v) solvent B [8% acetonitrile, 0.1% TFA] and held at  $30\text{ }^{\circ}\text{C}$ . Proteins were eluted by a gradient of solvent A from 10% to 90% in 40 min, and then held at 90% for 10 min before re-equilibrating the column in the starting conditions for an additional time of 10 min. Elution was followed by absorbance at 220 and 280 nm. From the 280-nm chromatogram, the fractions corresponding to the main peaks were collected according to their retention time: from 18.5 min to 20 min for thaumatin-like and from 30 min to 31.5 min for chitinase. These fractions were then dialyzed three times

against water (cut-off size 3 kDa) and freeze-dried. The protein purity assessment was then achieved by HPLC (same as previously described), mass spectrometry, and NMR spectroscopy. Lyophilized fractions were assumed to contain only pure proteins and used to set up calibration curves for UHPLC and ELISA quantifications. Protein storage in the form of lyophilized fractions ensured their stability over time.

#### Protein composition and purity assessment

##### *SDS-PAGE*

The protein composition of fractions at each purification step was assessed by SDS-PAGE as described previously. After electrophoresis, gels were either stained with colloidal Coomassie brilliant blue according to Anderson et al. [18] or electrophoretically transferred to nitrocellulose membranes in a Hoefer Transphor unit (Bio-Rad, Hercules, USA) at 100 V for 1 h at  $4\text{ }^{\circ}\text{C}$ .

##### *Western blots*

After transfer, membranes were blocked with 2% (w/v) skim milk powder (Merck, Germany) in Tris-buffered saline (TBS) overnight and rinsed three times with TBS. They were incubated for 3 h with primary antibody (TL or chitinase) appropriately diluted in TBS containing 1% (w/v) skim milk powder. Polyclonal antibodies were raised in rabbits after injection of pure TL and chitinase lyophilized fractions by the Proteogenix company (France). Membranes were then successively washed in TBS, TBS plus 0.05% (v/v) Tween 20, and TBS, and incubated with a goat anti-rabbit IgG peroxidase-conjugated antibody (Sigma-Aldrich, Steinheim, Germany) (1/1,500 dilution in TBS plus 1% skim milk powder) for 3 h at room temperature, and washed again successively with TBS, TBS plus 0.05% (v/v) Tween 20, and TBS. Peroxidase activity was revealed by staining using 4-chloro-1-naphthol (1.2 g/L) in ice-cold methanol and 0.4%  $\text{H}_2\text{O}_2$  in PBS.

##### *Mass spectrometry*

NanoLC-MS/MS was performed at the University of Strasbourg (France) as described by Jégou et al. [19].

##### *NMR spectroscopy*

Samples were prepared by dissolving 1 to 10 mg/mL of dried protein fractions in a mixture of 90% ultrapure water and 10%  $\text{D}_2\text{O}$  (v/v), adjusted to pH 6.  $^1\text{H}$  NMR spectra were measured at 298 K in a Bruker Avance AVIII500 NMR spectrometer equipped with a 10-A gradient amplifier

and with a 5-mm broadband fluorine observation (BBFO) probe including shielded z-gradients.

### Sample preparation

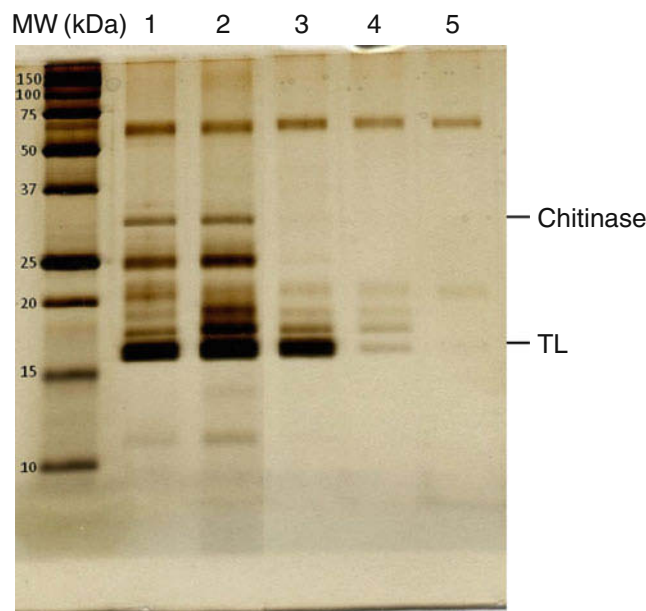
Grape juices and wines were centrifuged at 8,000g for 10 min. Supernatants were separated and filtered through a 0.45- $\mu$ m membrane (Alltech, Lexington, USA). Then 20 mL of filtered juices was dialyzed (cut-off size 3 kDa) and freeze-dried (three independent repetitions). The macromolecular fractions obtained were dissolved in ultrapure water (5 mg/mL), ready to be analyzed for the quantification experiments.

### UHPLC quantification

All the five prepared grape juice and wine samples and purified proteins at appropriate dilution used as standards were loaded (100  $\mu$ L) on a C18 column in a UHPLC system (Dionex, Sunnyvale, USA) using the Chromeleon version 6.80. The C18 column (4.6 $\times$ 250 mm, Nucleosil 300–5, Macherey-Nagel) was equilibrated in a mixture of 10% (v/v) solvent A [80% acetonitrile, 0.1% (v/v) TFA] and 90% (v/v) solvent B [8% acetonitrile, 0.1% TFA] at 1 mL/min and held at 30  $^{\circ}$ C. Proteins were eluted using the following gradient: 10% A for the first 5 min; 35% A in 1 min; 45% A in 5 min (non-linear gradient concave upward curve 8); 55% A in 1 min; 65% A in 5 min (convex upward curve 3); 90% A in 3 min. The column eluent was then held at 90% for 10 min before re-equilibrating in the starting conditions for an additional 10 min. Elution was followed by monitoring absorbance at 280 nm. TL and chitinase peaks were attributed via their retention times, and the protein quantification was achieved through the comparison of the peak areas with those of the purified proteins used as standards. Pure proteins were redissolved in ultrapure water prior to injection. The working range for protein detection was from 0.01 to 0.1 mg/mL. Pure amounts of proteins were also added to juices to assess the recovery and specificity (Electronic Supplementary Material Fig. S1). The range of recovery was 98.7–99.8% with a standard deviation of 0.1–0.5%. No other peak interfered with the TL and chitinase ones, so high method specificity was confirmed. The precision of the method was evaluated by performing three replicate analyses of each juice. The coefficients of variation ranged from 5 to 17 % for both proteins.

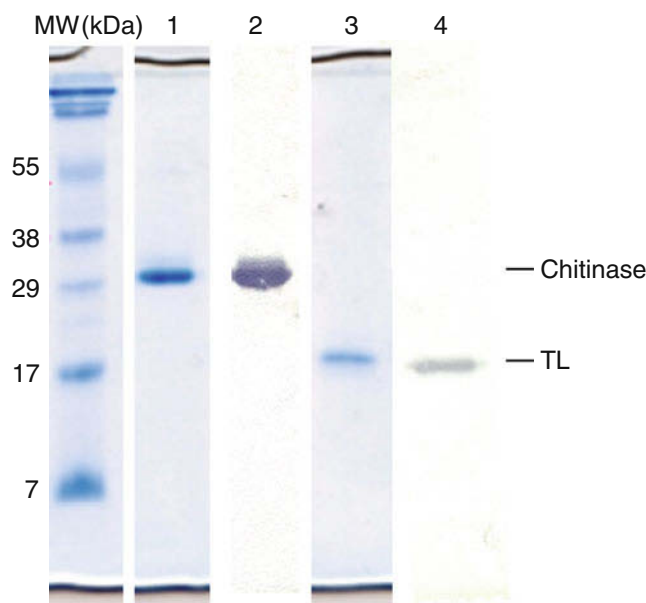
### ELISA

A PTA-ELISA was performed for the quantification of TL and chitinase in grape juice and wine samples using the protocol described by Dambrouck et al. [12] for the



**Fig. 1** Silver-stained SDS-PAGE of crude juice samples: 1 grape juice Chardonnay 2009; 2 grape juice Pinot Meunier 2009; 3 base wine Chardonnay 2008; 4 Champagne Chardonnay 2008; 5 Champagne Chardonnay 2006

quantification of invertase with minor modifications. Thus 96-well flat-bottom immunoplates (Immulon 2HB, Dynex Technologies, Chantilly, USA) were coated with standards or samples diluted (or not) in PBS (100  $\mu$ L/ well). After the plates were covered and left overnight at 4  $^{\circ}$ C, they were washed twice with PBST [PBS containing 0.05% (v/v) Tween 20] and once with PBS. The plates were then



**Fig. 2** F1p2 and F6p2 were analyzed by SDS-PAGE (respectively lanes 1 and 3) and by western blot (respectively lanes 2 and 4) using either anti-chitinase or anti-TL antibodies. F1p2 was thus assessed as the chitinase fraction, and F6p2 as the TL one

**Table 1** Proteins identified by nanoLC-MS/MS from the HIC fractions before (F1p2 and F6p2) and after HPLC purification (F1p2M and F6p2M)

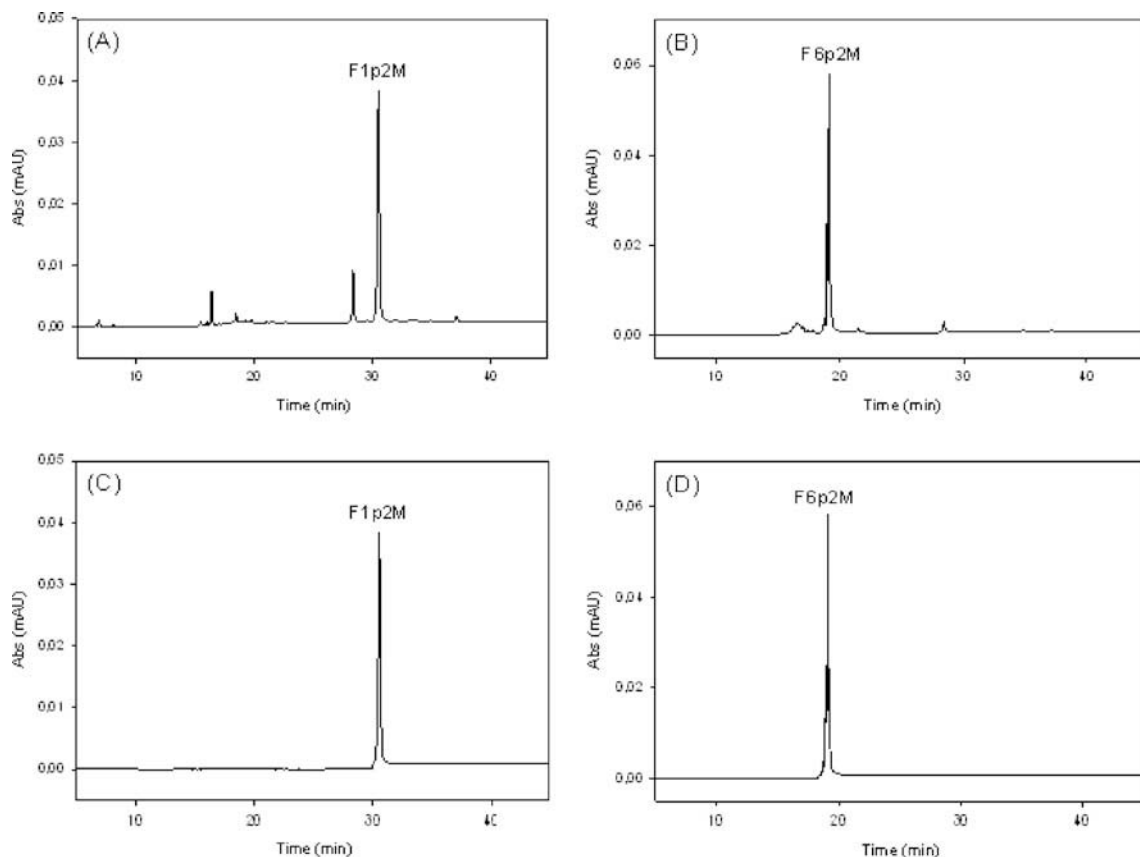
Sample	Protein name (organism)	NCBI accession number <sup>a</sup>	Theoretical molecular weight (kDa)	Sequence coverage (%) <sup>b</sup>	Number of peptides matched <sup>c</sup>
F1p2	Class IV chitinase ( <i>Vitis vinifera</i> )	gi 225434074	27.511	33	6
	PREDICTED: hypothetical protein ( <i>Vitis vinifera</i> )	gi 225426793	23.840	19	4
	Putative thaumatin-like protein ( <i>Vitis vinifera</i> )	gi 7406716	24.035	17	3
F6p2	VVTL1 ( <i>Vitis vinifera</i> )	gi 2213852	23.953	56	7
	Putative thaumatin-like protein ( <i>Vitis vinifera</i> )	gi 7406716	24.035	41	8
	PR-4 type protein ( <i>Vitis vinifera</i> )	gi 3511147	15.218	15	2
	Osmotin-like protein ( <i>Vitis vinifera</i> )	gi 1839046	23.859	13	2
	Class IV endochitinase protein ( <i>Vitis vinifera</i> )	gi 2306811	27.224	7	1
F1p2M	Class IV chitinase ( <i>Vitis vinifera</i> )	gi 225434074	27.511	33	6
F6p2M	VVTL1 ( <i>Vitis vinifera</i> )	gi 2213852	23.953	56	7

NCBI National Center for Biotechnology Information

<sup>a</sup> Protein identification number provided by the NCBI database

<sup>b</sup> Number of peptides identified by the MASCOT search engine that matched with the identified protein

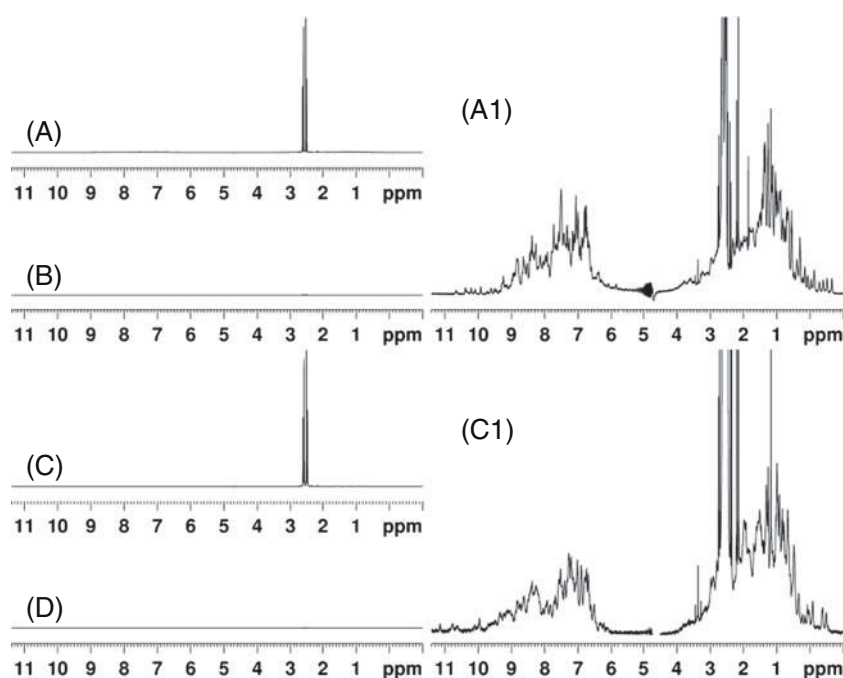
<sup>c</sup> Percentage ratio of all amino acids from valid peptides matched to the total number of amino acids in the reference protein



**Fig. 3** F1p2 and F6p2 were analyzed by HPLC on a C18 column at 280 nm (respectively **a** and **b**). Both chromatograms show small protein contaminant peaks. Fractions corresponding to the main peaks

(F1p2M and F6p2M) were collected and subjected again to HPLC (respectively **c** and **d**). Loads and washes are omitted

**Fig. 4**  $^1\text{H}$  NMR spectra of chitinase and TL fractions before (respectively F1p2, spectrum A and F6p2, spectrum C) and after HPLC purification (respectively F1p2M, spectrum B and F6p2M, spectrum D). The major peak at 2.6 ppm is assigned to citric acid. Zooms on spectra A and C reveal the protein signals (A1 chitinase; C1 TL)



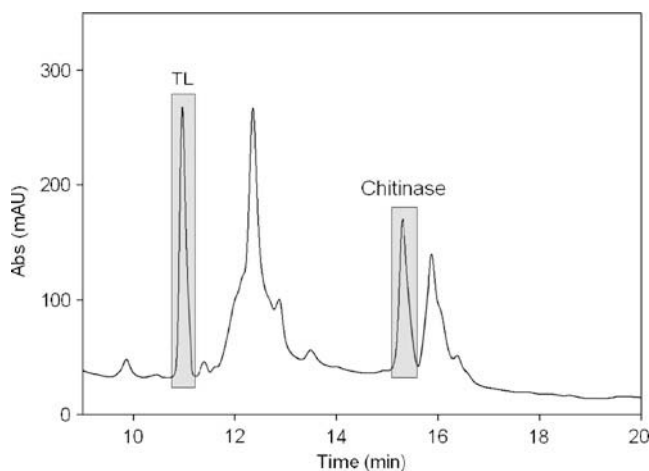
incubated in 3% (w/v) BSA-PBS (100  $\mu\text{L}$ /well) for 2 h at room temperature and washed twice with PBST and once with PBS. The plates were subsequently incubated with appropriate dilutions of the antiserum (diluted with PBS containing 1% (w/v) BSA, 100  $\mu\text{L}$ /well) for 2 h at room temperature and washed twice with PBST and once with PBS. The goat anti-rabbit peroxidase conjugate (diluted with PBS containing 1% (w/v) BSA, 100  $\mu\text{L}$ /well) was added and incubated for 2 h at room temperature. After the plates were washed twice with PBST and once with PBS, the peroxidase substrate was added (100  $\mu\text{L}$ /well). It was made of 25 mL phosphate-citrate buffer 0.1 M pH 5.0 plus 6  $\mu\text{L}$   $\text{H}_2\text{O}_2$  and 10 mg *ortho*-phenyldiamine. After the plate was covered and agitated for 15 min, the coloration was stopped by adding 50  $\mu\text{L}$  of 3 M HCl. The absorbance of each well was determined at 490 nm using an immunoplate device (Model 550, Bio-rad, Hercules, USA). TL and chitinase concentrations were detected up to 1 ng/mL. The precision of the method was evaluated by performing three replicate analyses of each juice. The coefficients of variation ranged from 9 to 16% for both proteins.

## Results and discussion

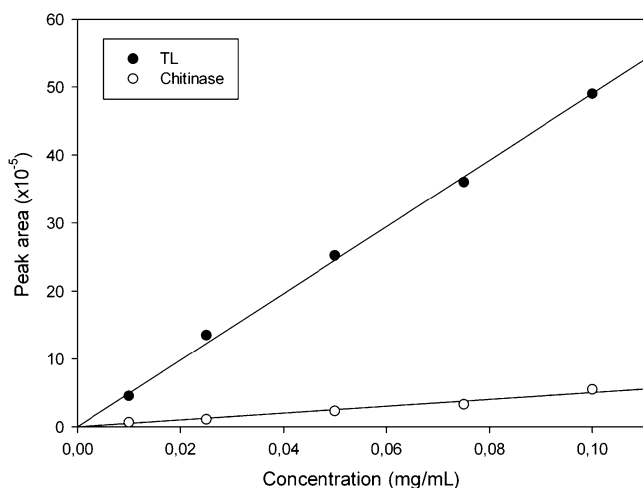
### SDS-PAGE analysis of grape juice and wine samples

The protein composition of five grape juices and wines was analyzed by silver-stained SDS-PAGE (Fig. 1): two grape juice samples, from two different grapevine varieties, Chardonnay and Pinot Meunier, harvested in 2009; one

base wine Chardonnay 2008; and two commercial Champagnes (Chardonnay 2008 and 2006). The silver-stain procedure is more sensitive than the Coomassie brilliant blue one, so it allows protein detection without prior concentration of grape juice and wine samples. TL (18 kDa) and chitinase (30 kDa) bands can be clearly observed in each grape juice sample, which confirms their abundance in grapes. Both bands were also detected in base wine, showing the concentration of chitinase clearly decreases throughout the vinification process. No chitinase band was thus detected in both Champagnes, but a slight band of TL can be observed. From Fig. 1, it then appears that these two proteins do not entirely persist from grape



**Fig. 5** UHPLC (C18) profile for the grape juice Chardonnay 2009. TL and chitinase peak areas were determined for protein quantification. Load and wash are omitted



**Fig. 6** Calibration curves for TL and chitinase performed on UHPLC (C18) using purified proteins as standards. A very good correlation is obtained (TL,  $r^2=0.9986$ ; chitinase,  $r^2=0.9750$ )

juice to Champagne, whereas they were shown to be resistant to yeast proteolysis [4]. The minor differences observed between both Champagnes could be a result of the different protein composition in the source grape juices. One can also observe that the band corresponding to invertase (60 kDa), another major grape protein [19], remains at around the same level from grape juice to Champagne. The intensity of chitinase bands appears lower than the TL one, but no comparison of their relative concentration can be made as stains could not react in the same way for each protein [15].

To perform an adequate determination of the actual concentration of chitinase and TL protein in grape juice and wine samples, pure fractions of both proteins used as standards are thus required.

#### Purification of TL and chitinase from grape juice

The main isoforms of chitinase (class IV chitinase, gi|225434074) and TL (VVTL1, gi|2213852) proteins were purified from Chardonnay grape juice according to the protocol developed by Van Sluyter et al. [15]. The fractionation of the crude juice was performed using strong cation exchange (SCX) followed by hydrophobic interaction (HIC)

FPLC (Electronic Supplementary Material Figs. S2 and S3). SDS-PAGE (stained with Coomassie brilliant blue) and western blot analyses were performed on the different HIC collected fractions. F1p2 was selected as the chitinase fraction, and F6p2 as the TL one (Fig. 2). After dialysis and freeze-drying, the purity of each fraction was analyzed by mass spectrometry (Table 1). As predicted, F1p2 was identified as class IV chitinase (gi|225434074), and F6p2 as VVTL1 (gi|2213852). Both mass spectrometry and HPLC analysis (on a C18 column) (Fig. 3) on each fraction revealed the presence of other minor grape proteins, which were not detected by SDS-PAGE. In addition,  $^1\text{H}$  NMR analyses were performed on each fraction to detect organic contaminants (Fig. 4). The results revealed a huge amount of citric acid in the fractions. This contaminant likely came from the HIC sodium citrate buffer and was not entirely filtered out by dialysis. The presence of other proteins and citric acid in the TL and chitinase fractions did not allow us to set up accurate calibration curves for the quantification methods. An additional HPLC purification step was thus performed for the two fractions (Fig. 3). The system was equipped with a collector device that allows one to collect and to pool the subfractions corresponding to the main peaks for each fraction: F1p1M and F6p2M. After dialysis and freeze-drying, they undergo the same mass spectrometry, HPLC, and  $^1\text{H}$  NMR analyses to assess their purity (Table 1, Figs. 3 and 4). As neither proteins nor organic contaminants were detected, pure fractions of TL and chitinase (respectively 12 mg and 10 mg) were finally obtained.

#### Quantification of chitinase and TL protein in grape juice and wine samples

To provide an accurate quantification of the main chitinase and TL proteins in grape juice and wine samples, two independent methods, UHPLC and ELISA, were performed in this study. For both techniques, the quantification relies on standard calibration curves made using exact weights of pure dried TL and chitinase.

UHPLC allows a better resolution and a shorter analysis time than HPLC [20]. The same column medium and buffers as for the protein purification were used but the

**Table 2** Chitinase and TL concentrations (mg/mL) in the different studied samples determined by UHPLC and ELISA

	UHPLC (mg/mL)		ELISA (mg/mL)	
	Chitinase	TL	Chitinase	TL
Grape juice CH09	124.3±6.1	105.1±3.2	118.3±10.3	89.0±13.9
Grape juice PM09	65.8±5.9	97.4±9.2	59.1±9.2	85.8±8.5
Base wine CH08	12.4±1.3	36.5±6.1	8.7±1.3	25.8±4.1
Champagne CH08	3.1±0.5	1.1±0.02	4.4±0.5	1.1±0.1
Champagne CH06	3.5±0.5	0.5±0.03	5.6±0.7	0.8±0.1

gradient time was reduced from 60 to 20 min. Juices were loaded (100  $\mu$ L) on the C18 column and the proteins were eluted with an increasing acetonitrile gradient (Fig. 5). The five grape juices and wines were concentrated in the same way: 5-mg samples of the macromolecular fraction obtained after dialysis and freeze-drying of each sample were redissolved in 1 mL of ultrapure water. All the juices roughly present the same protein profiles (data not shown), confirming the previous observation that there is no major discrepancy within the protein composition in different grape juice or wine samples. A calibration curve was set up for each protein, loading pure TL and chitinase in water at different concentrations (from 0.01 to 0.1 mg/mL) under the same chromatographic conditions (Fig. 6). As expected, the two proteins showed very different absorbance/concentration correlation, which emphasizes the absolute need to use as a standard the pure protein to be quantified in a sample. Moreover, one can wonder whether the previous approach developed by Marangon et al. [14] using BSA and Cyt C as standard proteins is truly accurate, considering BSA and Cyt C may not give an HPLC response equivalent to grape proteins. The TL and chitinase concentrations in grape juices and wines were calculated using the calibration lines and the dilution factor (Table 2).

ELISA was performed in parallel to measure the TL and chitinase concentrations in the same grape juice and wine samples according to the protocol developed by Dambrouck et al. [12] to quantify grape invertase. Polyclonal antibodies were raised against the purified TL and chitinase in rabbits. The antibodies were tested by using the western blot technique and showed a high specificity for the antigen: no cross-reaction was observed with other grape or wine proteins (data not shown). Calibration curves (absorbance at 490 nm versus the TL or chitinase concentration) were set up by coating wells of a microtiter plate with different concentrations of the purified proteins ranging from 50  $\mu$ g/mL to 100  $\mu$ g/mL. The protein concentration in grape juices and wines was then determined and are reported in Table 2.

The comparison of the TL and chitinase concentrations measured by UHPLC and ELISA shows very consistent results (Table 2). Relative standard deviations (RSD) calculated for chitinase and TL for both techniques are lower than 16%. UHPLC and ELISA measurements are based on completely different protein properties so their combined use is the best way to ensure an adequate and reliable protein concentration. The results confirmed the observation made during the SDS-PAGE analysis of the different samples: a general decrease of the protein concentration throughout the vinification process. However, TL and chitinase are concentrated at around the same level for each sample, which is not consistent with the band intensities observed on gel and confirms the different protein response to SDS-PAGE stains.

UHPLC and ELISA are thus both suitable for individual protein quantification in grape juices and wines. Both methods were performed on dialyzed and concentrated samples to suit UHPLC requirements, but one could consider performing ELISA on crude samples.

## Conclusion

The main isoforms of chitinase and TL proteins from grape juice were purified using the same preparative liquid chromatography procedure. Their high level of purity, assessed by HPLC, mass spectrometry, and NMR spectroscopy, allowed their use as standards for quantification techniques.

The results showed that both UHPLC and ELISA methods using the purified proteins as standards were successfully applied for the quantification of chitinase and TL protein in grape juices and wines. Knowledge of grape protein changes during winemaking is of great technological relevance as it can be applied to improve wine quality. More studies should be performed to correlate grape juice and wine composition with physicochemical characteristics of foam in sparkling wines. Structural studies achieved on pure proteins could also lead to a better understanding of the phenomena they are involved in.

**Acknowledgements** We thank the Champagne Ardenne Area, the General Council of Mame, the Ministry of Research, the Ville de Reims, the A.R.O.C.U. Association, and the Champagne Moët & Chandon for their support. We gratefully thank the wine cooperative of Nogent l'Abesse (France) for providing grape juice and wines samples. We also acknowledge Dr. S. Jégou (University of Reims Champagne-Ardenne, France) in appreciation of her contribution.

## References

1. Le Bourse D, Jégou S, Conreux A, Villaume S, Jeandet P (2010) Review of preparative and analytical procedures for the study of proteins in grape juice and wine. *Anal Chim Acta* 667:33–42
2. Robinson SP, Davies C (2000) Molecular biology of grape berry ripening. *Aust J Grape Wine Res* 6:175–188
3. Giribaldi M, Giuffrida MG (2010) Heard it through the grapevine: proteomic perspective on grape and wine. *J Proteomics* 73:1647–1655
4. Pocock KF, Hayasaka Y, McCarthy MG, Waters EJ (2000) Thaumatin-like proteins and chitinases, the haze-forming proteins of wine, accumulate during ripening of grape (*Vitis vinifera*) berries and drought stress does not affect the final levels per berry at maturity. *J Agric Food Chem* 48:1637–1643
5. Tattersall DB, Van Heeswijk R, Hoj PB (1997) Identification and characterization of a fruit-specific, thaumatin-like protein that accumulates at very high levels in conjunction with the onset of sugar accumulation and berry softening in grapes. *Plant Physiol* 114:759–769
6. Salzman RA, Tikhonova I, Bordelon BP, Hasegawa PM, Bressan RA (1998) Coordinate accumulation of antifungal proteins and hexoses constitutes a developmentally controlled defense response during fruit ripening in grape. *Plant Physiol* 117:465–472



7. Waters EJ, Alexander G, Muhlack R, Pocock KF, Colby C, O'Neill BN, Høj PB, Jones PR (2005) Preventing protein haze in bottled white wine. *Aust J Grape Wine Res* 11:215–225
8. Pocock KF, Alexander GM, Hayasaka Y, Jones PR, Waters EJ (2007) Sulfate – a candidate for the missing essential factor that is required for the formation of protein haze in white wine. *J Agric Food Chem* 55:1799–1807
9. Pocock KF, Waters EJ (2006) Protein haze in bottled white wines: how well do stability tests and bentonite fining trials predict haze formation during storage and transport? *Aust J Grape Wine Res* 12:212–220
10. Falconer RJ, Marangon M, Van Sluyter SC, Neilson KA, Chan C, Waters EJ (2010) Thermal stability of thaumatin-like protein, chitinase, and invertase isolated from sauvignon blanc and semillon juice and their role in haze formation in wine. *J Agric Food Chem* 58:975–980
11. Marangon M, Van Sluyter SC, Neilson KA, Chan C, Haynes PA, Waters EJ, Falconer RJ (2011) Roles of grape thaumatin-like protein and chitinase in white wine haze formation. *J Agric Food Chem* 59:733–740
12. Dambrouck T, Marchal R, Cilindre C, Parmentier M, Jeandet P (2005) Determination of the grape invertase content (using PTA-ELISA) following various fining treatments versus changes in the total protein content of wine. Relationships with wine foamability. *J Agric Food Chem* 53:8782–8789
13. Cilindre C, Castro AJ, Clément C, Jeandet P, Marchal R (2007) Influence of *Botrytis cinerea* infection on Champagne wine proteins (characterized by two-dimensional electrophoresis/immunodetection) and wine foaming properties. *Food Chem* 103:139–149
14. Marangon M, Van Sluyter SC, Haynes PA, Waters EJ (2009) Grape and wine proteins: their fractionation by hydrophobic interaction chromatography and identification by chromatographic and proteomic analysis. *J Agric Food Chem* 57:4415–4425
15. Van Sluyter SC, Marangon M, Stranks SD, Neilson KA, Hayasaka Y, Haynes PA, Ian Menz R, Waters EJ (2009) Two-step purification of pathogenesis-related proteins from grape juice and crystallization of thaumatin-like proteins. *J Agric Food Chem* 57:11376–11380
16. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
17. Rabilloud T, Vuillard L, Gilly C, Lawrence JJ (1994) Silver-staining of proteins in polyacrylamide gels: a general overview. *Cell Mol Biol* 40:57–75
18. Anderson NL, Esquer-Blasco R, Hofmann J-P, Anderson NG (1991) A two-dimensional gel database of rat liver proteins useful in gene regulation and drug effects studies. *Electrophoresis* 12:907–930
19. Jégou S, Conreux A, Villaume S, Hovasse A, Schaeffer C, Cilindre C, Van Dorsselaer A, Jeandet P (2009) One step purification of the grape vacuolar invertase. *Anal Chim Acta* 638:75–78
20. Nováková L, Vlčková H (2009) A review of current trends and advances in modern bio-analytical methods: chromatography and sample preparation. *Anal Chim Acta* 656:8–35

## Erratum to: Quantification of chitinase and thaumatin-like proteins in grape juices and wines

D. Le Bourse · A. Conreux · S. Villaume · P. Lameiras ·  
J.-M. Nuzillard · P. Jeandet

Published online: 8 May 2011  
© Springer-Verlag 2011

**Erratum to: Anal Bioanal Chem**  
**DOI 10.1007/s00216-011-4912-8**

Table 2 in the above article is incorrect. Below please find the corrected version where protein concentrations are given in mg/L. The authors apologize for the mistake.

**Table 2** Chitinase and TL concentrations (mg/L) in the different studied samples determined by UHPLC and ELISA

	UHPLC (mg/L)		ELISA (mg/L)	
	Chitinase	TL	Chitinase	TL
Grape juice CH09	124.3±6.1	105.1±3.2	118.3±10.3	89.0±13.9
Grape juice PM09	65.8±5.9	97.4±9.2	59.1±9.2	85.8±8.5
Base wine CH08	12.4±1.3	36.5±6.1	8.7±1.3	25.8±4.1
Champagne CH08	3.1±0.5	1.1±0.02	4.4±0.5	1.1±0.1
Champagne CH06	3.5±0.5	0.5±0.03	5.6±0.7	0.8±0.1

The online version of the original article can be found at <http://dx.doi.org/10.1007/s00216-011-4912-8>.

D. Le Bourse (✉) · A. Conreux · S. Villaume · P. Jeandet  
Laboratoire d'Oenologie et Chimie Appliquée,  
Université de Reims Champagne-Ardenne,  
URVVC-SE UPRES EA 2069, BP 1039,  
51687 Reims Cedex 2, France  
e-mail: [doriane.le-bourse@univ-reims.fr](mailto:doriane.le-bourse@univ-reims.fr)

P. Lameiras · J.-M. Nuzillard  
CNRS, Institut de Chimie Moléculaire de Reims,  
BP 1039, 51687 Reims Cedex 2, France

### 3.2 Propriétés tensioactives des protéines thaumatin-like et chitinase

L'objectif de cette partie était d'évaluer le rôle des protéines thaumatin-like et chitinase dans la formation des bulles et la stabilisation de la mousse d'un Champagne.

Une première expérience a été réalisée sur chacune des deux protéines purifiées (cf partie précédente) solubilisées dans de l'eau pure à 1 mg/mL, soit environ 1000 fois plus concentrée que dans un Champagne. Après une forte agitation, il apparaît que la solution de TL montre une forte propension à mousser alors que celle de chitinase ne présente pas la moindre bulle à sa surface (Figure 8).

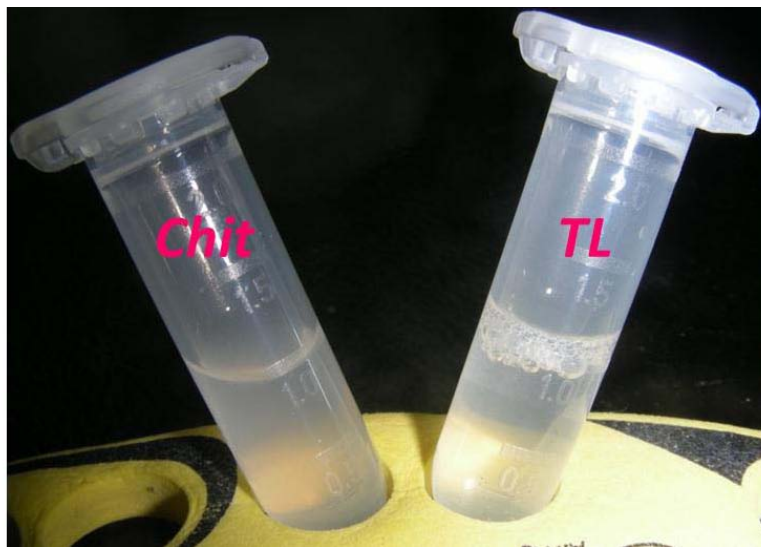


Figure 8 : Solutions de chitinase (à gauche) et TL (à droite) diluées à 1 mg/mL dans de l'eau pure après agitation.

Ce résultat visuel a été confirmé par tensiométrie à lame de Wilhelmy. A 20°C, l'étude de différentes solutions de TL diluée dans de l'eau pure de 0 à 500 mg/L a montré que cette protéine est un bon agent tensioactif, avec une pression de surface proche de 22 mN/m aux alentours de la concentration micellaire critique (Figure 9). Cependant, en reprenant la concentration de TL dans un Champagne située entre 0.5 et 1.1 mg/L [52], la pression de surface devient comprise entre 0.02 et 0.05 mN/m. La contribution de TL à l'interface gaz/liquide peut ainsi être considérée comme négligeable dans ces conditions. D'autre part, l'étude de chitinase sur le même principe a révélé son absence de propriété tensioactive : quelle que soit la concentration de l'échantillon aucune différence n'a été mesurable par rapport à une solution d'eau pure (pression de surface = 0 mN/m).

En partant du principe que le protocole de purification des protéines n'a pas altéré leur structure tridimensionnelle et que cette structure est conservée du jus au vin, il apparaît donc que ni TL ni chitinase ne peuvent être à elles seules responsables de l'effervescence d'un Champagne. Des expériences complémentaires menées sur des solutions de protéines diluées dans des vins modèles permettraient d'évaluer les effets synergiques ou antagonistes protéine/protéine ou protéine/autres composés, et ainsi confirmer ou infirmer définitivement leur implication en tant que tensioactif dans un Champagne.

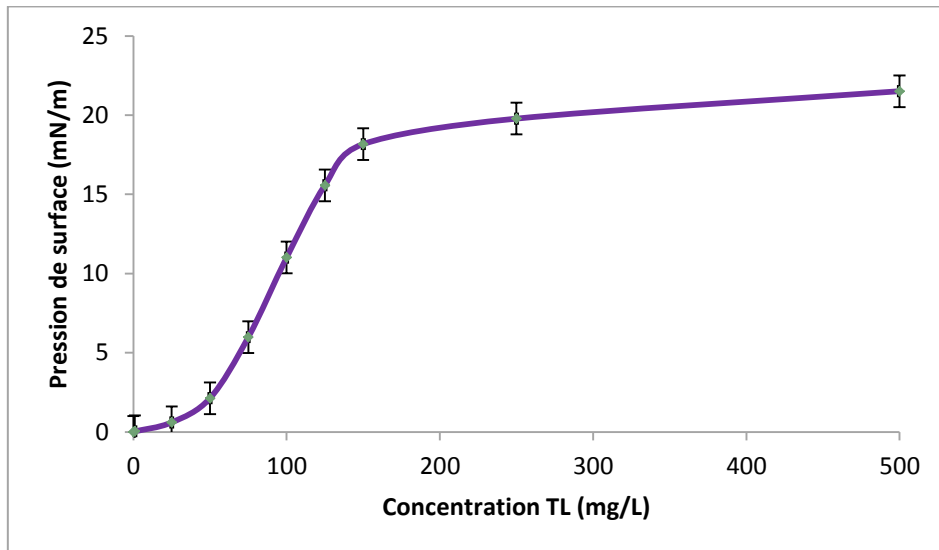


Figure 9 : Evolution de la pression de surface en fonction de la concentration en TL dans de l'eau pure.

### 3.3 Conclusion

La purification des protéines TL et chitinase à partir de jus de raisin a permis de réaliser des gammes standards dans le cadre de deux différentes méthodes de quantification de ces protéines du jus au Champagne. L'excellente corrélation des résultats obtenus par chromatographie liquide et par dosage immuno-enzymatique indique que l'une ou l'autre des méthodes est applicable pour la détermination de la concentration de TL et chitinase. Dans le cas de ces deux protéines, la technique UHPLC semble la technique la plus adaptée car elle est bien plus rapide et facile à mettre en place qu'une expérience ELISA (trente minutes vs deux jours). Cependant, cette dernière méthode trouve un grand intérêt lors de l'étude de protéines difficiles à faire passer sur une colonne de chromatographie, comme l'invertase de raisin [58].

Les propriétés tensioactives des deux protéines purifiées ont ensuite été mesurées par tensiométrie à lame de Wilhelmy. A partir des concentrations précédemment déterminées, il a été possible d'établir que la contribution individuelle de TL et de chitinase est négligeable dans le cadre des bulles et de la mousse d'un Champagne. Il reste cependant nécessaire de poursuivre ce travail en évaluant les interactions de ces protéines avec d'autres protéines et d'autres composés du vin qui pourraient modifier la capacité d'adsorption des agents tensioactifs aux interfaces gaz/liquide. L'approche systématique purification/quantification/tensiométrie menée sur toutes les protéines présentes dans un Champagne pourrait à terme permettre d'élucider la relation liant la composition du vin et la qualité de ses bulles et de sa mousse.

La compréhension des mécanismes dans lesquels les protéines sont impliquées, que ce soit au niveau d'une interface ou dans un autre contexte, notamment leur activité biologique chez la vigne, passe inévitablement par une étude structurale. Une première approche de la détermination de la structure de TL a ainsi été menée, ce qui fait l'objet de la partie suivante.

## Partie II : Etude de la structure d'une protéine thaumatin-like et d'un peptide

*Theoretical approaches to protein structure are fraught with difficulty, so much so that it is unclear when it will be possible to predict protein structure from sequence. In considering why the problem is so difficult, let me recount a conversation with Francis Crick (ndla : Prix Nobel de médecine en 1962 pour la découverte de la structure de l'ADN) in 1975. Crick stated that "it is very difficult to conceive of a scientific problem that would not be solved in the coming twenty years... [...] ...except for a model of brain function and protein folding". Although Crick was more interested in brain function, he did state that both problems were difficult because they involve many cooperative interactions in three-dimensional space.*

*Levitt M., Theory and simulation Through the breach, Curr. Opin. Struct. Biol. 1996, 2, 193-194*



# 1 Introduction

## 1.1 Structure des protéines

Une protéine est une macromolécule constituée d'un enchaînement d'acides aminés. Cet enchaînement type « collier de perles » révèle la structure primaire de la protéine, résultant de la traduction du gène codant cette protéine et déterminant son identité de manière unique. La chaîne d'acides aminés, soumise aux différentes interactions entre ses éléments et avec son milieu environnant, se replie localement et dans l'espace adoptant ainsi une structure tridimensionnelle qui traduit le résultat de la minimisation de l'énergie libre du système. La conformation spatiale d'une protéine lui permet d'acquérir ses propriétés physicochimiques et biologiques, caractérisant son fonctionnement et le rôle qu'elle peut jouer dans différents phénomènes.

### 1.1.1 Structure primaire : chaîne d'acides aminés

Le premier niveau de structuration d'une protéine correspond à l'enchaînement séquentiel des acides aminés qui la composent. Il existe dans la nature 20 acides aminés différents (Figure 10), constitués d'une fonction amine reliée à une fonction acide carboxylique par un carbone aliphatique ( $C_{\alpha}$ ). Ce  $C_{\alpha}$  porte une chaîne latérale (groupement R) qui caractérise l'acide aminé et dont la nature (hydrophobe, polaire ou chargée) va déterminer ses interactions avec son environnement. Excepté dans le cas de la glycine où  $R=H$ , le  $C_{\alpha}$  est un carbone asymétrique. Pour les acides aminés naturels, la configuration stéréochimique de ce centre chiral est en général L dans la nomenclature de Fischer.

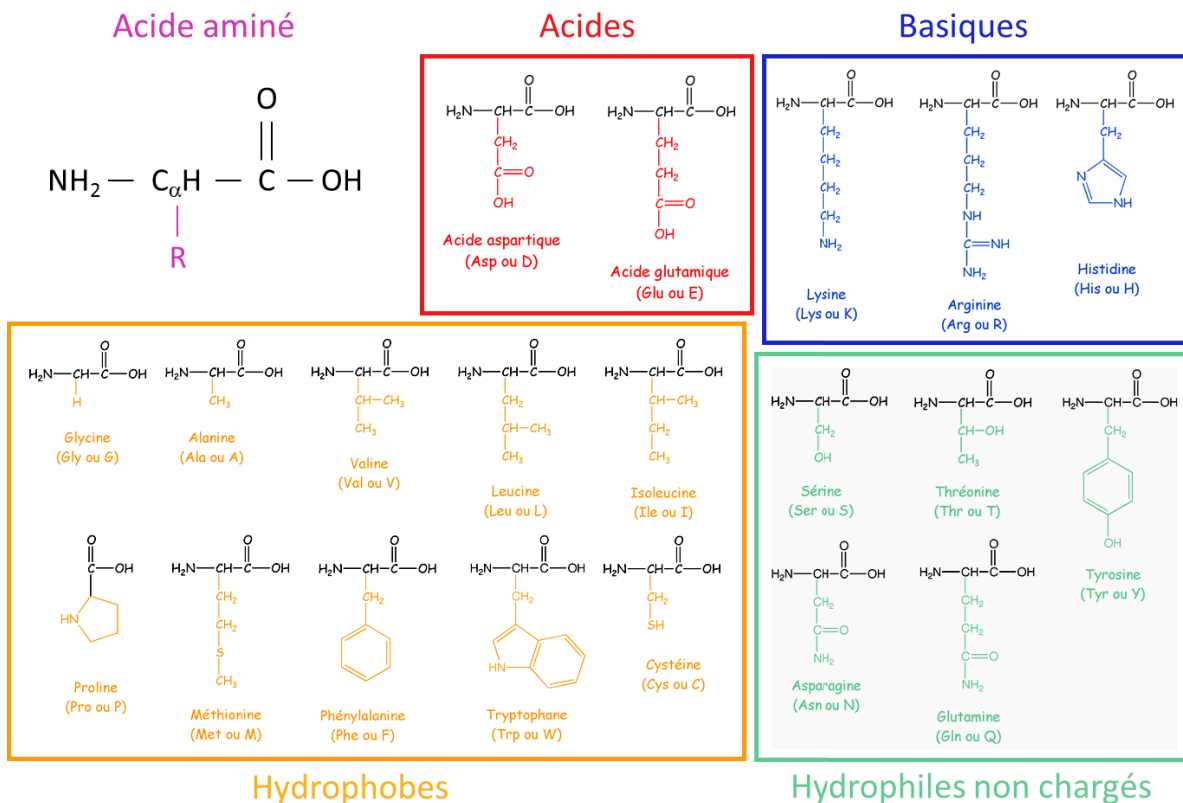


Figure 10 : Description des acides aminés.

Deux acides aminés ( $i$  et  $i+1$ ) peuvent s'enchaîner par la création d'une liaison peptidique reliant la fonction amine de l'un à la fonction acide carboxylique de l'autre (Figure 11). La séquence d'acides aminés ordonnée du résidu N-terminal au résidu C-terminal constitue ainsi la structure primaire d'une protéine.

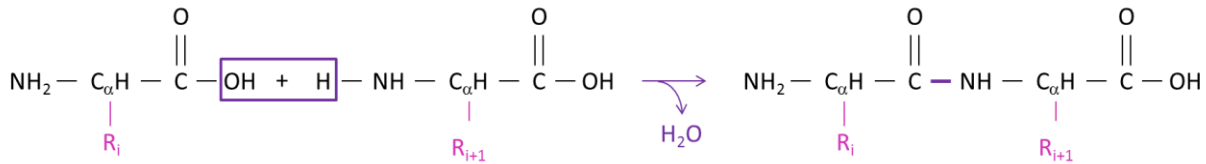


Figure 11 : Formation d'une liaison peptidique entre deux acides aminés.

### 1.1.2 Structure secondaire : repliement local

La structure secondaire représente le premier niveau d'agencement spatial d'un groupe d'acides aminés (brin peptidique), régi uniquement par les liaisons hydrogène qui peuvent se former entre les groupements C=O et H-N.

#### ▪ Notion d'angle dièdre

Les conformations adoptées par un brin (hélice, feuillet, coude...) sont imposées par les trois angles dièdres  $\phi$ ,  $\psi$  et  $\omega$  (Figure 12). Les angles dièdres sont les angles entre deux plans autour d'une liaison : l'angle  $\phi$  autour de la liaison N-C $_{\alpha}$ ,  $\psi$  autour de la liaison C $_{\alpha}$ -C,  $\omega$  autour de la liaison peptidique.

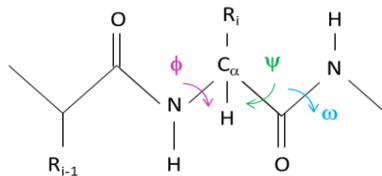


Figure 12 : Définition des angles dièdres dans un enchaînement d'acides aminés.

Les valeurs prises par ces angles sont soumises à des contraintes réduisant les géométries possibles. La conjugaison entre le doublet non-liant de l'azote et la liaison  $\pi$  du groupement carbonyle lors de la formation d'une liaison peptidique lui impose une conformation plane, ne permettant ainsi que deux valeurs pour l'angle  $\omega$  :  $0^\circ$  (forme cis) et  $180^\circ$  (forme trans). Cependant, l'encombrement stérique des chaînes latérales défavorise fortement la conformation cis. Les brins peptidiques s'arrangent donc généralement dans une conformation trans, où les C $_{\alpha}$  se retrouvent de part et d'autre de la liaison peptidique.

#### ▪ Diagramme de Ramachandran

L'encombrement stérique des chaînes latérales régit également les valeurs permises pour les combinaisons d'angles ( $\phi$ ,  $\psi$ ) [62]. Le diagramme de Ramachandran présente les valeurs corrélées de ces deux angles, définissant des régions interdites et des régions accessibles, impliquant le type de repliement favorisé associé (Figure 13). La proline et la glycine, en raison de leur structure particulière, présentent des cartes différentes de la carte générale.



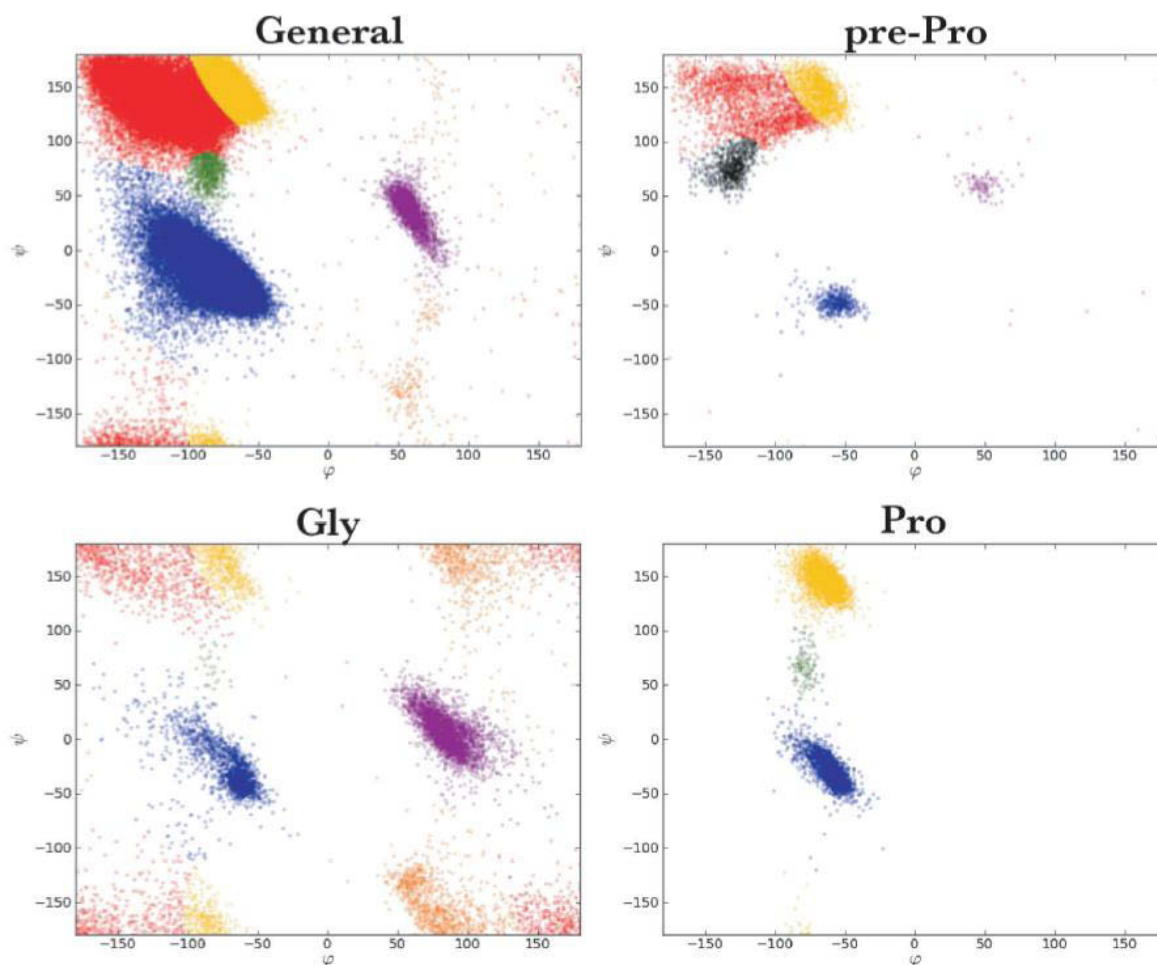


Figure 13 : Diagramme de Ramachandran présentant les domaines de conformations accessibles pour les couples ( $\phi$ ,  $\psi$ ) et non-atteintes (en blanc). Sept domaines sont représentés : feuillet  $\beta$  en rouge, hélice-droite en bleu, hélice-gauche en violet, coude  $\gamma$  en vert, hélice polyproline en jaune et orange,  $\zeta$  (pré-proline) en noir. Reproduit avec l'aimable autorisation de Matthieu Tanty, dans le cadre de son travail sur le logiciel RamaDA.

### ▪ Types de structure secondaire

Comme visualisé sur le diagramme de Ramachandran, les acides aminés peuvent s'organiser localement selon différents types de structure secondaire.

#### Hélice

L'hélice est une structure en forme de ressort. La plus fréquente est l'hélice  $\alpha$  (Figure 14), correspondant à la région ( $-170^\circ \leq \phi \leq -50^\circ$  ;  $-70^\circ \leq \psi \leq -40^\circ$ ) sur le diagramme de Ramachandran. Généralement droite (elle s'éloigne en tournant dans le sens des aiguilles d'une montre), elle contient 3.6 résidus par tour et son pas est de 5.41 Å. Elle est stabilisée par la formation de liaisons hydrogène entre l'atome d'oxygène du groupement C=O du résidu  $i$  et l'atome d'hydrogène du groupement N-H du résidu  $i+4$ .

D'autres types d'hélices peuvent être formés selon un réseau de liaisons hydrogène différent : l'hélice  $3_{10}$  (O du résidu  $i$  se liant avec NH du résidu  $i+3$ ) et l'hélice  $\pi$  (O du résidu  $i$  se liant avec NH du résidu  $i+5$ ). Dans tous les cas, les chaînes latérales des acides aminés sont dirigées vers l'extérieur.

### Feuillet $\beta$

Le feuillet  $\beta$  (Figure 14) représente un assemblage de brins peptidiques en forme d'accordéon. Il correspond à la région ( $-170^\circ \leq \phi \leq -50^\circ$  ;  $90^\circ \leq \psi \leq 175^\circ$ ) sur le diagramme de Ramachandran. L'agencement de deux ou plusieurs brins successifs peut se faire de manière parallèle (les bouts N-terminaux du même côté) ou anti-parallèle (sens contraire, les plus fréquents). Un feuillet  $\beta$  est stabilisé par la formation de liaisons hydrogène entre l'atome d'oxygène du groupement C=O d'un résidu sur un brin et l'atome d'hydrogène du groupement N-H d'un résidu sur le brin opposé. Il acquiert une forme plissée, les chaînes latérales se retrouvant alternativement au-dessus et en dessous du plan médian.

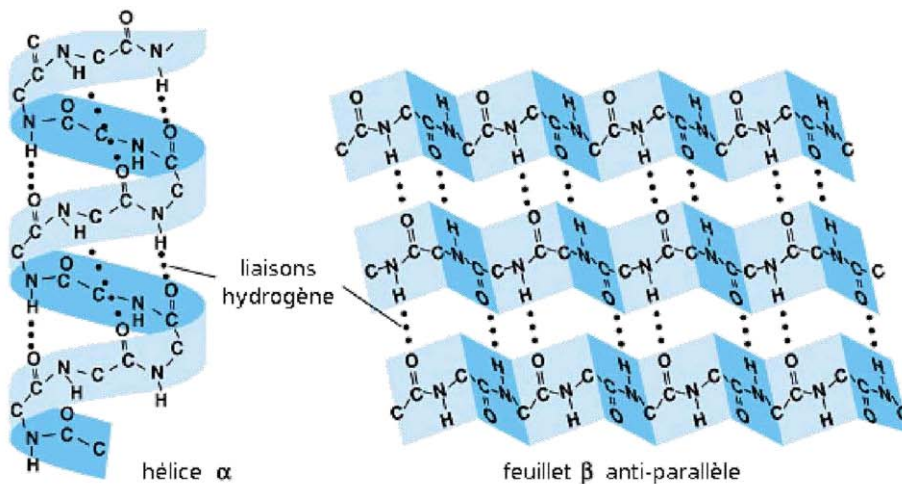


Figure 14 : Représentation d'une hélice  $\alpha$  (à gauche) et d'un feuillet  $\beta$  anti-parallèle (à droite).

### Autres structures

Bien que l'architecture d'une protéine soit constituée majoritairement d'hélices et de feuillets, le reste de la chaîne peut s'organiser sous forme de petits éléments structurés (boucle, coude) ou selon un état désordonné (pelote statistique).

Par exemple, les coudes  $\beta$  de type I ou II (Figure 15) sont caractérisés par une liaison hydrogène unique entre les résidus  $i$  et  $i+3$  servant à relier deux éléments de structure secondaire répétitive (hélice ou feuillet).

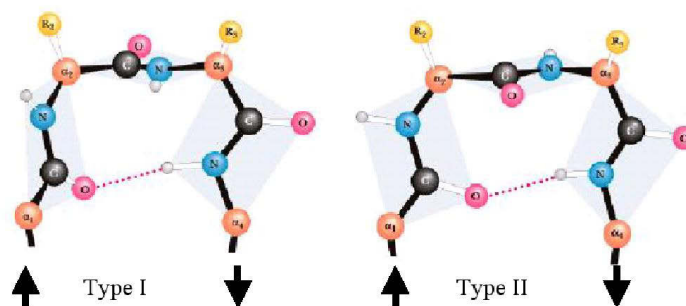


Figure 15 : Coudes  $\beta$  type I et II.

### 1.1.3 Structure tertiaire : agencement tridimensionnel

Les différents éléments de structure secondaire reliés par des connexions plus ou moins ordonnées s'agencent spatialement pour former la structure tertiaire de la protéine. Ce repliement tridimensionnel global résulte de l'interaction entre les différentes chaînes latérales des acides aminés de la séquence. Plusieurs types d'interaction peuvent stabiliser la structure (Figure 16) :

- Les ponts disulfures, liaisons covalentes qui résultent de l'oxydation de deux résidus cystéine.
- Les liaisons hydrogène, lorsqu'un hydrogène lié à un azote ou à un oxygène se retrouve à proximité d'un doublet électronique non partagé d'un autre azote ou d'un autre oxygène.
- Les ponts salins (interaction ionique), qui s'établissent entre deux groupements de charge opposée (acide/basique).
- Les interactions hydrophobes (forces de van der Waals), correspondant au rapprochement et à l'enfouissement des chaînes latérales hydrophobes afin de minimiser la perte entropique due à leur solvatation.

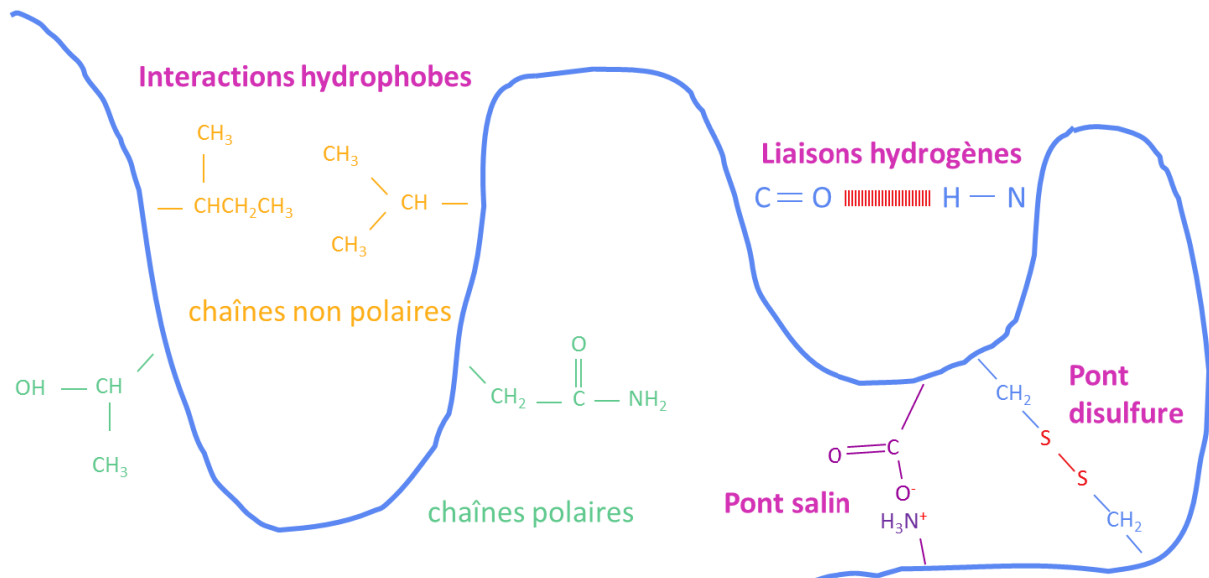


Figure 16 : Représentation des interactions qui gouvernent le repliement d'une protéine.

### 1.1.4 Structure quaternaire : assemblage de plusieurs chaînes

La forme biologiquement active d'une protéine peut être constituée d'un assemblage de plusieurs chaînes (appelées alors monomères ou sous-unités) agencées dans l'espace sous la forme d'une structure quaternaire pour former un complexe macromoléculaire.

## 1.2 Superfamille des protéines thaumatin-like : « such a sweet world »

Les protéines TL sont composées d'environ 200 acides aminés et présentent une séquence similaire à la thaumatin, une protéine au goût très sucré découverte dans le fruit d'un arbuste tropical, le Katemfe (*Thaumatococcus daniellii*) [63] utilisée aujourd'hui comme édulcorant (additif E957).

L'expression des protéines TL chez les végétaux est généralement induite par un stress biotique ou abiotique, mais elle peut exister de manière constitutive notamment chez la vigne [25-27]. Les protéines TL, impliquées dans les mécanismes de défense des végétaux, constituent la famille des PR5 (Pathogenesis-Related family 5) [29]. Outre chez les végétaux, des protéines TL ont été identifiées chez des espèces animales et des champignons [64-67].

### 1.2.1 Structure

Les séquences des protéines TL sont relativement bien conservées chez les végétaux. Elles partagent un motif caractéristique (thaumatin family signature PS00316) : G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C [68, 69]. Dans le cas de la protéine thaumatin-like de vigne VVTL1 (*Vitis Vinifera* Thaumatin-Like 1), ce motif est constitué de la séquence GRGKCE TGD CNG LLE C (acides aminés 56 à 71). La plupart des protéines TL possèdent seize résidus cystéine conservés, permettant la formation de huit ponts disulfures qui expliqueraient la très grande stabilité de ces protéines vis-à-vis de conditions extrêmes de température et de pH ainsi que leur résistance à la protéolyse chez la vigne [70, 71] et dans d'autres espèces [72, 73]. En plus de la thaumatin I [74] et de la thaumatin II [75], six structures de protéines TL ont été jusqu'à présent résolues par cristallographie : zeamatin de maïs (1DU5) [76], PR-5d (1AUN) [77] et osmotin (1PCV) [78] de tabac, Pru Av2 de cerise (2AHN) [79], Ba-TLP de banane (1Z3Q) [80] et NP24-I de tomate (2IOW) [81]. L'ensemble de ces structures révèle une organisation très conservée divisée en trois domaines peptidiques (Figure 17, Figure 18). Les domaines I et II forment une cavité acide contenant cinq acides aminés très conservés disséminés dans la séquence (REDD) [77].

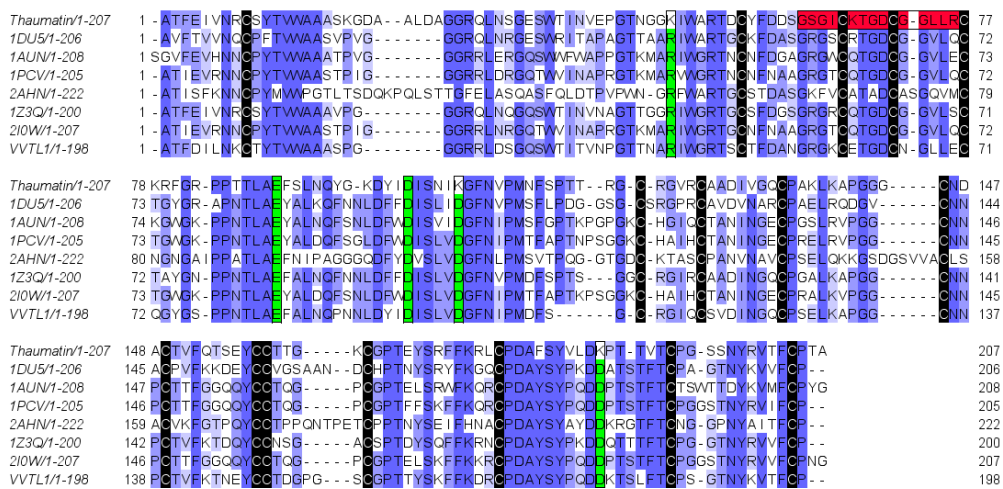


Figure 17 : Alignement des séquences de la thaumatin I, des six séquences des TL dont la structure a été résolue et de la TL de raisin VVTL1. L'intensité du fond bleu indique le degré de conservation des acides aminés. La signature thaumatin est colorée en rouge, les cystéines en noir et les résidus liés à l'acidité de la cavité entre les domaines I et II en vert.

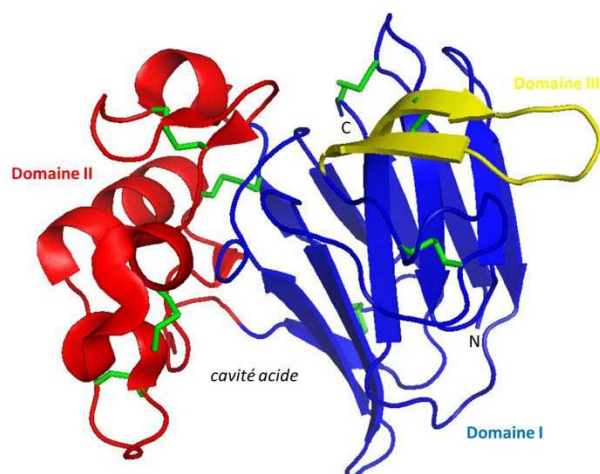


Figure 18 : Structure de la protéine TL de banane 1Z3Q. Les huit ponts disulfures sont indiqués en vert.

### 1.2.2 Fonction

La plupart des protéines TL dont le rôle biologique a été étudié montrent une activité antifongique *in vitro* [82, 83]. Le report de cette fonction sur un arbre phylogénétique des TL chez les eucaryotes montre qu'elle est commune à une grande majorité de la famille et probablement liée à la conservation des acides aminés REDD et de la cavité acide [84] dans la structure des protéines. Cependant le mécanisme de cette activité antifongique n'a pour le moment pas été entièrement élucidé. Différents modes d'action ont été proposés, parmi lesquels principalement une activité enzymatique endo- $\beta$ -1,3-glucanase [83]. En effet, le mode d'action antifongique pourrait être lié à la capacité des protéines TL de détruire la paroi cellulaire des champignons pathogènes composée de  $\beta$ -1,3-glucane par une activité enzymatique [85, 86]. Une première approche de docking protéine/ $\beta$ -1,3-glucane a été réalisée sur la TL de tomate 2I0W (Figure 19) [81], mais aucune interaction dynamique n'a pour l'instant été prise en compte.

Il apparaît cependant que l'activité endo- $\beta$ -1,3-glucanase n'est pas toujours corrélée avec l'activité antifongique des protéines TL [29] (Tableau 5). D'autres modes d'action ont ainsi été décrits, comme la perméabilisation des membranes [87], l'inhibition d'enzymes comme la xylanase [88], l' $\alpha$ -amylase et la trypsine [89], ainsi qu'un mécanisme d'induction d'apoptose reporté dans une levure [90]. Outre leur activité antifongique, les protéines TL peuvent également interagir avec des virus, comme reporté dans le cas du virus de la mosaïque du concombre [91]. En plus de leur rôle de défense contre un agent pathogène, des protéines TL ont montré une diversité de propriétés en interaction avec leur environnement : activité anti-gel [92, 93], protection contre un stress abiotique [94-96] et possibilité de liaison avec différentes protéines comme l'actine [97], des glycoprotéines de levure [98] et des GPCR (G-Protein Coupled Receptor) [90, 99, 100] ou avec des hormones comme les cytokinines [101]. Les protéines TL sont également connues pour pouvoir se lier avec l'immunoglobuline E humaine [81]. Elles présentent donc un fort potentiel allergénique pour l'homme, caractérisé notamment dans le raisin [102, 103], la pomme [73, 104, 105], la banane [80] et la farine de blé [106].

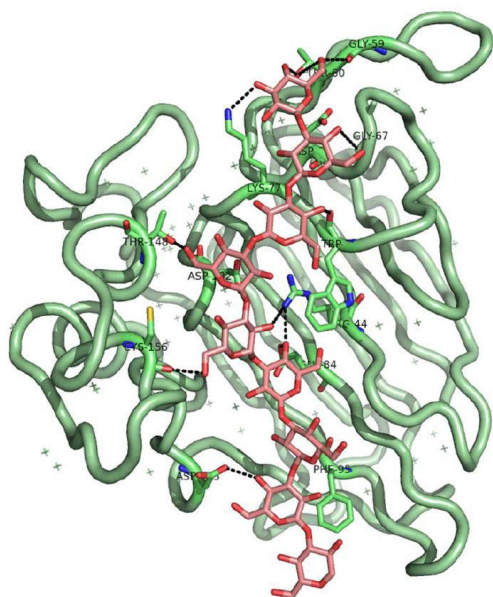


Figure 19 : Interaction entre un  $\beta$ -1,3-glucane (en rouge) et la protéine TL 210W (en vert) au cœur de la cavité acide. Reproduit de [81].

Protéine	Espèce	Activité		Référence
		antifongique <i>in vitro</i>	endo- $\beta$ -1,3-glucanase	
VVTL1	<i>Vitis vinifera</i>	X		[107]
TLP, OLP	<i>Vitis vinifera</i>	X		[108]
IFW19, Pr22-3 (feuilles)	<i>Hordeum vulgare</i>	X	X	[109, 110]
HvPR5c (graines)	<i>Hordeum vulgare</i>	X	non	[111]
TLP	<i>Irpex lacteus</i>	X	X	[66]
TLP, TLG1	<i>Lentinus edodes</i>		X	[66, 67]
AP24, NP24	<i>Lycopersicon esculentum</i>	X	X	[109]
P23	<i>Lycopersicon esculentum</i>	X		[112]
Mal-TLP	<i>Malus domestica</i>	non	X	[113]
Ban-TLP	<i>Musa acuminata</i>	non	X	[114]
Ban-TLP	<i>Musa acuminata</i>	X	non	[113]
TLP	<i>Musa basjoo</i>	X		[115]
Osmotin	<i>Nicotiana tabacum</i>	X		[116]
SE22, SE39b	<i>Nicotiana tabacum</i>		X	[109]
CHTL	<i>Prunus avium</i>	non	X	[109]
Pru-TLP	<i>Prunus avium</i>	non	X	[113]
TLP	<i>Rhizoctonia solani</i>	X	X	[66]
SniOLP	<i>Solanum nigrum</i>	X	non	[68]
SnOLP	<i>Solanum nigrum</i>	X		[117]
WAS-3a	<i>Triticum aestivum</i>	X		[118]
CdTLP	<i>Cassia didymobotrya</i>	X		[119]
CsTL1	<i>Castanea sativa</i>	X		[120]
CaTLP	<i>Castanopsis chinensis</i>	X		[121]

Tableau 5 : Activité identifiée chez différentes protéines TL. Adapté de [83] et [84].

## 2 Matériels et méthodes

La structure d'une protéine peut être étudiée, d'un point de vue statique ou dynamique, par un éventail de techniques expérimentales et théoriques. Le choix de l'approche utilisée dépend du type de protéine considérée (taille, accessibilité à une forme purifiée) et de l'information que l'on souhaite acquérir (protéine seule ou dans son environnement, implication dans des phénomènes biologiques ou physico-chimiques). La méthodologie appliquée pour caractériser la structure d'une protéine peut faire appel à une combinaison d'approches différentes, qui apportent des données complémentaires permettant d'intercorrélérer les résultats.

Cette section présente en premier lieu les outils utilisés pour l'étude de la protéine de la protéine de raisin VVTL1 (*Vitis Vinifera* Thaumatin-Like 1) divisés selon les niveaux successifs de structuration des protéines auxquels ils se rapportent. La méthodologie mise en place pour la détermination d'un peptide, issu de la séquence de VVTL1, par modélisation sous contraintes RMN est décrite en second lieu. Dans les deux cas, les principes des méthodes de modélisation et de dynamique moléculaires sont détaillés dans l'Annexe 2.

### 2.1 Etude de la structure de la protéine VVTL1

#### 2.1.1 Structure primaire : séquence déposée

Le point initial crucial pour l'étude de la structure d'une protéine est la connaissance de sa séquence d'acides aminés. Plusieurs techniques expérimentales peuvent être mises en place pour avoir accès à la séquence d'une protéine :

- Séquençage du gène codant, suivant les progrès de la biologie moléculaire.
- Séquençage total par spectrométrie de masse, requérant l'obtention préalable de la protéine sous forme pure.

L'étude de la structure de la protéine thaumatin-like ayant été initiée avant que nous soyons en mesure de l'obtenir expérimentalement sous forme pure, nous n'avons eu d'autre choix que de recourir à la séquence de la protéine VVTL1 [26] déposée dans la banque de données NCBI (National Center for Biotechnology Information, accessible à l'adresse <http://www.ncbi.nlm.nih.gov/>). Bien que la séquence déposée puisse potentiellement présenter de légères différences avec la séquence que l'on pourrait obtenir en séquençant la protéine purifiée, elle permet de fournir un point de départ à l'étude qui pourra être affinée lors de travaux ultérieurs.

La séquence de VVTL1 déposée contient un peptide signal initial de 24 acides aminés, non-présent dans la séquence finale. Les bornes de ce peptide ont été déterminées en utilisant les données croisées de la littérature, de séquences homologues et d'un logiciel de prédiction (signalP 3.0 [122], accessible à l'adresse <http://www.cbs.dtu.dk/services/SignalP/>).

La séquence de VVTL1 est donc finalement composée de 198 acides aminés :

```
ATFDILNKCTYTVWAAASPGGRRRLDSGQSWITVNPGETTNARIWGRTSCTFDANGRGKCEGTGDCNGL  
LECQGYGSPNNTLAEFALNQPNNLDYIDISLVDGFNIPMDFSGCRGIQCSVDINGQCPSELKAPGGCN  
NPCTVFKTNEYCCTDGPSCGPTTYSKFFKDRCPDAYSYPQDDKTSLFTCPSGTNYKVTFCP
```

### 2.1.2 Structure secondaire : dichroïsme circulaire

Le dichroïsme circulaire (DC) repose sur la capacité d'un échantillon optiquement actif (molécule chirale) à absorber différemment les composantes circulaires droite et gauche de la lumière. Dans le cas des protéines, la mesure du DC dans l'UV lointain (180-250 nm) permet de déterminer leur contenu en structures secondaires et de suivre les changements conformationnels en fonction des conditions expérimentales (concentration de l'échantillon, température, présence de sels...) [123].

#### ▪ Principe général

Une lumière polarisée plane peut être considérée comme la somme de deux ondes polarisées tournant circulairement en sens contraire (gauche et droite). Après la traversée d'un échantillon optiquement actif, chaque composante est encore polarisée circulairement, mais l'amplitude et la vitesse de propagation sont différentes pour les deux composantes. La lumière résultante, qui est la somme des deux composantes, n'est plus polarisée linéairement mais elliptiquement.

L'absorption de la lumière par un échantillon peut être reliée à son coefficient d'extinction molaire  $\varepsilon$ , à sa concentration molaire  $C$  et à la longueur de la cuve contenant l'échantillon (parcours optique)  $l$  par la loi de Beer-Lambert :

$$A = \varepsilon l C$$

Le pouvoir dichroïque d'un échantillon, défini par la différence d'absorption entre les deux composantes gauche (G) et droite (D) de la lumière, s'exprime selon la relation :

$$\Delta\varepsilon = \varepsilon_G - \varepsilon_D = \frac{A_G - A_D}{l C}$$

En pratique, on peut également définir l'ellipticité molaire  $[\theta]$  (en  $\text{degré} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$ ) et pour une protéine, l'ellipticité molaire par résidu ( $[\theta]/(\text{nombre de résidus})$ ) :

$$[\theta] = 3300 \Delta\varepsilon$$

Dans le cas de peptides et de protéines, le groupe amide est le chromophore dont la contribution est la plus importante, avec une absorption dans la zone 180-250 nm. C'est la géométrie des liaisons peptidiques les unes par rapport aux autres qui détermine la forme du signal, chaque structure secondaire présente donc un spectre dichroïque caractéristique (Figure 20). En considérant que les contributions des différents types de structure sont additives lors de l'étude du spectre d'une protéine ou d'un peptide, le pourcentage de chaque structure dans la conformation peut être calculé. Différents programmes ont ainsi été développés pour réaliser cette décomposition (ou déconvolution) des spectres [124].



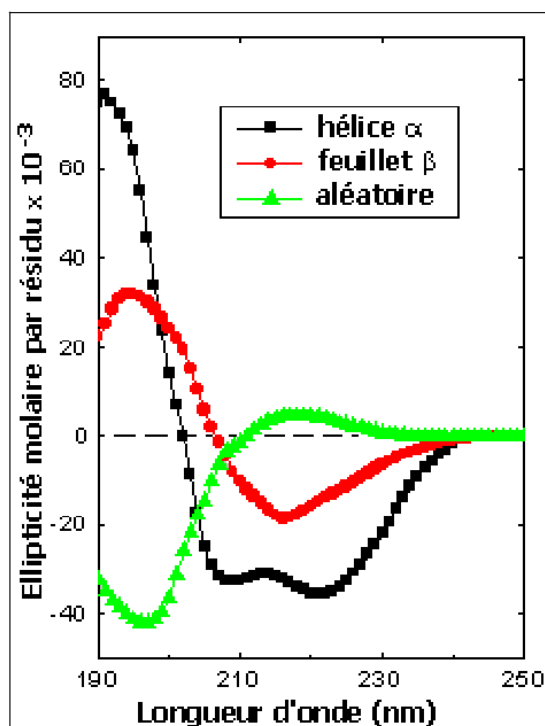


Figure 20 : Courbes de références [125] pour les principaux types de structure secondaire (hélice  $\alpha$ , feuillet  $\beta$ , conformation aléatoire).

#### ▪ Etude de la protéine VVTL1

Les spectres dichroïques de la protéine VVTL1 purifiée dans le cadre de la première partie de notre travail ont été réalisés sur un spectropolarimètre Chirascan (Applied Photophysics) à l'Université de Lyon (IBCP UMR5086) avec un trajet optique de 0.1 cm. La protéine a été diluée dans de l'eau ultrapure (8.8  $\mu$ M) puis les spectres ont été enregistrés de 185 à 255 nm à différentes températures. La déconvolution des spectres pour la quantification des structures secondaires a été réalisée par le logiciel Dichroweb [126].

### 2.1.3 Structure tertiaire : méthodes *in silico*

La bioinformatique structurale fournit des outils très puissants pour la détermination de la structure tertiaire des protéines, ainsi que pour l'étude de son comportement dans un environnement donné. Nous avons ainsi utilisé la modélisation par homologie afin de construire un modèle de la structure tridimensionnelle de la protéine VVTL1, basé sur les structures connues de protéines thaumatin-like d'autres espèces végétales. Ce modèle a ensuite été minimisé et solvato dans une boîte d'eau pour évaluer son comportement dynamique et sa stabilité au cours d'une simulation de 150 ns. L'analyse des modes normaux du modèle construit a d'autre part permis d'obtenir des informations sur la flexibilité conformationnelle de la protéine, ouvrant la voie à la compréhension de son mode d'action enzymatique.

### ▪ Modélisation moléculaire par homologie

---

La modélisation d'une protéine par homologie repose sur le principe que des séquences très proches (fort taux d'homologie) vont présenter des structures 3D similaires. L'existence préalable de protéines homologues dont la structure est résolue conditionne ainsi souvent la pertinence de la construction d'un modèle d'une protéine dont la séquence seule est connue. Cette méthode est donc particulièrement adaptée lorsqu'il s'agit de travailler sur une protéine appartenant à une famille déjà relativement bien étudiée, ce qui est le cas pour les protéines thaumatin-like.

Un modèle de la protéine VVTL1 a ainsi été construit selon un protocole de modélisation par homologie se divisant en quatre étapes successives :

- Sélection des protéines homologues : utilisation du programme BLASTP (NCBI, [127]) afin de sélectionner un ensemble de cinq thaumatin-like homologues (*templates*) puis récupération des structures cristallographiques à haute résolution de chaque protéine (Protein Data Bank [128]).
- Alignement des séquences : utilisation du programme CLUSTALW2 [129] pour réaliser l'alignement de la protéine VVTL1 (*target*) et des protéines homologues.
- Construction des modèles : utilisation du logiciel Modeller 9v7 [130], basé sur la satisfaction des contraintes spatiales et suivant un protocole de recuit simulé, afin de générer 50 modèles de la protéine VVTL1. Modeller permet d'extraire des contraintes de distance des structures templates puis de les appliquer à la séquence cible sur la base des alignements.
- Sélection et validation d'un modèle : le choix du meilleur modèle a été basé sur la fonction objective de Modeller (Discrete Optimized Protein Energy score [131]) et la qualité de la géométrie de la conformation a été vérifiée par les logiciels PROCHECK [132] et RamaDA (développé par Matthieu Tanty et Marc-André Delsuc, accessible à l'adresse <http://ramada.u-strasbg.fr/>), permettant de placer tous les acides aminés de la protéine sur le diagramme de Ramachandran.

### ▪ Dynamique moléculaire

---

#### *Stabilité dans l'eau au cours du temps*

Au-delà de la structure tridimensionnelle d'une protéine, il est important d'étudier sa stabilité et son comportement dynamique. Ainsi, dans un premier temps, le modèle de la protéine VVTL1 construit par homologie a été minimisé, solvato dans une boîte d'eau périodique et neutralisé par des contre-ions. A partir de cet état, plusieurs simulations de dynamique moléculaire ont été réalisées pour une durée totale de 150 ns. Ces simulations ont été exécutées en utilisant le logiciel NAMD [133], et les résultats visualisés sous le logiciel VMD [133]. L'ensemble des simulations numériques a été effectué sur le cluster de calcul CLOVIS de l'URCA.

La trajectoire obtenue a ensuite été analysée en calculant l'écart quadratique moyen des carbones du squelette peptidique (« Root Mean Square Deviation » : RMSD). Le RMSD permet de comparer la structure obtenue au temps  $t$  à la structure initiale après alignement dans l'espace de leurs atomes respectifs. Plus cette valeur est faible, plus les structures sont proches alors que des valeurs élevées indiquent des régions très flexibles.

Le RMSD entre deux structures  $i$  et  $j$  possédant un nombre d'atomes  $N$  est calculé selon :

$$RMSD(i, j) = \sqrt{\frac{1}{N} \sum_{n=1}^N \|\vec{r}_{i,n} - \vec{r}_{j,n}\|^2}$$

avec  $\vec{r}_{i,n}$  la position de l'atome  $n$  dans la conformation  $i$ .

### Flexibilité conformationnelle

Dans un second temps, des simulations conduisant aux modes normaux ont été réalisées pour tenter de comprendre les mouvements plus amples de la structure, souvent corrélés aux mouvements fonctionnels.

L'analyse des modes normaux de basse fréquence d'une protéine consiste à déterminer les modes de vibration de la molécule, indiquant des déplacements atomiques dans toutes les directions de l'espace et révélant ainsi la flexibilité conformationnelle du système. Ces changements de conformation peuvent se révéler particulièrement intéressants dans le cas notamment de la fixation d'un substrat à une enzyme [134].

Deux webserveurs ont été utilisés en parallèle pour réaliser l'étude des modes normaux du modèle construit de la protéine VVTL1 :

- « the elastic network model » proposé par eINémo [135], pour la réalisation du mode 7 avec une gamme de perturbation -40/+40 %.
- WEBnm@ [136], pour l'étude complète des modes 7 à 12 avec tous les paramètres par défaut.

## 2.2 Etude d'un peptide par modélisation moléculaire sous contraintes RMN

### 2.2.1 Sélection et synthèse du peptide

Un peptide de 13 acides aminés (en bleu) a été sélectionné dans la séquence de VVTL1 (résidus 158-170) :

ATFDILNKCTYTVWAAASPGGGRRLD SGQSWTITVNP GTTNARIWGR TSCTFDANGRGK CETGDCNGL  
LECGYGSPPNTLAEFALNQPNNDYIDISLVDGFNIPMDFSGCRGIQCSVDINGQCPSKAPGGC  
CTVFKTNEYCCTDGP GSCG **PTTYSKFFKDRCP** DAYSY PQDDKTSLFTCPSGTNYKVTFCP

Ce peptide, désigné par la suite sous le nom de « PP13 », a été synthétisé chimiquement (pureté > 98%) sans marquage isotopique par l'entreprise Protéogenix (Strasbourg).

### 2.2.2 Modélisation de novo

Dans la structure modélisée de VVTL1, le peptide PP13 sélectionné présente une structure hélice  $\alpha$ . Afin de confirmer que cette structuration est stabilisée intrinsèquement et qu'elle ne résulte pas d'interactions avec le reste de la séquence/structure de VVTL1, le peptide a été modélisé en premier lieu par une approche *de novo*.

A la différence de la modélisation par homologie utilisée pour prédire la structure de VVTL1 en se basant sur des structures déjà résolues, la modélisation *de novo* permet de construire une structure protéique à partir de la séquence seule.

Les méthodes *de novo* sont à distinguer des méthodes *ab initio* dites pures, basées uniquement sur les lois physiques régissant le repliement des protéines, car elles exploitent des informations statistiques obtenues par l'analyse de structures 3D connues.

La prédiction de structure *de novo* peut ainsi être réalisée à l'aide d'alphabets structuraux [137]. Ce concept repose sur l'observation que les structures protéiques sont constituées d'un nombre limité de structures locales récurrentes. L'exploration de structures connues permet ainsi de constituer une bibliothèque de fragments protéiques couvrant l'ensemble de ces structures locales et formant ainsi un ensemble de blocs structuraux élémentaires (alphabet structural). La structure tertiaire est alors prédite à partir de la combinaison des fragments qui va satisfaire au mieux aux différentes contraintes (énergie minimale, enfouissement des résidus hydrophobes, appariement des brins  $\beta$ ). L'exploration de l'espace conformationnel est ainsi beaucoup plus rapide que pour les méthodes *ab initio* pures car la taille de ce dernier est réduite.

Le serveur PEP-FOLD (portail RPBS Mobyly accessible à l'adresse <http://mobyly.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py?form=PEP-FOLD#>) a été utilisé pour prédire la structure du peptide PP13. PEP-FOLD [138, 139] utilise l'alphabet structural HMM-SA (Hidden Markov Model Structural Alphabet) composé de 27 lettres contenant des fragments de 4 acides aminés [140-142]. Les fragments sont assemblés par un algorithme glouton basé sur une stratégie locale de reconstruction utilisant le potentiel énergétique OPEP (Optimized Potential for Efficient structure Prediction) [143, 144]. Les paramètres de ce champ de forces sont développés dans l'Annexe 2.2.3.

### 2.2.3 Dichroïsme circulaire

---

Les peptides courts possédant une faible capacité à se structurer spontanément, une première approche expérimentale a été menée par dichroïsme circulaire afin de déterminer le solvant ou mélange de solvants structurant optimal pour l'analyse par RMN [145-148].

Selon le même principe que celui présenté dans la section 2.1.2, les spectres dichroïques ont été réalisés sur un spectropolarimètre JASCO J810 à l'Université d'Orléans (Centre de Biophysique Moléculaire, CNRS) avec un trajet optique de 1 mm. Le peptide PP13 a été dilué à 0.2 mg/mL (concentration déterminée par la mesure d'absorbance à 280 nm) dans différents solvants (eau, trifluoroéthanol (TFE) et méthanol (MeOH) purs ou en mélange) à 20°C :

- H<sub>2</sub>O
- H<sub>2</sub>O/TFE 60:40 (v/v)
- H<sub>2</sub>O/MeOH 50:50 (v/v)
- MeOH

L'analyse des spectres dichroïques de PP13 obtenus a été réalisée par le logiciel CDpro [149]. Les spectres ont été déconvolués selon les trois programmes CONTIN, SELCON et CDSSTR. La qualité de la reconstruction a été estimée en comparant au spectre original. Les mauvaises reconstructions n'ont pas été prises en compte. Pour chaque spectre au moins deux analyses correctes ont été moyennées.

## 2.2.4 Expériences RMN

### Principe général

La RMN repose sur le comportement du moment magnétique des noyaux atomiques de spin non nul (comme les noyaux  $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}$ ) sous l'influence d'un champ magnétique statique intense, d'un rayonnement électromagnétique externe et des champs locaux induits par les noyaux environnants. C'est l'observation de l'interaction entre des noyaux voisins, en termes de proximité dans la séquence ou dans l'espace lors du repliement, qui va ainsi permettre de caractériser la structure d'une protéine ou d'un peptide. La réalisation de spectres RMN mono- et bidimensionnels permet d'obtenir des informations sur la géométrie de la structure, par une liste de distances interatomiques H-H (intra- ou inter-résidus) et de mesures des angles dièdres  $\phi$  et  $\psi$ . Ces informations représentent les contraintes qui sont ensuite utilisées lors de la modélisation sous contraintes RMN.

Le travail préalable à l'affectation des contraintes géométriques structurant une protéine ou un peptide est l'attribution des déplacements chimiques de chaque atome d'hydrogène et de carbone. Pour cela, nous avons mis en œuvre différentes expériences 2D homo- ( $^1\text{H}$ - $^1\text{H}$ ) et hétéronucléaires ( $^1\text{H}$ - $^{13}\text{C}$  et  $^1\text{H}$ - $^{15}\text{N}$ ), adaptées à l'étude d'un petit peptide sans marquage isotopique (abondance naturelle de  $^{13}\text{C}$  et  $^{15}\text{N}$ ).

### Attribution des déplacements chimiques $^1\text{H}$ , $^{13}\text{C}$ et $^{15}\text{N}$

#### ➤ Spectre 1D $^1\text{H}$

Chaque proton possède un déplacement chimique propre dont la valeur dépend de son environnement chimique. L'influence principale provient du groupement auquel il est attaché, permettant de définir des zones de déplacement chimique attendues pour les protons de la chaîne principale et des chaînes latérales selon des tables théoriques de déplacements chimiques (Figure 21) [150].

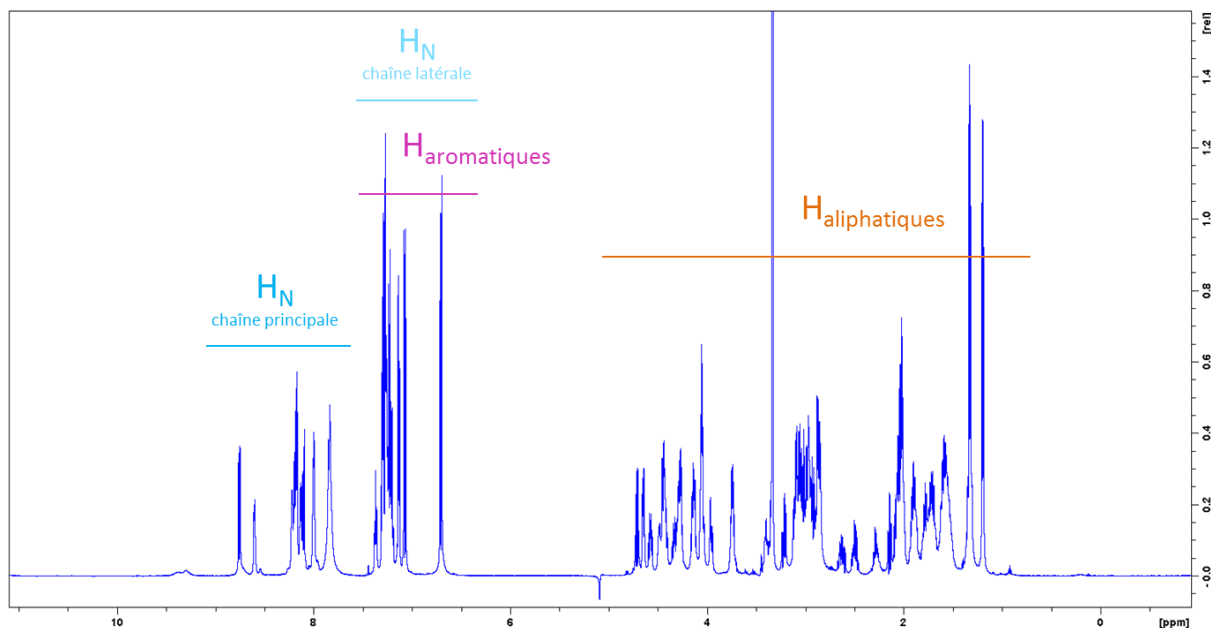


Figure 21 : Zones caractéristiques des déplacements chimiques  $^1\text{H}$  lors de l'analyse d'une protéine ou d'un peptide. Le spectre représenté est celui du peptide PP13.

Par ailleurs, les protons portés par un même atome (séparation par 2 liaisons : H géminés sur les groupements  $\text{CH}_2$  ou  $\text{NH}_2$ ) ou des atomes adjacents (séparation par 3 liaisons : H vicinaux C/N H – CH) induisent des couplages se traduisant par une multiplicité du signal sur un spectre 1D  $^1\text{H}$ . Tout proton couplant avec un proton voisin se traduit par un dédoublement des raies, l'intensité relative des pics suivant le triangle de Pascal (Figure 22).

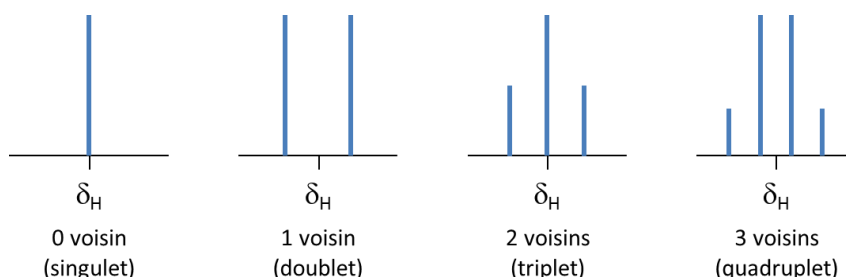


Figure 22 : Multiplicité des signaux  $^1\text{H}$  du singulet au quadruplet pour des constantes de couplage identiques.

La triple information fournie par un spectre 1D  $^1\text{H}$  (valeur du déplacement chimique dans une zone caractéristique, multiplicité du signal et aire des résonances) donne une première indication dans l'attribution des protons d'une molécule mais n'est pas suffisante lors de l'étude d'un peptide en raison de la superposition des signaux sur le spectre, formant souvent des massifs non-identifiables. Il faut donc avoir recours à des spectres à deux dimensions afin d'améliorer la résolution et d'obtenir des informations complémentaires facilitant l'attribution sur les couplages homo- ou hétéronucléaires.

### ➤ Spectres 2D

L'utilisation d'un spectre à deux dimensions permet l'observation de taches de corrélation traduisant un couplage entre deux atomes du peptide ( $^1\text{H}$ ,  $^{13}\text{C}$  ou  $^{15}\text{N}$ ). La Figure 23 illustre ce principe en montrant deux protons  $\text{H}_\text{A}$  et  $\text{H}_\text{B}$  de déplacements chimiques  $\delta_{\text{H}_\text{A}}$  et  $\delta_{\text{H}_\text{B}}$  (en ppm) qui corrélient entre eux (taches vertes symétriques). Les taches de corrélation  $\text{H}_\text{A}$ - $\text{H}_\text{A}$  et  $\text{H}_\text{B}$ - $\text{H}_\text{B}$  sur la diagonale (taches bleues) sont également observables.

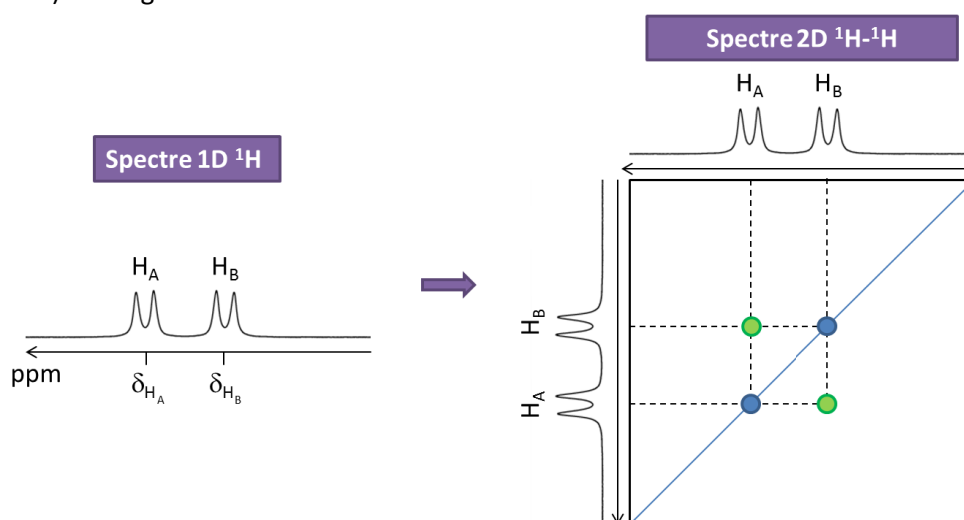


Figure 23 : Observation du couplage entre deux protons  $\text{H}_\text{A}$  et  $\text{H}_\text{B}$  lors d'une expérience 2D homonucléaire  $^1\text{H}$ - $^1\text{H}$ .

Les régions caractéristiques de déplacements chimiques identifiées sur les spectres 1D  $^1\text{H}$  se retrouvent lors de l'étude des spectres 2D  $^1\text{H}$ - $^1\text{H}$ , permettant de définir des zones de couplages (Figure 24).

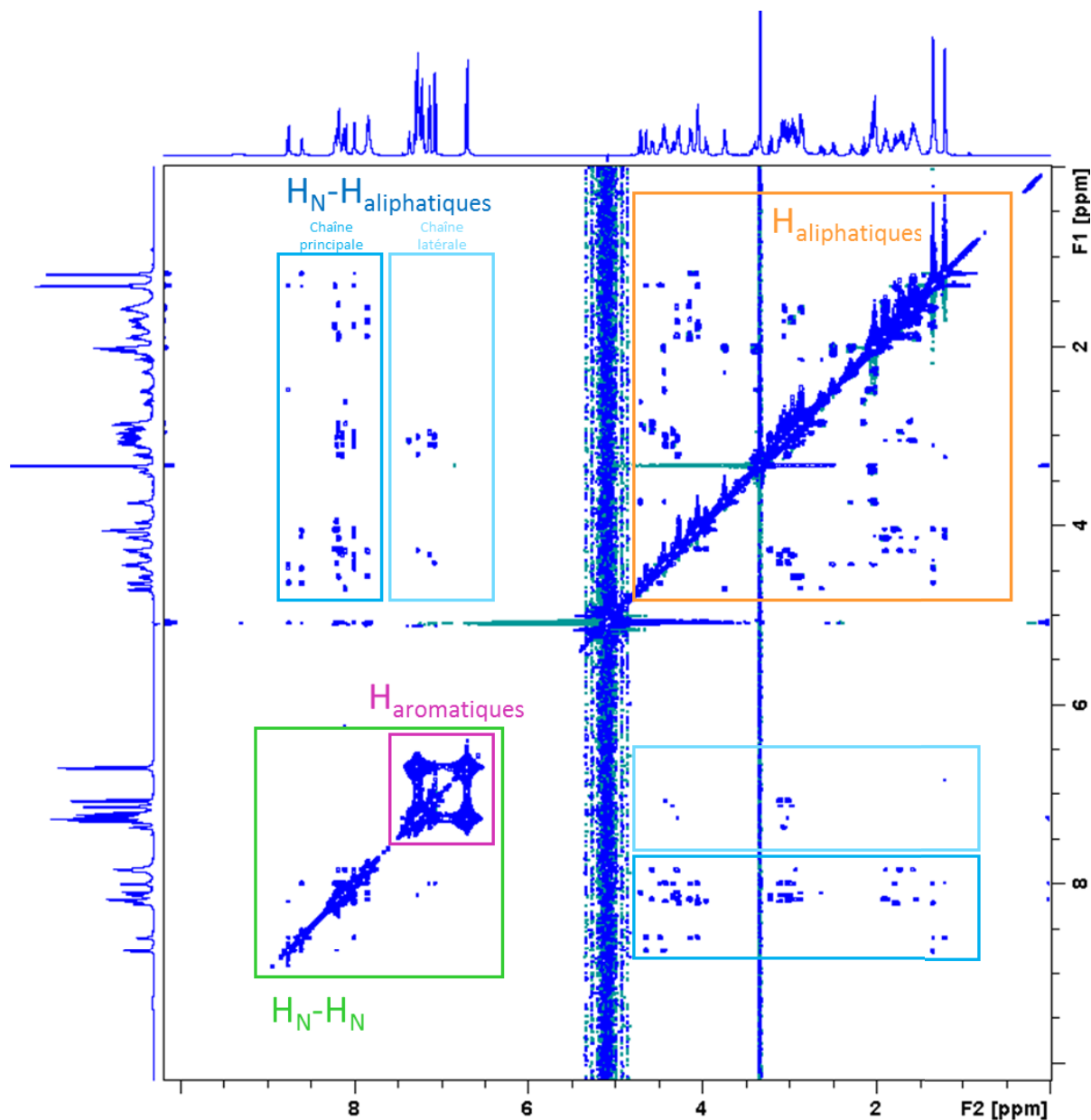


Figure 24 : Zones caractéristiques des déplacements chimiques  $^1\text{H}$  et des couplages  $^1\text{H}$ - $^1\text{H}$  lors de l'analyse d'une protéine ou d'un peptide. Le spectre représenté est le spectre NOESY du peptide PP13.

Les expériences 2D homonucléaires COSY (COrrélation SpectroscopY) [151], TOCSY (TOtal Correlated SpectroscopY) [152] et NOESY (Nuclear Overhauser Enhancement SpectroscopY) [153] permettent d'observer les couplages  $^1\text{H}$ - $^1\text{H}$ , tandis que les expériences 2D hétéronucléaires HSQC (Heteronuclear Single Quantum Correlation spectroscopy) [154] et HMBC (Heteronuclear Multiple Bond Correlation spectroscopy) [155] permettent d'observer les couplages  $^1\text{H}$ - $^{13}\text{C}$  ou  $^1\text{H}$ - $^{15}\text{N}$ . Selon l'expérience, les couplages peuvent être scalaires (relais de l'aimantation à travers les liaisons covalentes) ou dipolaires (relais de l'aimantation à travers l'espace) et donner des informations sur les contacts intra- ou inter-résidus.

L'observation de toutes les taches de corrélation intra-résidu sur les spectres COSY et TOCSY permet d'identifier la position des atomes  $^1\text{H}$  dans chaque résidu (identification des systèmes de spins). Ces deux expériences, basées sur les couplages scalaires  $^n\text{J}$ , permettent d'identifier les protons d'un même acide aminé qui corrélient entre eux à travers les liaisons. Dans le cadre d'une expérience COSY, seuls les couplages entre deux protons géminés ( $n = 2$ ) ou vicinaux ( $n = 3$ ) sont observés, tandis qu'une expérience TOCSY permet d'obtenir les corrélations entre tous les protons d'un même acide aminé ( $n = 2,3,4,5\dots$ ). Aucun couplage scalaire entre les protons de deux résidus voisins ne peut avoir lieu car le carbonyle de la liaison peptidique empêche le relais d'aimantation. Les 20 acides aminés produisent des systèmes de spins caractéristiques représentés sur la Figure 25.

Une fois tous les systèmes de spins identifiés, l'expérience NOESY permet de déterminer l'enchaînement séquentiel de ces systèmes par les contacts inter-résidus ( $i$  et  $i+1$ ). Cette expérience est basée sur l'effet Overhauser nucléaire (NOE), qui correspond à un transfert d'aimantation entre deux spins proches dans l'espace (distance  $< 6 \text{ \AA}$ ). Un spectre NOESY met ainsi en évidence les corrélations dipolaires  $^1\text{H}$ - $^1\text{H}$  d'un même système de spins (corrélation intra-résidu) mais également les corrélations dipolaires séquentielles  $d_{\text{H}\alpha\text{-HN}}(i,i+1)$ ,  $d_{\text{HN-HN}}(i,i+1)$  et  $d_{\text{H}\beta\text{-HN}}(i,i+1)$ . L'effet NOE séquentiel permet donc d'observer les corrélations inter-résidus, donnant une information sur la localisation des acides aminés dans la séquence peptidique. En cas de repliement de la structure, un effet NOE est également observé entre des  $^1\text{H}$  proches en termes de proximité spatiale indépendamment de la proximité séquentielle.

En connaissant au préalable la séquence du peptide étudié, la confrontation des trois expériences COSY, TOCSY et NOESY représente ainsi la stratégie de Wüthrich [156] qui permet d'attribuer tous les déplacements chimiques  $^1\text{H}$  et d'effectuer l'enchaînement séquentiel. Les expériences hétéronucléaires HSQC et HMBC ( $^1\text{H}$ - $^{13}\text{C}$  et  $^1\text{H}$ - $^{15}\text{N}$ ) basées sur les couplages scalaires peuvent venir en complément pour lever des ambiguïtés sur l'attribution de signaux  $^1\text{H}$  lorsque des taches se superposent sur les spectres 2D homonucléaires et fournissent l'attribution des déplacements chimiques  $^{13}\text{C}$  et  $^{15}\text{N}$ .

Un résumé des expériences RMN 2D donnant leurs caractéristiques et utilités est présenté dans le Tableau 6.



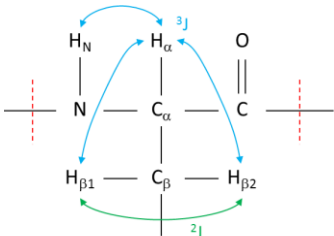
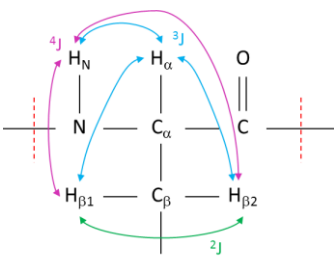
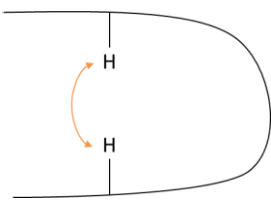
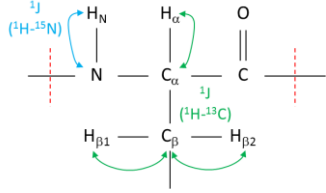
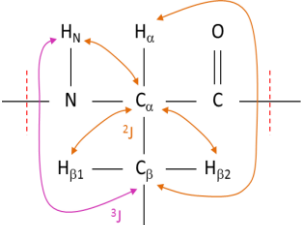
Expérience 2D		Couplage observé	Information obtenue
Homonucléaire	<b>COSY</b> $^1\text{H}-^1\text{H}$	<b>Scalaire intra-résidu</b> <i>relais de l'aimantation à travers 2 ou 3 liaisons d'un résidu (<math>^2J</math> et <math>^3J</math>)</i> 	Identification des systèmes de spins (déplacements chimiques de tous les $^1\text{H}$ de chaque acide aminé)
	<b>TOCSY</b> $^1\text{H}-^1\text{H}$	<b>Scalaire intra-résidu</b> <i>relais de l'aimantation à travers toutes les liaisons du résidu (<math>^nJ</math>, <math>n \geq 2</math>)</i> 	
Hétéronucléaire	<b>NOESY</b> $^1\text{H}-^1\text{H}$	<b>Dipolaire</b> <b>intra- et inter-résidu</b> <i>transfert d'aimantation en cas de distance interatomique <math>&lt; 6 \text{ \AA}</math></i> 	Enchaînement des systèmes de spins (reconstruction de la séquence peptidique)
	<b>HSQC</b> $^1\text{H}-^{13}\text{C}$ ou $^1\text{H}-^{15}\text{N}$	<b>Scalaire intra-résidu</b> <i>corrélation <math>^{13}\text{C}</math> ou <math>^{15}\text{N}</math> et <math>^1\text{H}</math> liés (<math>^1J</math>)</i> 	Aide à la détermination des $^1\text{H}$ et attribution des déplacements chimiques $^{13}\text{C}$ et $^{15}\text{N}$
	<b>HMBC</b> $^1\text{H}-^{13}\text{C}$	<b>Scalaire intra-résidu</b> <i>corrélation <math>^1\text{H}</math> et <math>^{13}\text{C}</math> non-liés (<math>^nJ</math>, <math>n \geq 2</math>)</i> 	

Tableau 6 : Résumé des différentes expériences RMN 2D réalisées pour déterminer l'attribution des déplacements chimiques du peptide PP13.

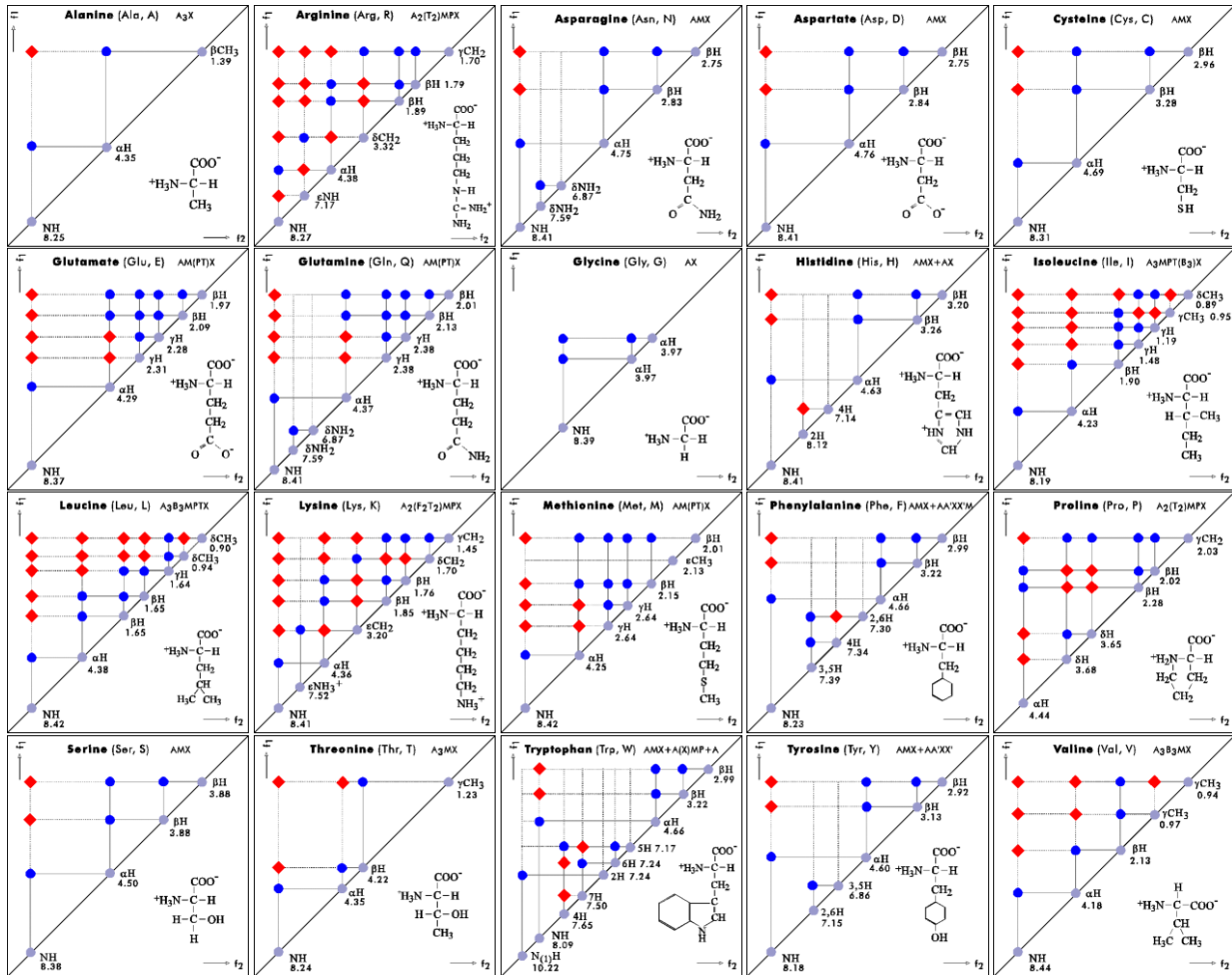


Figure 25 : Pics COSY (●) et TOCSY (◆) des 20 acides aminés.

### Détermination des contraintes structurales

Une fois tous les déplacements chimiques attribués, les paramètres structuraux du peptide étudié peuvent être dévoilés. De multiples informations peuvent être potentiellement tirées de l'analyse de spectres RMN, mais nous nous concentrerons sur trois approches principales, suffisantes pour résoudre la structure d'un peptide de taille modérée lorsqu'elles sont combinées. Les indices de déplacements chimiques en premier lieu fournissent une indication sur la probabilité qu'un acide aminé d'être impliqué dans un type de structure secondaire donné. Ensuite, la mesure des constantes de couplage  $^3J_{\text{HN-H}\alpha}$  (obtention des valeurs d'angle dièdre) et la mesure des effets NOE (obtention des valeurs de distance inter-protons) permettent de dresser une liste de contraintes expérimentales qui sont utilisées pour la construction d'un modèle 3D du peptide.

#### ➤ Indices de déplacements chimiques

La notion d'indices de déplacements chimiques repose sur l'observation que ces déplacements sont très sensibles à leur environnement. Dans le cas des protéines, l'étude statistique de structures protéiques résolues montre que la nature de la structure secondaire dans laquelle est impliqué un atome influence la valeur de son déplacement chimique, en particulier dans le cas des  $H_\alpha$  et des  $C_\alpha$  [150]. Les indices de déplacements chimiques, ou CSI (pour Chemical Shift Index) sont

alors définis comme l'écart entre la valeur expérimentale d'un déplacement et un déplacement de référence déterminé pour chaque atome des 20 acides aminés dans le cas de peptides courts non-structurés [157].

Une structure de type hélice est probable lorsque le CSI  $^1\text{H}$  ( $\delta_{\text{H}\alpha\text{exp}} - \delta_{\text{H}\alpha\text{ref}}$ ) est inférieur à -0.1 ppm pour quatre acides aminés consécutifs, tandis qu'une structure de type feuillet est indiquée lorsque la différence est supérieure à 0.1 ppm. Un indice  $^1\text{H}$  compris entre -0.1 et 0.1 ppm suggère la présence d'une région non-structurée [158]. Il en est de même pour le CSI  $^{13}\text{C}$  ( $\delta_{\text{C}\alpha\text{exp}} - \delta_{\text{C}\alpha\text{ref}}$ ), qui traduit une structure hélice lorsqu'il est supérieur à 1 ppm, l'hypothèse d'une structure feuillet lorsqu'il est inférieur à -1 ppm et pas de structuration entre -1 et 1 ppm [159]. Le calcul du CSI pour tous les  $\text{H}_\alpha$  et des  $\text{C}_\alpha$  des acides aminés de la séquence du peptide étudié permet ainsi d'établir un profil potentiel d'enchaînement d'éléments de structures secondaires.

### ➤ Constantes de couplage

La mesure d'une constante de couplage  $^3J_{\text{HN-H}\alpha}$  est réalisée sur le spectre 1D  $^1\text{H}$ , définie par la distance en Hz entre les deux sommets du doublet représentant le signal du  $\text{H}_\text{N}$ . L'angle dièdre de torsion  $\phi$  entre les liaisons  $\text{H-N-C}_\alpha\text{-H}_\alpha$  peut être déduit de cette constante par la relation de Karplus :

$$^3J_{\text{HN-H}\alpha} = A \cos^2\theta - B \cos\theta + C$$

avec  $\theta = |\phi - 60|$ . Les valeurs empiriques des coefficients A, B et C diffèrent selon les auteurs [160-162], les plus souvent utilisés étant les coefficients de Pardi (A = 6.4, B = 1.4, C = 1.9) [163]. A partir de la courbe représentative de l'équation correspondante (Figure 26) et en tenant compte des contraintes stériques (diagramme de Ramachandran), on peut définir des intervalles de valeurs permises pour l'angle  $\phi$ . Ainsi, lorsque la constante de couplage  $^3J$  est supérieure à 9 Hz, l'angle  $\phi$  est contraint dans l'intervalle  $-120 \pm 30^\circ$  (caractérisant une structure secondaire de type feuillet) et lorsqu'elle est inférieure à 6 Hz l'angle est contraint dans l'intervalle  $-60 \pm 30^\circ$  (caractérisant une structure secondaire de type hélice). De la même manière que le calcul des CSI, la mesure de la constante de couplage  $^3J_{\text{HN-H}\alpha}$  de chaque acide aminé de la séquence du peptide étudié permet ainsi d'établir un profil potentiel d'enchaînement d'éléments de structures secondaires. On note cependant que la proline ne peut pas être caractérisée car cet acide aminé ne possède pas de  $\text{H}_\text{N}$  lorsqu'il est engagé dans une liaison peptidique.

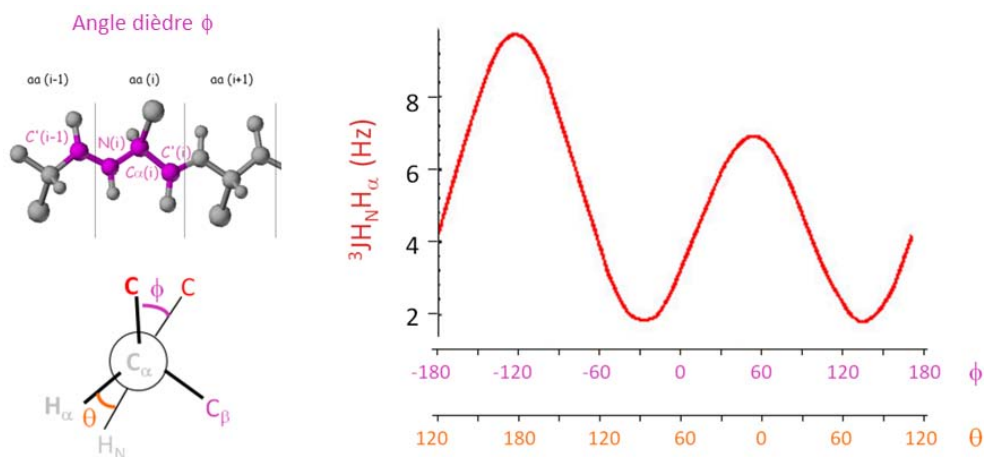


Figure 26 : Variation de la constante de couplage  $^3J_{\text{HN-H}\alpha}$  en fonction de l'angle  $\phi$  déterminée avec les coefficients de Pardi.

➤ **Mesure des distances inter-protons**

Chaque tache présente sur le spectre NOESY indique un couplage entre deux  $^1\text{H}$  voisins (distance inférieure à 6 Å). L'intensité  $I$  de la tache (volume intégré) est proportionnelle à la distance  $d$  entre les deux  $^1\text{H}$  corrélés :

$$I \propto \frac{1}{d^6}$$

Un couple intégrale/distance de référence (protons géminés  $d = 1.8 \text{ \AA}$  ou protons aromatiques  $d = 2.46 \text{ \AA}$ ) est défini afin d'établir le coefficient de proportionnalité et ainsi convertir toutes les corrélations du spectre en distances par la relation  $d = d_{ref} * (I_{ref}/I)^{1/6}$ .

Trois gammes d'effet NOE peuvent être caractérisées :

- effet NOE fort :  $1.8 \leq d < 3.0$
- effet NOE moyen :  $1.8 \leq d < 3.5$
- effet NOE faible :  $1.8 \leq d < 6.0$

Chaque type de structure secondaire est défini par des distances inter-protons caractéristiques (Tableau 7) [164]. Les motifs structuraux sont donc facilement repérables sur un diagramme de connectivités NOE (Tableau 8).

	Hélice $\alpha$	Hélice $3_{10}$	Feuillet $\beta$ antiparallèle	Feuillet $\beta$ parallèle	Coude $\beta$ de type I	Coude $\beta$ de type II
$d\alpha\text{N}(i, i+1)$	3,5	3,4	2,2	2,2	3,4 3,2	2,2 3,2
$d\alpha\text{N}(i, i+2)$	4,4	3,8			3,6	3,3
$d\alpha\text{N}(i, i+3)$	3,4	3,3			3,1-4,2	3,8-4,7
$d\alpha\text{N}(i, i+4)$	4,2					
$d\text{NN}(i, i+1)$	2,8	2,6	4,3	4,2	2,6 2,4	4,5 2,4
$d\text{NN}(i, i+2)$	4,2	4,1			3,8	4,3
$d\beta\text{N}(i, i+1)$	2,5-4,1	2,9-4,4	3,2-4,5	3,7-4,7	2,9-4,4	3,6-4,6
$d\alpha\beta(i, i+3)$	2,5-4,4	3,1-5,1				

Tableau 7 : Distances inter-protons (en Å) caractéristiques pour les principaux types de structures secondaires [164].

	Hélice $\alpha$	Hélice $3_{10}$	Feuillet $\beta$	Coude $\beta$ de type I	Coude $\beta$ de type II
$d\text{NN}(i, i+1)$					
$d\alpha\text{N}(i, i+1)$					
$d\alpha\text{N}(i, i+2)$					
$d\alpha\text{N}(i, i+3)$					
$d\alpha\text{N}(i, i+4)$					
$d\alpha\beta(i, i+3)$					
$d\text{NN}(i, i+2)$					
Résidus	1 2 3 4 5 6 7	1 2 3 4 5 6	1 2 3 4 5 6	1 2 3 4	1 2 3 4

Tableau 8 : Diagramme des connectivités NOE pour les principaux types de structures secondaires [164]. L'intensité des barres reflète l'intensité de l'effet NOE (fort, moyen, faible).

### ▪ Réalisation des spectres

Tous les spectres RMN ont été réalisés sur un spectromètre Bruker AVANCE III 600 MHz équipé d'une cryosonde 5 mm TCI  $^2\text{H}/^1\text{H}-^{13}\text{C}-^{15}\text{N}$ . Le peptide PP13 a été dissous à une concentration de 5 mM dans 600  $\mu\text{L}$  de solvant deutéré ( $\text{CD}_3\text{OH}$ ).

Des spectres 1D  $^1\text{H}$  ont été enregistrés avant et après chaque série d'expériences 2D (COSY, TOCSY, NOESY, HSQC, HMBC) afin de s'assurer que la structure du peptide étudié reste inchangée au cours du temps.

#### *Suppression du signal du solvant*

Lors d'une expérience RMN  $^1\text{H}$ , on cherche à observer les protons du soluté. Cependant, le solvant (aqueux ou organique) dans lequel est dissous le soluté contient lui aussi des protons, qui vont se superposer à ceux du soluté sur les spectres en produisant des signaux très intenses. L'utilisation de solvants deutérés permet de minimiser ce phénomène mais deux limites se posent : d'une part les produits deutérés ne sont jamais à un taux d'enrichissement isotopique de 100% donc produisent toujours un signal  $^1\text{H}$  résiduel (d'où l'observation du signal  $-\text{CH}_3$  du méthanol), d'autre part la capacité d'échange entre les protons du soluté et les deutériums du solvant impliquerait la disparition des signaux du soluté en cas de solvant totalement deutéré.

Les solvants organiques et l'eau produisent donc des signaux très intenses sur les spectres RMN, sources de distorsions de la ligne de base et pouvant recouvrir les signaux de l'échantillon analysé. Il faut donc impérativement les supprimer avant de réaliser des spectres exploitables.

Lors de l'analyse de PP13, deux signaux  $^1\text{H}$  provenant du solvant méthanol sont observés :  $-\text{OH}$  vers 5.0 ppm et  $-\text{CH}_3$  vers 3.3 ppm. Le signal le plus important à éliminer est celui du groupement hydroxyle, identique à celui observé dans le cas du signal de l'eau.

La température est le premier facteur qui peut être modifié pour minimiser la superposition du signal  $-\text{OH}$  sur les signaux du peptide. La Figure 27 montre la région du signal  $-\text{OH}$  sur les spectres  $^1\text{H}$  de PP13 à 278 K et à 288 K. Le pic du méthanol se décale de 5.0 à 5.1 ppm lorsque la température diminue, améliorant la qualité de la ligne de base pour les  $\text{H}_\alpha$  très proches (4.8 ppm). La température de 278 K a donc été choisie pour la réalisation de tous les spectres RMN.

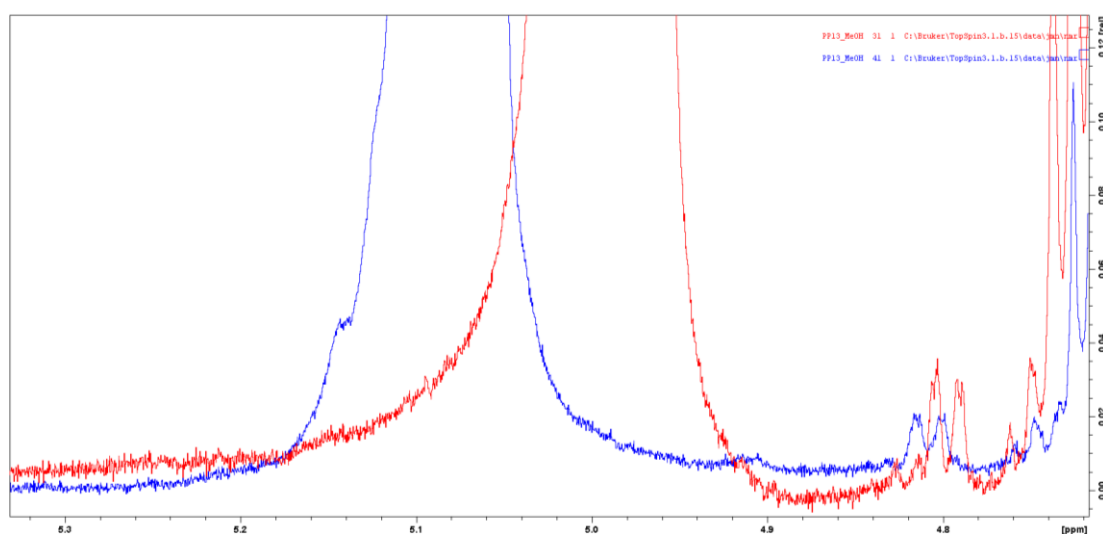


Figure 27 : Spectres 1D  $^1\text{H}$  de PP13 zoomés sur la zone du signal du  $-\text{OH}$  du méthanol à 278 K (bleu) et 288 K (rouge).

Le signal  $-OH$  du méthanol a ensuite été sélectivement supprimé par l'utilisation d'impulsions gradients de champ magnétique selon une séquence « Excitation Sculpting » (Figure 28) [165]. La Figure 29 montre les spectres avant et après suppression du signal  $-OH$  sur le spectre de PP13, révélant la qualité de la séquence utilisée.

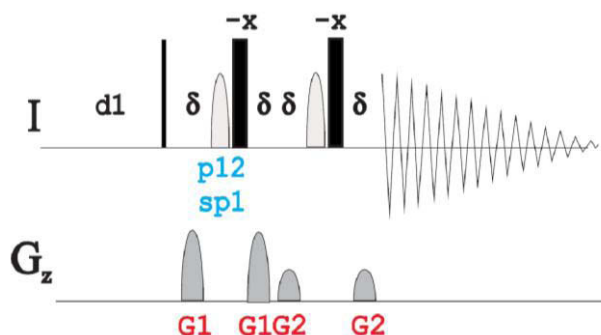


Figure 28 : Séquence d'impulsions « Excitation Sculpting » : zgesgp. Reproduit avec l'aimable autorisation de Pedro Lameiras.

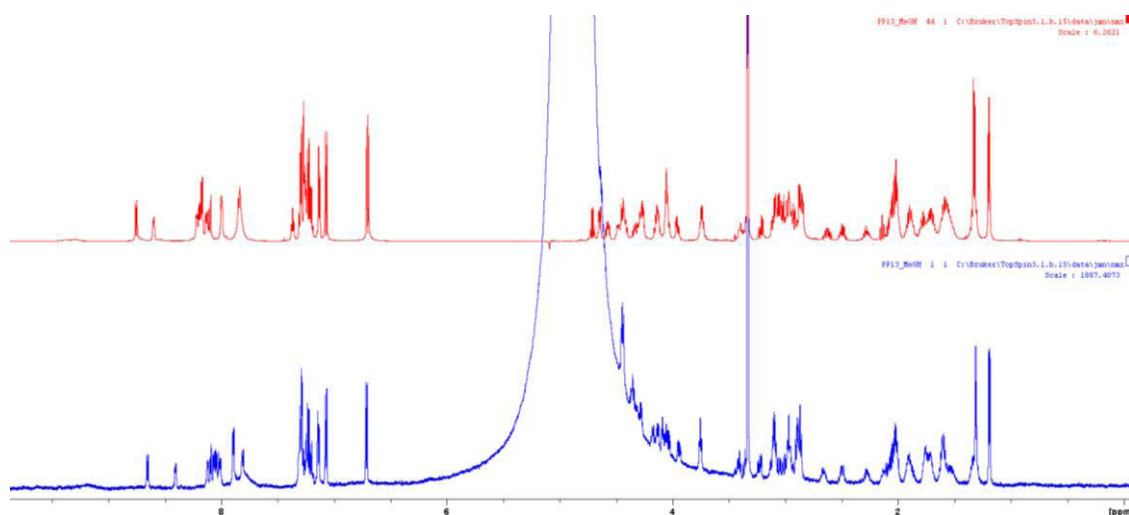


Figure 29 : Spectres  $^1H$  de PP13 avant (bleu) et après (rouge) suppression du signal du  $-OH$  du méthanol. Le pic fin intense à 3.3 ppm représente le groupement  $-CH_3$  du méthanol (non supprimé).

### Acquisition des spectres 2D

Les expériences 2D ont été réalisées selon les paramètres présentés dans le Tableau 9. Les séquences d'impulsions sont données selon la nomenclature Bruker. Les deux séquences TOCSY et NOESY ont été modifiées pour ajouter une séquence de suppression du signal du solvant (présaturation) à une séquence supprimant les interférences zéro-quantum [166]. La séquence HSQC  $^1H$ - $^{15}N$  a également été modifiée pour intégrer une présaturation.

Le temps de mélange  $\tau_m$  a été fixé à 100 ms pour les expériences TOCSY et à 120 ms pour les expériences NOESY, déterminé par NOE « build up ». Cette technique permet de minimiser l'erreur due à la diffusion de spin sur la mesure de l'intensité des effets NOE en réalisant une série de spectres NOESY avec des temps de mélange différents (de 50 à 250 ms) puis des courbes d'intensité NOE  $I$  en fonction du temps de mélange. Ces courbes réalisées sur un certain nombre de pics de corrélation NOE indiquent le temps de mélange pour lequel la diffusion de spin est négligeable, correspondant à la partie linéaire de la courbe  $I = f(\tau_m)$ , afin de choisir l'expérience NOESY à intégrer.

Expérience 2D	Séquence d'impulsions	TD (F1) x TD (F2)
COSY $^1\text{H}$ - $^1\text{H}$	cosydfesgpph	1024 x 4096
TOCSY $^1\text{H}$ - $^1\text{H}$	tocsyzqpr_pl	512 x 4096
NOESY $^1\text{H}$ - $^1\text{H}$	noesyqpr_pl	1024 x 4096
HSQC $^1\text{H}$ - $^{15}\text{N}$	hsqcetf3gpsipr_pl	256 x 1714
HSQC $^1\text{H}$ - $^{13}\text{C}$	hsqcedetgpsisp2.4	1024 x 1714
HMBC $^1\text{H}$ - $^{13}\text{C}$	hmbcedetgpl3nd	1024 x 4096

Tableau 9 : Acquisition des expériences 2D homo- et hétéronucléaires du peptide PP13.

### 2.2.5 Modélisation moléculaire sous contraintes RMN

La structure 3D d'une protéine ou d'un peptide, *ie* la position relative des atomes dans l'espace, dont on ne connaît que la séquence peut être obtenue par différents protocoles de modélisation moléculaire. Nous avons précédemment utilisé la modélisation par homologie (construction d'un modèle à partir de protéines de séquences similaires dont la structure est connue) dans la section 2.1.3 et la modélisation *de novo* (construction d'un modèle en exploitant des informations statistiques tirées d'un large éventail de structures connues) dans la partie 2.2.2. Ces deux techniques ont recours aux informations issues des banques de données de structures de protéines et n'exploitent ainsi aucune donnée expérimentale obtenue sur la protéine ou le peptide considéré. Bien que ces approches puissent fournir des résultats fiables de manière rapide, le recueil de données expérimentales peut s'avérer très utile de manière indépendante ou en complément des autres méthodes.

La réalisation d'expériences RMN permet de collecter des paramètres structuraux, en termes de distances interatomiques (via les effets NOE) et d'angles dièdres (via les constantes de couplage). Ces données représentent des contraintes structurales expérimentales qui vont être intégrées dans un processus de modélisation moléculaire. La modélisation sous contraintes RMN permet ainsi de générer des modèles de structures qui respectent à la fois les lois physiques régissant le repliement des protéines et les données expérimentales sur la géométrie de la structure et l'organisation des atomes. Les contraintes RMN de distances et d'angles sont converties sous forme de termes d'énergie et intégrés dans le protocole de génération de structures d'énergie totale minimale détaillé dans l'Annexe 2.

La modélisation sous contraintes RMN du peptide PP13 a été réalisée à l'aide du logiciel CNS (Crystallography and NMR System) [167, 168]. CNS fonctionne de manière similaire à Modeller, *ie* par exécution successive de scripts prenant en compte des fichiers d'entrée (ici séquence peptidique et listes de contraintes) et générant des fichiers de sortie (structures modélisées). Le protocole de modélisation débute par la génération de la structure étendue du peptide servant de modèle initial avant l'application des contraintes RMN qui vont induire le repliement du peptide lors de la génération des structures modélisées qui seront finalement analysées et validées.

### ■ Génération de la structure étendue initiale

A partir de la séquence peptidique, CNS génère une structure étendue du peptide par tirage aléatoire des angles dièdres  $\phi$  ( $C_{i-1}, N_i, C_{\alpha i}, C_i$ ) et  $\psi$  ( $N_i, C_{\alpha i}, C_i, N_{i+1}$ ). Pour cela, le logiciel utilise un champ de forces simplifié défini par les paramètres standard des fichiers de topologie (décrivant tous les acides aminés avec le type d'atomes, de liaison et d'angles qui les constituent ainsi que la charge partielle et la masse des atomes) et de paramètres (définissant le champ de forces à partir des valeurs d'équilibre des liaisons et des angles dièdres et impropres relatives aux différents termes énergétiques). La structure obtenue servira de point de départ pour les calculs de dynamique moléculaire par recuit simulé sous contraintes RMN.

### ■ Introduction des contraintes RMN

Les constantes de couplage  ${}^3J_{H_N H_\alpha}$  sont mesurées sur les spectres 1D  ${}^1H$  et intégrées manuellement dans un fichier de contraintes. CNS établit la conversion en contrainte d'angle dièdre  $\phi$  selon la relation de Karplus avec les coefficients de Pardi [163].

Les contraintes de distances inter-protons sont obtenues par intégration manuelle (TopSpin) des volumes des taches de corrélation sur les spectres NOESY selon la relation  $d = d_{ref} * (I_{ref}/I)^{1/6}$ . Dans le cas de PP13, les protons aromatiques de la tyrosine (résidu 4) ont servi de référence pour la calibration des volumes ( $d = 2.46 \text{ \AA}$ ). Etant donné le nombre de taches intégrées sur le spectre NOESY (plus de 200), la liste des contraintes de distances dans le format requis par CNS a été générée automatiquement par un script (awk). Dans le cas où les signaux de deux protons différents se superposent (déplacements chimiques identiques ou très proches), l'attribution d'une contrainte de distance devient ambiguë. Cette ambiguïté se traduit dans l'écriture de la contrainte par un « or » à l'intérieur d'une ligne « assign ».

Le résumé des contraintes structurales issues des données RMN est présenté dans le Tableau 10.

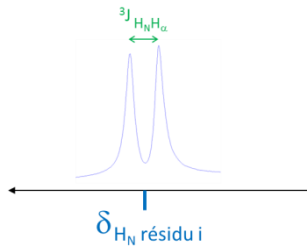
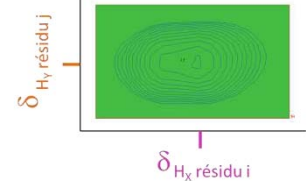
Expérience RMN	Mesure sur le spectre	Contrainte obtenue	Rédaction format CNS
1D ${}^1H$		${}^3J_{H_N H_\alpha}$ (en Hz) Relation de Karplus ${}^3J_{H_N H_\alpha} = 6,4 \cos^2\theta + 1,4 \cos\theta + 1,9$ Angle $\phi$ (en °) entre les liaisons H-N-C $_{\alpha}$ -H $_{\alpha}$	assign (resid i-1 and name C) (resid i and name N) (resid i and name CA) (resid i and name C) “ ${}^3J$ ” “intervalle”
2D NOESY ${}^1H$ - ${}^1H$		Intégrale (valeur absolue) $I \propto \frac{1}{d^6}$ Couple connu I et d de référence Distance entre H $_x$ et H $_y$ (en Å)	assign (resid i and name HX) (resid j and name HY) “d” “bornes sup et inf”

Tableau 10 : Contraintes structurales pour la modélisation sous contraintes RMN. L'intervalle de confiance pour la valeur  ${}^3J$  est fixé à 0.5 Hz, les bornes supérieures et inférieures pour les distances sont affectées selon l'intensité forte, moyenne ou faible de l'effet NOE :  $1.8 \leq d < 3.0$  « 0.8 0.8 » ;  $1.8 \leq d < 3.5$  « 1.2 1.2 » ;  $1.8 \leq d < 6.0$  « 2.0 2.0 ».



### ▪ Modélisation des structures

---

CNS applique un protocole de recuit simulé basé sur la dynamique des angles de torsion [169] pour générer des structures d'énergie minimale à partir de la structure étendue et des contraintes RMN.

En reprenant les principes de la mécanique et de la dynamique moléculaires, la méthode du recuit simulé permet d'explorer l'espace conformationnel à haute température et de dégager les minima d'énergie locaux (structures stables). La définition du champ de forces et le calcul des termes énergétiques spécifiques sont détaillés dans l'Annexe 2.

Brièvement, le protocole de modélisation peut se diviser en quatre étapes :

- Obtention de la structure étendue minimisée
- Introduction des contraintes RMN d'angles dièdres et de distance inter-protons sous forme de termes énergétiques et chauffage du système jusqu'à 50000 K.
- Refroidissement du système jusqu'à 300 K par paliers de 250 K.
- Minimisation des structures calculées

Une première série de 100 structures est générée à partir du fichier de contraintes de distances RMN. L'analyse des bilans énergétiques des structures, du bilan des erreurs de contact entre les atomes (distance < 1.8 Å) et du bilan des violations de distances (différence contrainte/modèle > 0.2 Å) permet d'estimer la validité des structures construites. Selon ces informations, le fichier de contraintes de distances est modifié manuellement afin de lever toute attribution ambiguë et réintégré dans le protocole de modélisation jusqu'à l'obtention de séries de structures contenant un minimum de violations de distance et présentant une énergie minimale.

A partir du fichier de contraintes de distance affiné, une série finale de 50 structures est générée en intégrant les contraintes d'angles dièdres.

### ▪ Validation des structures

---

Les 50 structures obtenues sont en premier lieu analysées par le logiciel Whymmod (développé par Julien Grosjean et Pedro Lameiras, téléchargeable à l'adresse <http://em2.crihan.fr>). Le module de clustering par RMSD a été utilisé afin de définir des familles de structures (amas de conformations). Seul le cluster contenant le plus grand nombre de structures a été retenu. Les structures de ce cluster ont finalement été étudiées par l'intermédiaire des logiciels de visualisation PyMOL [170] et MOLMOL [171]. Le logiciel RamaDA (développé par Matthieu Tanty et Marc-André Delsuc, accessible à l'adresse <http://ramada.u-strasbg.fr/>) a été utilisé pour valider l'ensemble des modèles générés.

## 3 Résultats et discussion

### 3.1 Détermination et comportement dynamique de la structure de VVTL1

La structure tridimensionnelle de la protéine VVTL1 a été obtenue par modélisation par homologie, basée sur la structure de cinq protéines TL de végétaux déjà résolues. Les pourcentages de chaque type de structures secondaires obtenues dans le modèle ont été comparés à ceux obtenus par dichroïsme circulaire sur la VVTL1 purifiée à partir de jus de raisin (cf Partie 1). Les deux répartitions se sont révélées très cohérentes. Cette corrélation modèle théorique/données expérimentales suggère à la fois la validité du modèle et celle du protocole de purification, qui conserve les éléments de structures secondaires de la protéine. La stabilité de VVTL1 thermique et au cours du temps a été déterminée respectivement par dichroïsme circulaire et dynamique moléculaire dans une boîte d'eau.

La flexibilité conformationnelle du modèle construit de VVTL1 a été étudiée par l'analyse de ses modes normaux. Trois domaines de la structure sont impliqués dans la majorité des mouvements de la protéine : mécanisme de mâchoire des domaines I et II autour de la cavité acide et flexibilité du domaine III vers l'extérieur. L'observation de ces mouvements laisse à penser qu'ils pourraient être conjointement liés à l'activité biologique de VVTL1, et des protéines TL en général. Le rôle majeur de la cavité acide, hautement conservée dans cette famille, avait précédemment été suggéré [81] mais aucune étude n'a jusqu'à présent été menée sur le mystérieux domaine III. La recherche de ce motif structural au sein de toute la Protein Data Bank révèle qu'il est absolument spécifique aux protéines TL, représentant peut-être un élément clé pour l'élucidation de leur mode d'action biologique.

Ce projet a donné lieu à la mise en place d'une collaboration avec le laboratoire SiRMA de l'Université de Reims Champagne-Ardenne pour toute la partie bioinformatique structurale et avec l'IGBMC de l'Université de Strasbourg pour l'application du logiciel RamaDA dans le cadre de la validation du modèle construit par homologie et de la recherche de motifs structuraux.

L'article suivant, « Insights into structure and function of a pathogenesis-related thaumatin-like protein from grape », présente les résultats de cette partie. Il est en cours de soumission au journal *Proteins : Structure, Function and Bioinformatics*.

# Insights into structure and function of a pathogenesis-related thaumatin-like protein from grape

D. Le Bourse<sup>1\*</sup>, M. Tanty<sup>2,3</sup>, N. Belloy<sup>4</sup>, A. Conreux<sup>1</sup>, S. Villaume<sup>1</sup>, P. Lameiras<sup>5</sup>, J-M. Nuzillard<sup>5</sup>, M-A. Delsuc<sup>2,3</sup>, P. Jeandet<sup>1</sup> and M. Dauchez<sup>4</sup>

<sup>1</sup>Laboratoire d'Oenologie et Chimie Appliquée, Université de Reims Champagne-Ardenne, URVVC-SE UPRES EA 2069, BP 1039, 51687 Reims Cedex 2, France

<sup>2</sup>IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), UMR 7104, 1 rue Laurent Fries, BP 10142, 67404 Illkirch-Graffenstaden Cedex, France

<sup>3</sup>NMRtec, Bioparc B, boulevard Sébastien Brandt, 67400 Illkirch-Graffenstaden, France

<sup>4</sup>Laboratoire SiRMa, UMR 6237 MEDyC, Université de Reims Champagne-Ardenne, BP 1039, 51687 Reims Cedex 2, France

<sup>5</sup>ICMR, Université de Reims Champagne-Ardenne, BP 1039, 51687 Reims Cedex 2, France

\*doriane.le-bourse@univ-reims.fr

## Abstract

A three-dimensional model of thaumatin-like (TL) protein from grape was constructed by homology modeling using five highly similar TL templates from other plant species. The overall model quality was assessed by PROCHECK and RamaDA programs. Confrontation of secondary structures quantification from both the homology model and the purified protein from grape juice showed very consistent results. The TL structure was found to be stable over time, as demonstrated by a 150 ns unconstrained molecular dynamics simulation in water, and stable over a large temperature range, as monitored by circular dichroism spectroscopy. Normal modes analysis (NMA) was performed to give an insight into conformational dynamics of the protein. It revealed that the acidic cleft shared among the whole TL superfamily exhibits a jaw-like mechanism which could be linked to the alleged glucanase activity of these proteins. NMA also pointed out the high flexibility of the  $\beta$ -hairpin from domain III. This structural pattern was found to be absolutely specific to TL proteins and could thus be a key to elucidate their exact biological activity.

## Introduction

Thaumatin-like (TL) proteins are made of around 200 amino acids and share sequence similarities with thaumatin, a sweet-tasting protein [1]. They are classified as pathogenesis-related (PR) family 5 due to their inducible expression by environmental stresses such as fungal attacks [2]. Despite its diversification among plant species, the superfamily of TL proteins shares well-defined structural domains, especially an acidic cleft assumed to be relevant to protein specific receptor binding for an antifungal activity [3]. This activity is alleged to be linked to the ability of TL proteins to degrade the cell walls of pathogenic fungi which are commonly composed of  $\beta$ -1,3-glucans [4-7]. While the antifungal activity of TL proteins was clearly evidenced [8], other modes of action such as the functioning of TL proteins as enzyme inhibitors may confer a particular biological function to these proteins [3]. Determination of their exact role in plants thus remains a largely unsolved issue. TL proteins are some of the most abundant proteins in grapes (*Vitis vinifera*) [9-11]. As for other TL in plants, their antifungal activity was previously observed but not fully elucidated [12, 13]. Understanding their full biological activity would allow improvements in protecting vineyards against devastating diseases, such as gray mold or powdery mildew infections. TL proteins also play a leading role in the wine industry concern. They can persist in wines due to their high stability over environmental stresses [14] and can induce the haze formation phenomenon in white wines, attributed to the protein aggregation during bottle storage [15-17]. This protein family is also alleged to have an influence on foaming properties of sparkling wines and Champagne [18] but no evidence of its involvement in gas/liquid interfaces has been yet demonstrated.

Despite recent advances in protein purification, the access to experimental data for structure determination from X-ray crystallography remains challenging [19]. NMR analysis of the full protein is even more inaccessible considering its size and the isotope labeling required involving expression in heterologous systems. The aim of this work is thus to provide a first study of the three-dimensional structure of the main TL protein from grape (VVTL1, gi|2213852 [20]) using homology modeling and to give an insight into its structural behavior. Purified VVTL1 from Chardonnay grape juice was used to perform circular dichroism (CD) spectroscopy in order to confront the quantification of secondary structures with the predicted model and to monitor the protein stability over temperature. The structure stability was also studied by a 150 ns molecular dynamics simulation. The protein flexibility, studied by conformational dynamics, could get a first clue to elucidate the biological activity of VVTL1 in the field of plant-pathogen interactions. Moreover, a conformational study can be led on VVTL1 to increase the information on the protein in order to understand its behavior. Many hypotheses may be expressed thanks to this kind of study.

## Materials and Methods

### Protein purification

VVTL1 was obtained from Chardonnay grape juice harvested in 2009 according to a previous study [14]. Briefly, 20 L of grape juice were first fractionated using a FPLC cation exchange column. Collected fractions were then subjected to a second FPLC purifying step (hydrophobic interaction). Extra-cleansing was achieved by UHPLC. Protein assessment and purity level determination were

realized by mass spectrometry and NMR spectroscopy analyses, respectively. VVTL1 fraction was finally freeze-dried and stored at room temperature.

## Circular dichroism spectroscopy

CD analyses were performed on a Chirascan (Applied Photophysics), using cells with a 0.1 cm path length. The purified protein was dissolved in ultrapure water (8.8  $\mu$ M) and CD spectra were recorded at different temperatures (4°C - 80°C range) from 255 nm to 185 nm. Quantification of secondary structures was performed by Dichroweb.

## Homology modeling

The primary sequence of VVTL1 was obtained from NCBI protein database (accession number AAB61590). The 24 first aminoacids were assessed to be a signal peptide [20] so they were not considered for protein structure building.

A set of templates (1Z3Q [21], 2I0W [7], 1PCV [6], 1AUN [5] and 1DU5 [4]) showing strong sequence similarity with VVTL1 was found using BLASTP program at NCBI [22]. High-resolution X-ray crystallography structure of each template was obtained from the Protein Data Bank. Multiple sequence alignment of templates and VVTL1 (target) was performed using ClustalW2 [23].

Program MODELLER 9v7 [24] was used to generate 50 models of VVTL1 based on the 5 chosen templates. The most accurate model was selected based on the MODELLER objective function for further analyses.

## Model evaluation

The overall stereochemical quality of the final model of VVTL1 was assessed by the program PROCHECK [25]. Moreover, the RamaDA web-server (<http://ramada.u-strasbg.fr/>) was used to highlight putative local misfoldings on the protein.

## Molecular dynamics

VVTL1 model was minimized and solvated in a water periodic box using the NAMD module from VMD [26]. Four Na<sup>+</sup> ions were added at random positions to satisfy the electroneutrality condition. An unconstrained 150 ns simulation was performed using the computational resources of the Clovis cluster (University of Reims). The simulation convergence was analyzed in terms of root mean square deviation (RMSD) from the initial model structure.

## Conformational dynamics

The conformational flexibility of VVTL1 was studied by normal mode analysis (NMA) using the elastic network model as embedded in the eINémo webserver [27] where mode 7 was performed on a -40/40 % perturbation range. WEBnm@ server [28] was also used to study modes 7 to 12 with all parameters set to their default values.

## Structural patterns

The conformational assignment of VVTL1 was performed by the RamaDA web-server. The assignment between C50 and C71 was first used as a pattern to find other files adopting this conformation over the entire PDB.

Then, thaumatin and thaumatin-like protein files were searched for into the PDB web-server. A conformational pattern which boundaries are the same as previously was designed to detect each of these files. Files found by this pattern and not identified as a thaumatin-like protein were studied. Patterns given in the following sections will use the nomenclature of regular expressions. Conformational domains will be represented by a letter : H for R-helices, B for  $\beta$ , P for PPII, L for L-helices, G for  $\gamma$ , Z for  $\zeta$  and Q for PPII<sub>R</sub>. Conformational states *extended* and *random-coil*, as for them, will be represented by e and R, respectively.

## Results and discussion

### Structure determination

#### **Homology modeling**

VVTL1 is a grape protein belonging to the thaumatin-like superfamily. It consists of 198 amino acids (including 16 cysteine residues), with a predicted molecular mass of 21.287 kDa and an isoelectric point of 4.55.

As no 3D experimental data could be obtained from X-ray crystallography or NMR spectroscopy, structural prediction based on comparative modeling was a method of choice to determine the 3D structure of VVTL1. This approach allows to generate models using structural information from previously resolved proteins (templates) showing a substantial sequence similarity with the target protein (VVTL1). Five thaumatin-like proteins deposited in the Protein Data Bank (1Z3Q, 2I0W, 1PCV, 1AUN and 1DU5) were chosen as templates. All the five templates revealed a great sequence identity with VVTL1 (Fig. 1), confirming the homology-based approach relevance to build a great quality protein model.

The three-dimensional structure of VVTL1 (Fig. 2) exhibits the characteristic thaumatin-like folding. The overall structure of the protein and all the eight disulfide bonds (marked in orange) are highly conserved in thaumatin [29] and thaumatin-like proteins [3]. The protein structure can be divided in three domains. Domain I is a lectin-like  $\beta$ -barrel that forms the compact central core of the molecule. It is made up of 11  $\beta$ -strands organized in a  $\beta$ -sandwich structure stabilized by two disulfide bonds (residues 9-197 and 112-186). Another disulfide bond can be found between domain I and domain II (residues 117-169, respectively) and a cleft can be clearly observed between these two domains. The domain II consists of an  $\alpha$ -helix (residues 160-168) and three  $\alpha$ -helical segments (residues 121-124, 138-142, 145-148). Three disulfide bonds stabilize the domain II (residues 125-135, 139-148, 149-157). On the other side of domain I, domain III consists of a 13 residues  $\beta$ -hairpin (residues 50-62) stabilized by a disulfide bond (residues 50-60). The long loop following the  $\beta$ -hairpin is stabilized by a disulfide bond (residues 65-71).

VVTL1	1 -	ATFDILNKCTYTVWAAASP	-	GGGRRLDSGQSWT	ITVNP	GTTNAP	IWGR	TSCTFD	ANGR	GRKCE	TGDC	NGLLE	EC	71
1Z3Q	1 -	ATFEIVNRCSYTVWAAAVP	-	GGGRQLNQGSWT	INVN	AGTTGGR	IWGR	TGCS	FDGS	GRGR	QQTG	DCGG	VLS	SC
2I0W	1 -	ATIEVRNNOPTVWAASTP	I	GGGRRLNRGQTWV	INAP	RGT	KMAR	IWGR	TGCF	NA	AGR	TQ	TGDC	CGV
1PCV	1 -	ATIEVRNNOPTVWAASTP	I	GGGRRLDRGQTWV	INAP	RGT	KMAR	VWGR	TNCF	NA	AGR	TQ	TGDC	CGV
1AUN	1	SGVFEVHNNOPTVWAAATP	V	GGGRRLER	GQSW	FWAPP	GT	KMAR	IWGR	TNCF	DG	AGR	WC	TGDC
1DU5	1 -	AVFTVNVNOCFFTWAAASV	V	GGGRQLNRGESWR	ITAP	AGT	TAA	IWART	GCK	FDAS	GRGS	ORT	TGDC	CGV
VVTL1	72	QYGSPPNTLAE	FAL	NQPNLDYID	ISLV	GFNIP	MDFS	-----	G	C	RG	IQ	CSVD	ING
1Z3Q	72	TAYGNPPNTLAE	FAL	NQFNLDFFD	ISLV	GFNV	PMDFS	PT	S	T	S	---	GGC	RG
2I0W	73	TGWGKPPNTLAE	YAL	DQFSNLD	FDW	ISLV	GFNIP	M	T	F	A	P	T	K
1PCV	73	TGWGKPPNTLAE	YAL	DQFSGLD	FDW	ISLV	GFNIP	M	T	F	A	P	T	K
1AUN	74	KGWGKPPNTLAE	YAL	NQFSNLD	FDW	ISV	IGFNIP	M	S	F	G	P	T	K
1DU5	73	TGYGRAPNTLAE	YAL	KQFNLD	FFD	ISL	IGFN	V	P	M	S	E	L	P
VVTL1	137	NPCTVFKTNEY	CCTD	GP	-	SCGPT	TSK	FFKDR	CP	DAYS	YPQ	DKT	SL	FT
1Z3Q	141	NPCTVFKTDQY	CCNSG	---	AC	SPTD	YS	QFFKRN	CP	DAYS	YPK	DD	Q	T
2I0W	145	NPCTTFGGQQY	CCTQG	---	PC	GPT	EL	SK	FFKRC	CP	DAYS	YPQ	DD	P
1PCV	145	NPCTTFGGQQY	CCTQG	---	PC	GPT	FF	SK	FFKRC	CP	DAYS	YPQ	DD	P
1AUN	146	NPCTTFGGQQY	CCTQG	---	PC	GPT	EL	SR	WFKRC	CP	DAYS	YPQ	DD	P
1DU5	144	NACPVEKKDEY	CCV	GSA	AND	CH	P	T	N	S	R	Y	F	K
VVTL1	198	NPCTVFKTNEY	CCTD	GP	-	SCGPT	TSK	FFKDR	CP	DAYS	YPQ	DKT	SL	FT
1Z3Q	200	NPCTVFKTDQY	CCNSG	---	AC	SPTD	YS	QFFKRN	CP	DAYS	YPK	DD	Q	T
2I0W	207	NPCTTFGGQQY	CCTQG	---	PC	GPT	EL	SK	FFKRC	CP	DAYS	YPQ	DD	P
1PCV	205	NPCTTFGGQQY	CCTQG	---	PC	GPT	FF	SK	FFKRC	CP	DAYS	YPQ	DD	P
1AUN	208	NPCTTFGGQQY	CCTQG	---	PC	GPT	EL	SR	WFKRC	CP	DAYS	YPQ	DD	P
1DU5	206	NACPVEKKDEY	CCV	GSA	AND	CH	P	T	N	S	R	Y	F	K

Fig. 1: Sequence alignment (ClustalW2 manually optimized) of VVTL1 (target) and the templates 1Z3Q, 2I0W, 1PCV, 1AUN, 1DU5 used for homology modeling. The sequence identity between VVTL1 and each template is respectively 73.1 %, 63.8 %, 63.5 %, 62.6 %, 62 %. The red box (residues 58 to 73) shows the TL protein family signature (PS00316) in thaumatin, G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C. Green background (45R, 85E, 98D, 103D, 189D) indicates the five conserved aminoacids involved in the protein acidic cleft.

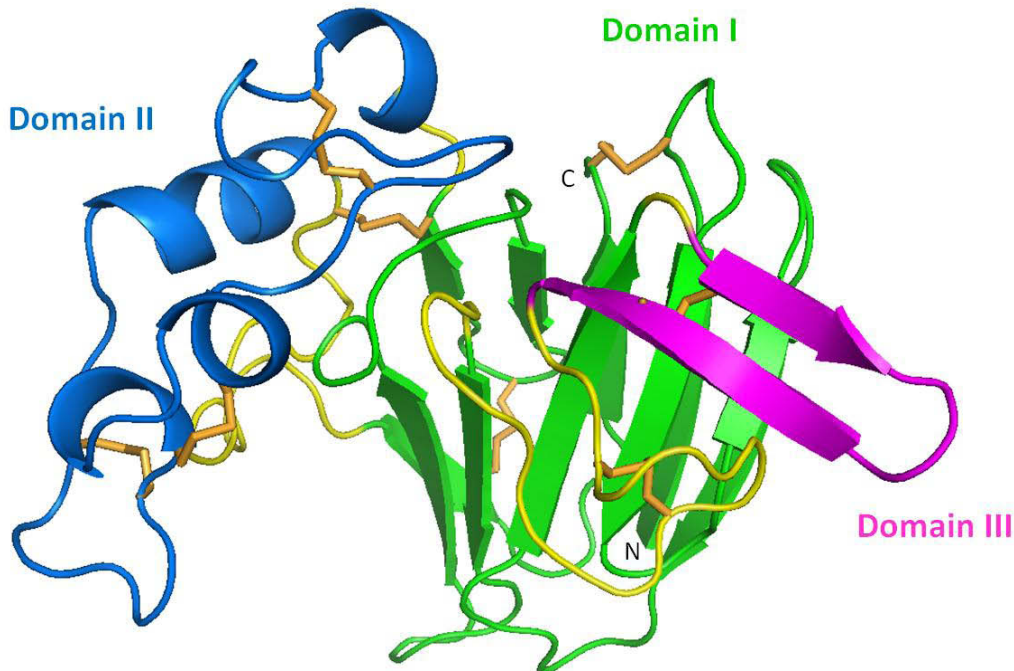


Fig. 2: Cartoon representation of homology-modeled VVTL1. The three structured domains I, II and III are respectively represented in green, blue and pink. Disulfide bonds are indicated in orange.

### Model evaluation

The VVTL1 model accuracy was assessed by determining the distribution of the main chain torsion angles phi and psi using a Ramachandran plot (Fig. 3). Detailed results (Tab. 1) show that all the residues are either in most favored or additional allowed regions and none in generously allowed or disallowed regions, suggesting the high quality of the model. The stereochemical quality of the model was also assessed by RamaDA revealing a great z-score of 1.30 (Fig. 4).

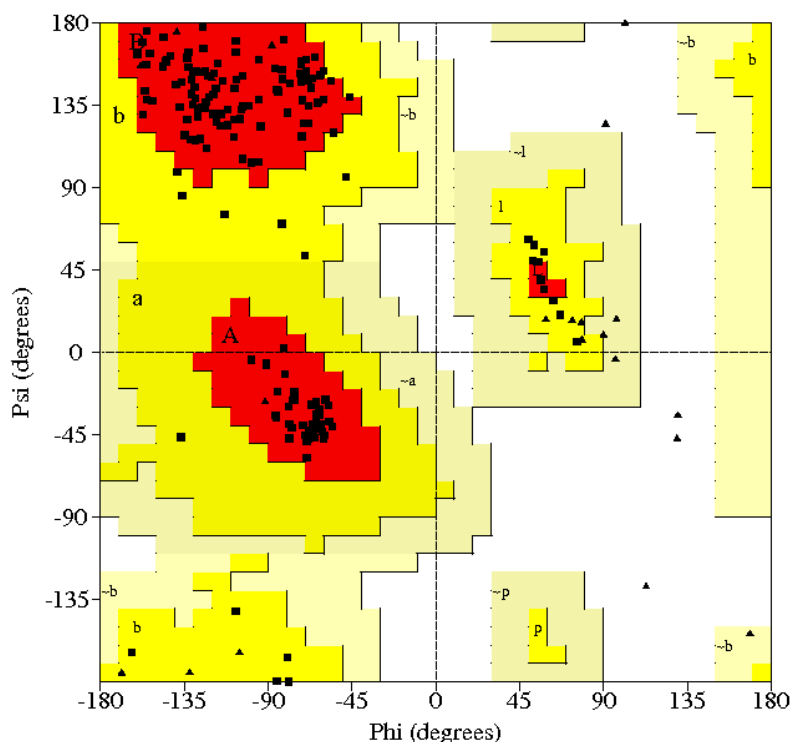


Fig. 3: Ramachandran plot of the homology-modeled structure of VVTL1 by PROCHECK. The different colored areas indicate “disallowed” (white), “generously allowed” (light yellow), “additional allowed” (yellow) and “most favored” (red) regions.

Residues in most favored regions	142 (88.8%)
Residues in additional allowed regions	18 (11.2%)
Residues in generously allowed regions	0 (0.0%)
Residues in disallowed regions	0 (0.0%)
Number of non-glycine and non-proline residues	160
Number of end-residues (excl. Gly and Pro)	1
Number of glycine residues (shown as triangles)	22
Number of proline residues	15
Total number of residues	198

Tab. 1 : PROCHECK evaluation of the homology-modeled structure of VVTL1.



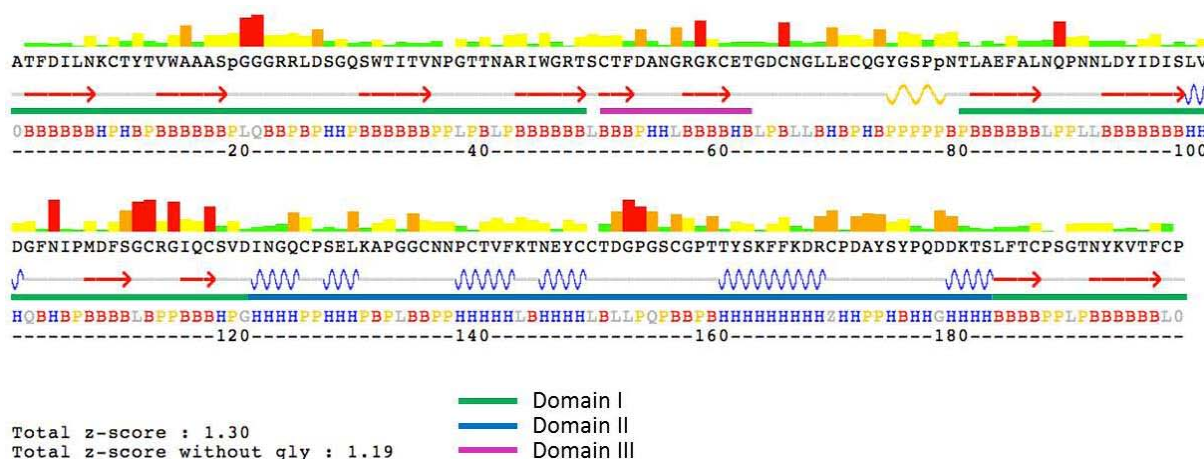


Fig. 4 : RamaDA analysis of the homology-modeled structure of VVTL1. First line shows the histogram of z-scores. Green is used for z-scores lower than 1, yellow for z-scores between 1 and 2, orange for z-scores between 2 and 3 and red for z-scores higher than 3. Second line corresponds to the protein sequence. Cis-prolines are detected thanks to their  $\omega$  dihedral angle and indicated by p. Third line indicates putative secondary structure elements (blue waves for helices, yellow waves for PPII helices and red arrow for  $\beta$  strands). Last line gives the conformational domains assignment (H for R-helices, B for  $\beta$ , P for PPII, L for L-helices, G for  $\gamma$ , Z for  $\zeta$  and Q for PPII<sub>R</sub>).

### Confrontation with experimental data

Purified VVTL1 obtained from Chardonnay grape juice after three liquid chromatography steps [14] was used to perform circular dichroism spectroscopy (Fig. 5). Spectrum deconvolution was achieved through the k2d method. Experimental distribution of secondary structures was thus assessed in water at room temperature and compared to results obtained through DSS assignment (PyMOL algorithm) applied on the solvated homology model (Tab. 2). Both approaches showed very consistent values for each type of secondary structure, self-confirming model quality and non-alteration of the purified protein structure.

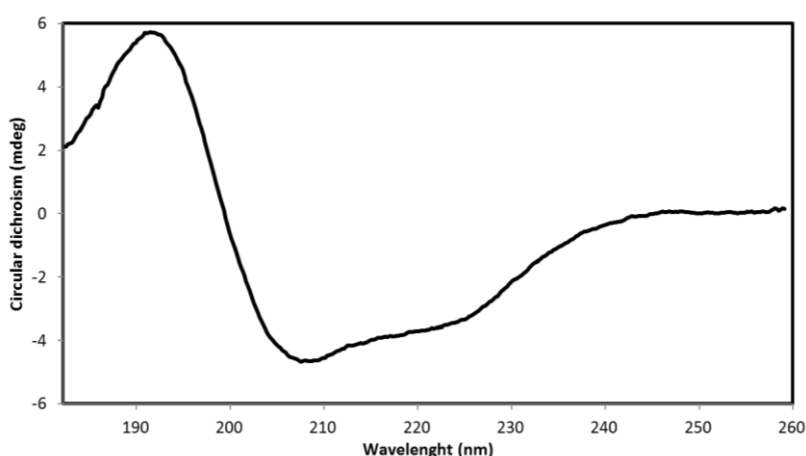


Fig. 5 : CD spectra of purified VVTL1 dissolved in water (8  $\mu$ M) recorded at 25°C.

---

	Homology modeling	Circular dichroism
$\alpha$ -helixes (%)	13	13
$\beta$ -sheets (%)	40	38
Others (%)	47	49

---

Tab. 2 : Quantification of secondary structures from the modeled structure and the purified VVTL1.

## Structure stability over time and temperature

TL proteins are known to be very stable and resistant to environmental stresses (temperature/pH variations and proteolysis) [3].

A molecular dynamics simulation was thus performed to evaluate the stability of VVTL1 structure at room temperature and at normal pressure. The overall stability of the solvated model was studied over 150 ns. Results show that the system remains in equilibrium during the entire simulation. While high motion of two external loops can be observed, the protein core appears to be very stable during the whole trajectory (Fig. 6).

Purified VVTL1 was also studied by circular dichroism spectroscopy over a large range of temperature to collect experimental information about its structural behavior. VVTL1 revealed a great stability over temperature variations considering that only a slight denaturation begins to occur around 60°C, confirming previously reported differential scanning calorimetry results [15]. No structural variation was observed over the 4°C-25°C temperature range. Given that VVTL1 could be involved in gas/liquid interfaces stabilization, the different foamabilities observed while serving a Champagne wine at different temperature [30] is likely not related to intrinsic structural protein changes.

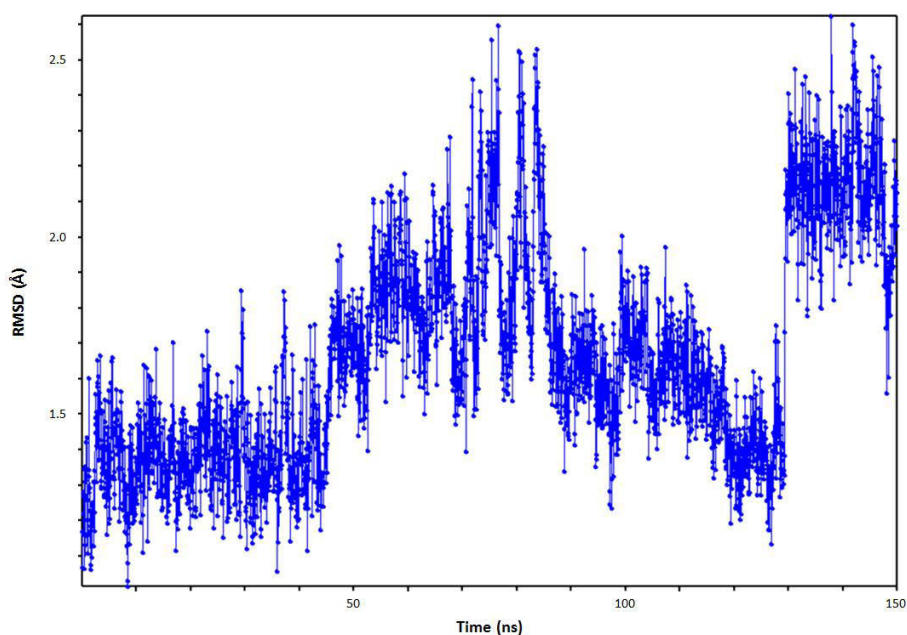


Fig. 6: Graphical representation of VVTL1 backbone RMSD observed with respect to time throughout the molecular dynamics simulation.

**Conformational flexibility and structural patterns: insight into function prediction ?**

To explore the protein flexibility, conformational dynamics was carried out through normal mode analysis. This method allows to express the motions of a protein in terms of collective motions of a group of atoms [31]. Each mode represents fluctuations of the protein around a minimum energy conformation.

Overall analysis of modes 7 to 12 (Fig. 7) revealed three very flexible areas in VVTL1 : from residues 145 to 159 ( $\alpha$ -helix and loop from the cleft part of domain II), from residues 86 to 93 ( $\beta$ -sheet and loop from the cleft part of domain I), from residues 50 to 62 ( $\beta$ -hairpin from domain III). The cleft between domains I and II and the domain III thus appear to be key elements to understand the protein function.

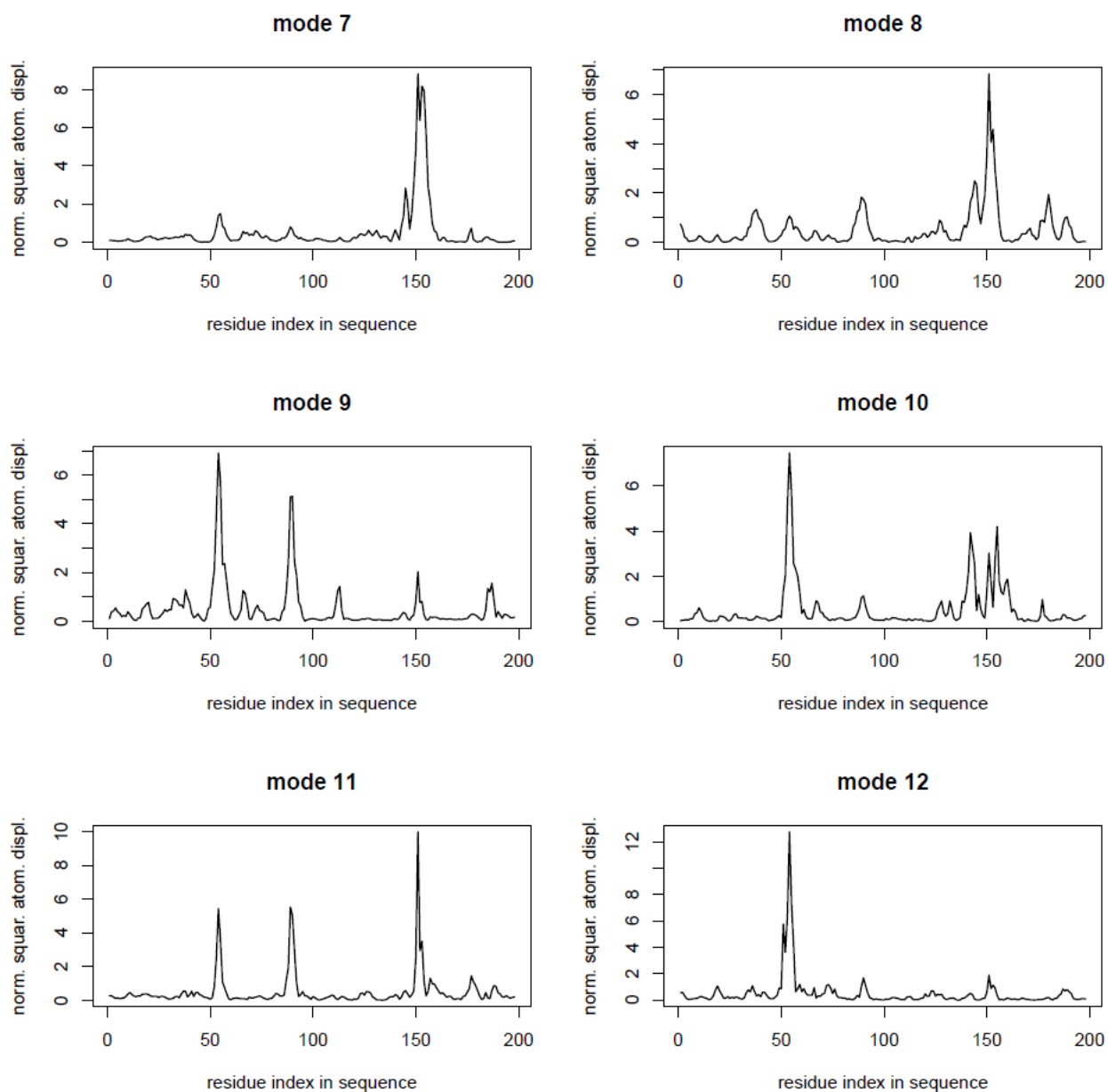


Fig. 7: Atomic displacements of each VVTL1  $C_{\alpha}$  atom for modes 7 to 12 from WEBnm@.

### ***The acidic cleft***

Electrostatic potential surface of VVTL1 indicates a high acidity in the cleft region (Fig. 8). Five amino acids (arginine, glutamic acid and three aspartic acid residues) are responsible for the cleft acidity and are highly conserved among TL proteins with known antifungal activity (Fig. 1).

A TL protein from tomato (2I0W) was reported to bind a linear glucan in the interdomain cleft but no dynamics studies were reported to understand the behavior of the docked complex [7]. Considering that intrinsic flexibility of a protein can facilitate the conformational changes required for its function, the mode 7 from NMA was of great interest. It reveals a jaw-like mechanism between domains I and II (Fig. 8), opening and closing the acidic cleft region. The potential glucanase activity of VVTL1 could thus be considered as a three steps mechanism: 1) open cleft 2) ligand binding / closed cleft 3) open cleft / degraded ligand released.

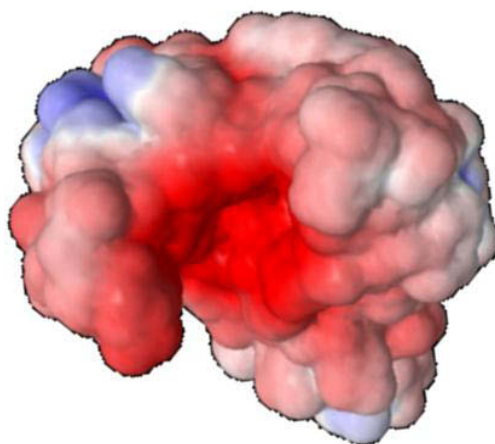


Fig. 8: Electrostatic potential surface of VVTL1 (using APBS), clearly indicating the highly acidic cleft between domains I (on the right) and II (on the left). The positively and negatively charged surface regions are shown in blue and red, respectively.

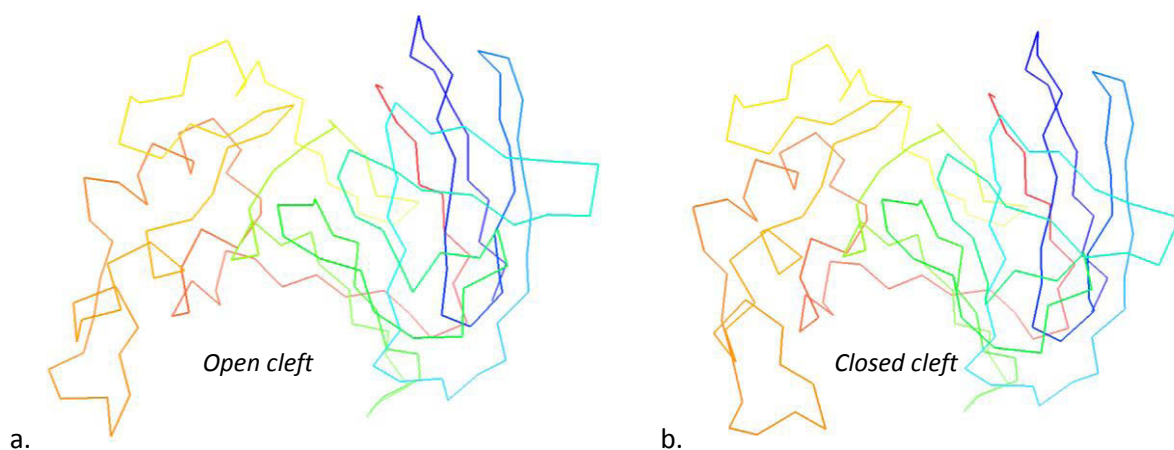


Fig. 8 : NMA of VVTL1 (mode 7) showing open (a.) and closed (b.) cleft region between domains I and II.

### ***The mysterious domain III***

Despite the understandable flexibility of domains I and II around the acidic cleft, the conformational changes observed for the  $\beta$ -hairpin of domain III allege of its potential role in protein function which yet remains unknown.

The conformational assignment of domain III given by RamaDA (BBBPHHLBBBBHBLPBLLBHBP), once used as a pattern searched for over the entire PDB, allows the detection of 9 files (1KWN, 1LYO, 1PP3, 1THW, 2G4Y, 2VI1, 2VI3, 2VU6 and 3AOK). All these files describe thaumatin-like protein structures with exactly the same domain III amino-acid sequence (except 3AOK which presents a glutamine instead of an arginine at the 76<sup>th</sup> position). Although none of this file is one of the models used for the homology modeling, they all are high-resolution files (below 2 Å), emphasizing on the good quality of the VVTL1 structure proposed here.

In the PDB, 45 files are identified as describing thaumatin or thaumatin-like proteins. Among them, only one can not be analyzed by RamaDA because it only provides  $\alpha$ -carbon positions. Thanks to all these files, it is possible to design a new conformational pattern which is able to detect them all. This pattern is  $eBB e.[HQ][LB]e[BQ]BBHBLP.L.\{0,1\}[LH]e[HG][Ge][PH]$ . 45 files can be found with this pattern meaning that only one of these files is not identified as thaumatin-like. This file is 1PCV, one of the model for homology modeling, describing an antifungal protein structure.

Thus, it clearly appears that domain III conformational assignment is very specific to thaumatin and TL proteins. As conformations are closely linked to biological activities, it can be suggested that domain III is important for the thaumatin function. Moreover, since the conformational assignment is only found in anti-fungal proteins, domain III should take part of the mechanism of this activity. To assess this statement, more biological studies must be conducted. However, a simple conformational study led with RamaDA is enough to express strong hypotheses on a protein activity.

## **Conclusion**

Homology modeling allowed to determine the structure of TL protein from grape. Considering the issue raised by TL proteins biological activity, NMA turned out to be a powerful tool to identify flexible - thus functional - protein areas before full docking studies.

Apart from biological activity occurring in grapes, VVTL1 can be involved in various phenomena in the oenology field. Revealing the structural behavior of this protein appears to be a prerequisite to understand gas/liquid or liquid/solid adsorption mechanisms and could open a way toward a better understanding of the role played by proteins in wines. In particular, this study could lead to investigate the surfactant properties of TL proteins and their unfolding on gas/liquid interfaces in a sparkling wine.

It also becomes possible to work on the anti-fungal activity of TL proteins thanks to the model obtained by homology modeling. A short conformational study of domain III suggests that it is very important for this activity. Moreover, this kind of conformation can not be found in any other type of protein in the PDB, showing that this folding must be associated to a very specific property.

Getting a structure of VVTL1 was important in order to study the behaviour of TL proteins. Now there is one, a lot of hypotheses on its mechanism can be expressed. All these hypotheses may lead to a better understanding of thaumatin and TL proteins.

## Acknowledgements

We thank the Champagne Ardenne Region, the General Council of Marne, the Ministry of Research, the Ville de Reims, the A.R.O.C.U. Association, and the Champagne Moët & Chandon for their support. We also thank Nicolas Borie from ICMR, URCA, for protein freeze-drying.

## References

1. van der Wel, H. and K. Loeve, *Isolation and Characterization of Thaumatin I and II, the Sweet-Tasting Proteins from Thaumatooccus daniellii Benth.* European Journal of Biochemistry, 1972. **31**(2): p. 221-225.
2. van Loon, L.C., M. Rep, and C.M.J. Pieterse, *Significance of Inducible Defense-related Proteins in Infected Plants.* Annual Review of Phytopathology, 2006. **44**(1): p. 135-162.
3. Liu, J.J., R. Sturrock, and A.K.M. Ekramoddoullah, *The superfamily of thaumatin-like proteins: Its origin, evolution, and expression towards biological function.* Plant Cell Reports. **29**(5): p. 419-436.
4. Batalia, M.A., et al., *The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like, PR-5 protein family.* Nat Struct Biol, 1996. **3**(1): p. 19-23.
5. Koiwa, H., et al., *Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins.* J Mol Biol, 1999. **286**(4): p. 1137-45.
6. Min, K., et al., *Crystal structure of osmotin, a plant antifungal protein.* Proteins: Structure, Function, and Bioinformatics, 2004. **54**(1): p. 170-173.
7. Ghosh, R. and C. Chakrabarti, *Crystal structure analysis of NP24-I: a thaumatin-like protein.* Planta, 2008. **228**(5): p. 883-90.
8. Velazhahan, R. and S. Muthukrishnan, *Transgenic Tobacco Plants Constitutively Overexpressing a Rice Thaumatin-like Protein (PR-5) Show Enhanced Resistance to Alternaria alternata.* Biologia Plantarum, 2003. **47**(3): p. 347-354.
9. Le Bourse, D., et al., *Review of preparative and analytical procedures for the study of proteins in grape juice and wine.* Analytica Chimica Acta. **667**(1-2): p. 33-42.
10. Robinson, S.P. and C. Davies, *Molecular biology of grape berry ripening.* 2000, Blackwell Publishing Ltd. p. 175-188.
11. Giribaldi, M. and M.G. Giuffrida, *Heard it through the grapevine: Proteomic perspective on grape and wine.* Journal of Proteomics. **73**(9): p. 1647-1655.
12. Salzman, R.A., et al., *Coordinate Accumulation of Antifungal Proteins and Hexoses Constitutes a Developmentally Controlled Defense Response during Fruit Ripening in Grape.* Plant Physiology, 1998. **117**(2): p. 465-472.
13. Monteiro, S., et al., *Osmotin and Thaumatin from Grape: A Putative General Defense Mechanism Against Pathogenic Fungi.* Phytopathology, 2003. **93**(12): p. 1505-1512.
14. Le Bourse, D., et al., *Quantification of chitinase and thaumatin-like proteins in grape juices and wines.* Analytical and Bioanalytical Chemistry, 2011. **401**(5): p. 1545-1553.
15. Falconer, R.J., et al., *Thermal stability of thaumatin-like protein, chitinase, and invertase isolated from sauvignon blanc and semillon juice and their role in haze formation in wine.* Journal of Agricultural and Food Chemistry, 2010. **58**(2): p. 975-980.
16. Marangon, M., et al., *Roles of grape thaumatin-like protein and chitinase in white wine haze formation.* Journal of Agricultural and Food Chemistry, 2011. **59**(2): p. 733-740.

17. Marangon, M., et al., *Effects of Ionic Strength and Sulfate upon Thermal Aggregation of Grape Chitinases and Thaumatin-like Proteins in a Model System*. Journal of Agricultural and Food Chemistry. **59**(6): p. 2652-2662.
18. Cilindre, C., et al., *Foaming properties of various Champagne wines depending on several parameters: Grape variety, aging, protein and CO<sub>2</sub> content*. Analytica Chimica Acta. **660**(1-2): p. 164-170.
19. Van Sluyter, S.C., et al., *Two-Step Purification of Pathogenesis-Related Proteins from Grape Juice and Crystallization of Thaumatin-like Proteins*. Journal of Agricultural and Food Chemistry, 2009. **57**(23): p. 11376-11382.
20. Tattersall, D.B., R. Van Heeswijck, and P.B. Hoj, *Identification and characterization of a fruit-specific, thaumatin-like protein that accumulates at very high levels in conjunction with the onset of sugar accumulation and berry softening in grapes*. Plant Physiology, 1997. **114**(3): p. 759-769.
21. Leone, P., et al., *Resolution of the structure of the allergenic and antifungal banana fruit thaumatin-like protein at 1.7-Å*. Biochimie, 2006. **88**(1): p. 45-52.
22. Altschul, S.F., et al., *Basic local alignment search tool*. Journal of Molecular Biology, 1990. **215**(3): p. 403-410.
23. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0*. Bioinformatics, 2007. **23**(21): p. 2947-2948.
24. Sali, A. and T.L. Blundell, *Comparative Protein Modelling by Satisfaction of Spatial Restraints*. Journal of Molecular Biology, 1993. **234**(3): p. 779-815.
25. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures*. Journal of Applied Crystallography, 1993. **26**(2): p. 283-291.
26. Humphrey, W., A. Dalke, and K. Schulten, *VMD: Visual molecular dynamics*. Journal of Molecular Graphics, 1996. **14**(1): p. 33-38.
27. Suhre, K. and Y.-H. Sanejouand, *EINémo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement*. Nucleic Acids Research, 2004. **32**(suppl 2): p. W610-W614.
28. Hollup, S., G. Salensminde, and N. Reuter, *WEBnm@: a web application for normal mode analyses of proteins*. BMC Bioinformatics, 2005. **6**(1): p. 52.
29. Ogata, C.M., et al., *Crystal structure of a sweet tasting protein thaumatin I, at 1.65 Å resolution*. Journal of Molecular Biology, 1992. **228**(3): p. 893-908.
30. Liger-Belair, G.r., et al., *Kinetics of CO<sub>2</sub> Fluxes Outgassing from Champagne Glasses in Tasting Conditions: The Role of Temperature*. Journal of Agricultural and Food Chemistry, 2009. **57**(5): p. 1997-2003.
31. Tirion, M.M., *Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis*. Physical Review Letters, 1996. **77**(9): p. 1905.

### 3.2 Détermination de la structure d'un peptide issu de VVTL1 par RMN

Après la détermination de la structure de la protéine VVTL1, plusieurs domaines peptidiques d'intérêt ont pu être observés. Dans le cadre d'une première approche, nous avons choisi d'étudier une hélice  $\alpha$ , située vers l'extérieur de la protéine donc pouvant à priori facilement interagir avec son environnement et incluse dans le motif allergénique identifié dans la superfamille des TL [80, 81].

Le peptide PP13, Pro-Thr-Thr-Tyr-Ser-Lys-Phe-Phe-Lys-Asp-Arg-Cys-Pro, a ainsi été sélectionné sur la séquence de VVTL1 entre les résidus 158 et 170, caractérisant une hélice  $\alpha$  (en bleu sur la Figure 30). Les deux acides aminés en random coil avant et après l'hélice ont été inclus afin de relâcher la structure générale.

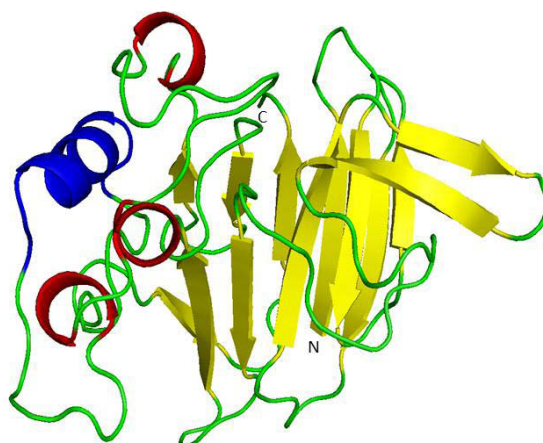


Figure 30 : Peptide PP13 (en bleu) sélectionné sur la structure modélisée de la protéine VVTL1.

Ce peptide a été synthétisé chimiquement sans marquage isotopique et analysé par RMN. A partir des contraintes de distances inter-protons et d'angles dièdres, sa structure a pu être modélisée et comparée avec une structure prédite par une modélisation *de novo*.

Ce projet a donné lieu à la poursuite des collaborations avec le laboratoire ICMR de l'Université de Reims Champagne-Ardenne pour la réalisation des expériences RMN et avec l'IGBMC de l'Université de Strasbourg pour l'application du logiciel RamaDA.

#### 3.2.1 Choix du solvant

Les peptides courts ayant une faible capacité à se structurer spontanément, une première approche expérimentale a été menée par dichroïsme circulaire afin de déterminer le solvant optimal pour l'analyse structurale par RMN du peptide PP13. Les solvants organiques TFE et MeOH ayant une capacité connue de structuration des peptides, les solvants et mélanges suivants ont été testés :

- H<sub>2</sub>O
- H<sub>2</sub>O/TFE 60:40 (v/v)
- H<sub>2</sub>O/MeOH 50:50 (v/v)
- MeOH



Les spectres CD obtenus indiquent l'augmentation de la structuration, typiquement sous forme d'hélice, avec l'augmentation du pourcentage de solvant structurant (Figure 31). L'analyse de la déconvolution des spectres indique que le peptide PP13 présente un très faible taux d'hélice dans l'eau (6 %) mais se structure jusqu'à 41 % d'hélice dans le méthanol pur. Ce dernier solvant a donc été retenu pour la réalisation des spectres RMN.

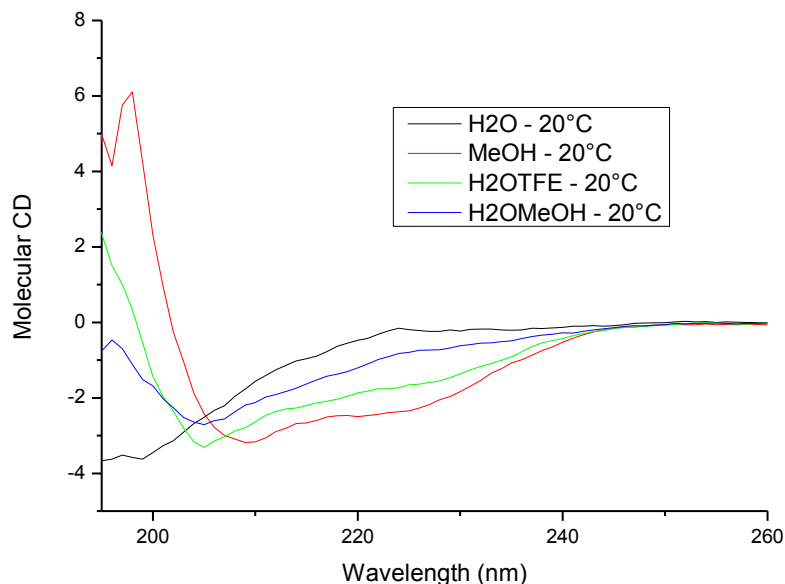


Figure 31 : Spectres CD du peptide PP13 dans différents solvants.

### 3.2.2 Attribution des déplacements chimiques

Les différentes expériences RMN 2D (COSY  $^1\text{H}$ - $^1\text{H}$ , TOCSY  $^1\text{H}$ - $^1\text{H}$ , NOESY  $^1\text{H}$ - $^1\text{H}$ , HSQC  $^1\text{H}$ - $^{13}\text{C}$  et  $^1\text{H}$ - $^{15}\text{N}$ , HMBC  $^1\text{H}$ - $^{13}\text{C}$ ) réalisées sur le peptide PP13 ont permis d'attribuer l'ensemble des déplacements chimiques  $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}$  (Tableau 11).

AA	H <sub>N</sub>	H <sub>α</sub>	H <sub>β</sub>	H <sub>γ</sub>	H <sub>δ</sub>	H <sub>ε</sub>	H <sub>ζ</sub>
1P		4,44	2,49 / 2,07	2,03	3,40/ 3,35		
2T	8,75	4,64	4,48	1,32			
3T	8,60	4,05	4,12	1,19			
4Y	7,99	4,42	3,09 / 3,00		7,07	6,70	
5S	8,19	4,26	4,05 / 3,95				
6K	8,22	4,04	1,77	1,33	1,59	2,86	
7F	8,00	4,33	3,10/2,97		7,13	7,22	7,20
8F	8,13	4,28	3,21/3,02		7,26	7,30	7,25
9K	8,17	4,14	1,90	1,53	1,71	2,94 / 2,86	
10D	8,17	4,57	2,93 / 2,86				
11R	7,84	4,28	1,89 / 1,72	1,57	3,05		
12C	8,10	4,70	2,86/2,62				
13P		4,44	2,28/2,00	2,02	3,74		

AA	N	C <sub>α</sub>	C <sub>β</sub>	C <sub>γ</sub>	C <sub>δ</sub>	C <sub>ε</sub>	C <sub>ζ</sub>
1P		59,5	29,9	23,6	46,1		
2T	110,1	58,2	68,3	18,5			
3T	113,8	61,3	66,1	19,3			
4Y	118,8	57,3	35,8		129,7	114,9	
-5S	114,7	58,2	60,6				
6K	121,3	56,0	29,6	22,2	26,9	34,9	
7F	116,9	57,2	36,3		128,7	128,0	126,4
8F	117,2	57,4	36,6		128,9	128,2	126,5
9K	115,3	54,6	29,7	22,2	26,5	34,9	
10D	118,0	51,1	34,9				
11R	116,8	53,0	28,0	24,5	40,4		
12C	118,5	53,9	24,8				
13P		59,4	28,8	24,4	47,1		

Tableau 11 : Déplacements chimiques <sup>1</sup>H, <sup>13</sup>C et <sup>15</sup>N du peptide PP13.

### 3.2.3 Indices de déplacements chimiques

Les indices de déplacements chimiques H<sub>α</sub> et C<sub>α</sub> ont été calculés en comparant les valeurs des déplacements chimiques expérimentaux et les valeurs théoriques en random coil [150, 157-159]. Ces indices indiquent potentiellement une conformation prédominante en hélice entre les acides aminés 3T et 8F (Figure 32). Les acides aminés chargés 9K, 10D et 11R ne sont pas détectés comme étant impliqués dans cette structuration.

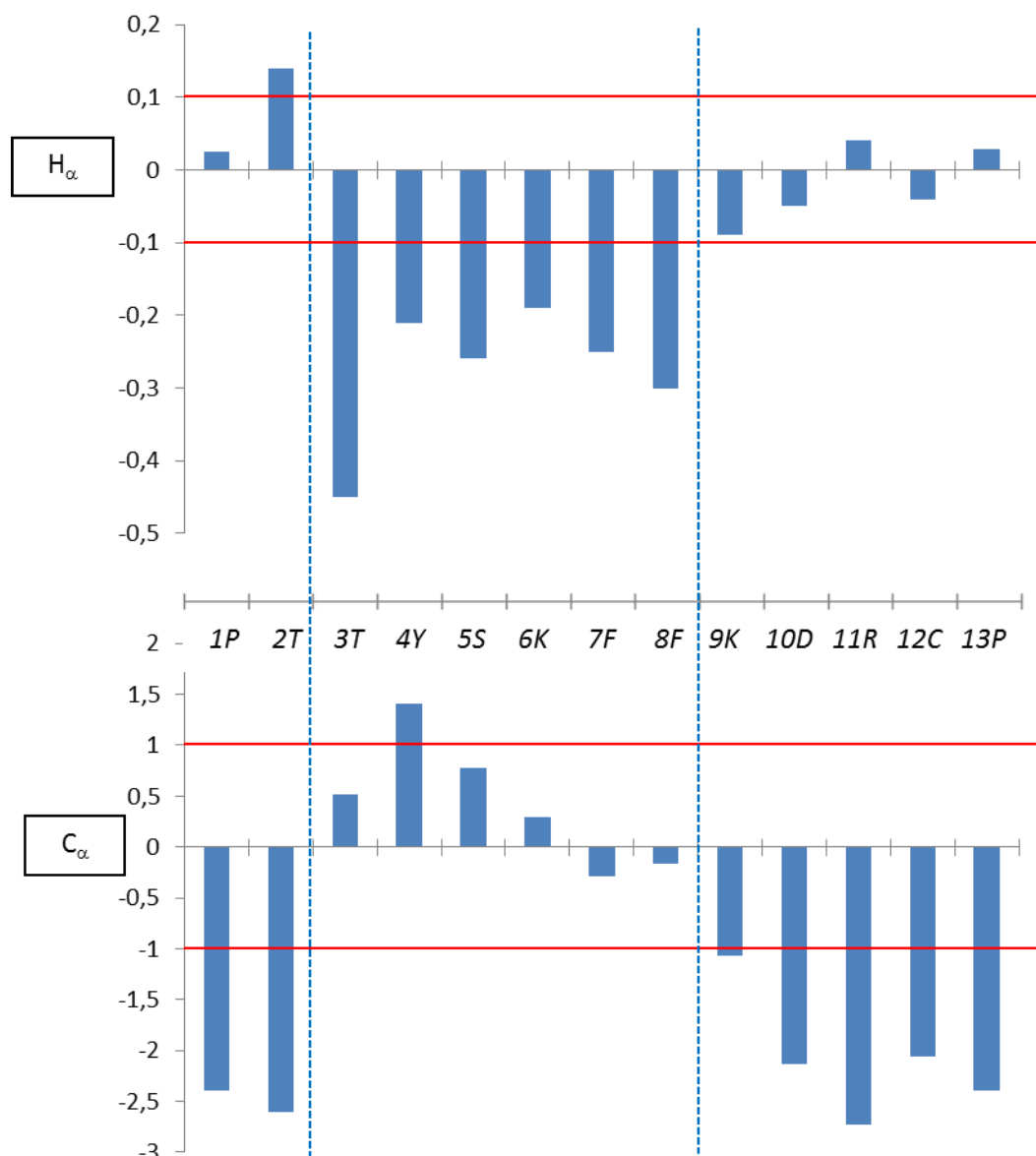


Figure 32 : Indices de déplacements chimiques des  $H_{\alpha}$  et  $C_{\alpha}$  du peptide PP13.

### 3.2.4 Détermination des contraintes structurales

Les constantes de couplage  ${}^3J_{HN-H\alpha}$  ont été mesurées sur le spectre 1D  ${}^1H$  puis intégrées dans un fichier de contraintes CNS, qui a établi la conversion en contrainte d'angle dièdre  $\phi$  selon la relation de Karplus (Annexe 3.2). La proline (résidus 1 et 13) ne possédant pas de  $H_N$  aucune constante de couplage ne peut être déterminée. Les résidus 3 à 10 présentent une constante de couplage de l'ordre de 6 Hz ou inférieure (en vert), confirmant la probabilité de la structuration du peptide en hélice déjà précédemment observée (Tableau 12).

AA	$^3J$ (Hz)
2T	8,22
3T	5,11
4Y	5,17
5S	5,28
6K	4,93
7F	5,69
8F	6,25
9K	6,38
10D	6,38
11R	8,28
12C	8,70

Tableau 12 : Constantes de couplage  $^3J_{\text{HN-H}\alpha}$  du peptide PP13.

En parallèle, une liste de 129 contraintes de distances inter-protons a été générée après intégration des taches NOE avec les protons aromatiques de 4Y comme volume de référence (Annexe 3.1). Le diagramme des connectivités NOE expérimentales entre les résidus de  $i$  à  $i+4$  obtenu indique également des motifs caractéristiques d'une hélice (Figure 33).

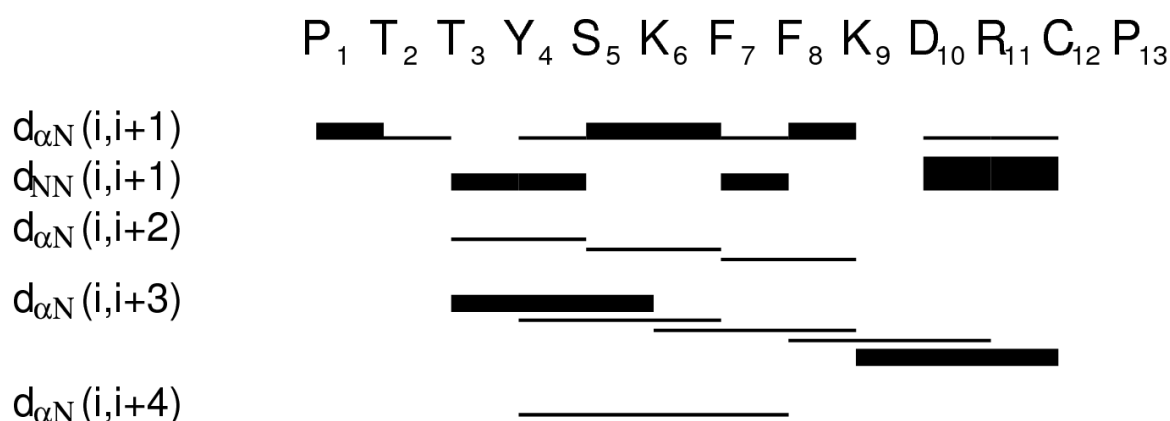


Figure 33 : Diagramme des connectivités NOE observées dans le peptide PP13. L'intensité des barres reflète l'intensité de l'effet NOE (fort, moyen, faible).

### 3.2.5 Modélisation moléculaire sous contraintes RMN

Les contraintes d'angles dièdres et de distances inter-protons ont été intégrées dans le logiciel CNS afin de générer une série de 50 structures ne présentant aucune violation de structure (distance  $< 0.2 \text{ \AA}$  et angle  $< 10^\circ$ ). L'utilisation du logiciel Whymmod a permis de réaliser un clustering par RMSD de ces modèles. Un cluster principal regroupant 22 structures a été obtenu (RMSD backbone  $< 1 \text{ \AA}$ ), définissant l'ensemble des structures retenues. La population des angles ( $\phi$ ,  $\psi$ ) est indiquée dans la Figure 34 (logiciel MOLMOL). Le logiciel RamaDA, appliqué sur les modèles obtenus par CNS et sur le modèle obtenu par prédiction *de novo*, a permis d'étudier la répartition de ces angles sur le diagramme de Ramachandran, confirmant la qualité des modèles obtenus (Figure 35).

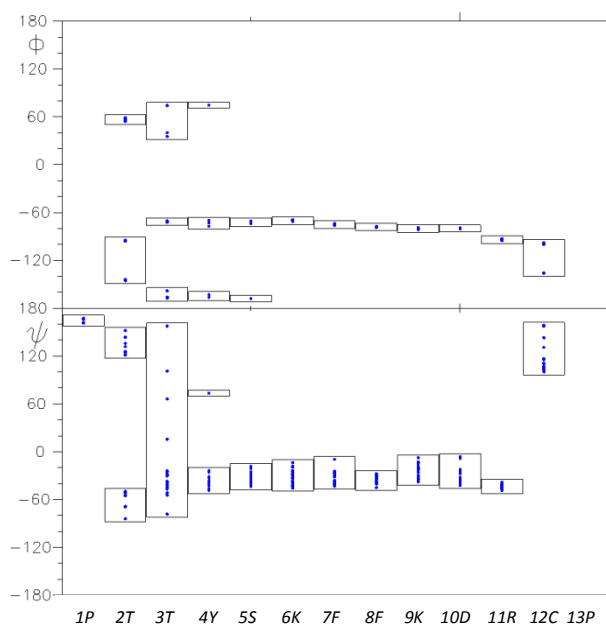


Figure 34 : Population des angles ( $\phi$ ,  $\psi$ ) sur les 22 structures modélisées retenues de PP13.



Figure 35 : Analyse par RamaDA du modèle de PP13 obtenu par prédiction *de novo*. L'intensité des barres de l'histogramme au-dessus de la séquence reflète la valeur du z-score de chaque acide aminé. La valeur moyenne obtenue pour l'ensemble du peptide est de 1.94.

La Figure 36, représentant l'ensemble des 22 structures modélisées, indique une quasi-parfaite superposition de la zone centrale du peptide, alors que les deux extrémités sont très flexibles (forme « entonnoir »). L'analyse des RMSD réalisée résidu par résidu sur le squelette traduit cette observation en montrant une forte déviation pour les résidus 1 et 2 (respectivement 2.03 et 1.04 Å), alors qu'elle est inférieure à 0.9 Å pour tous les autres.

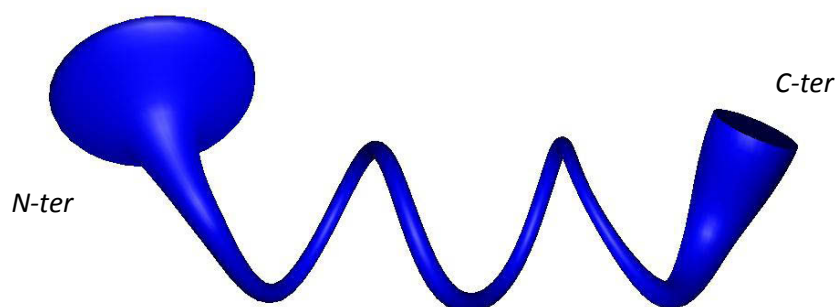


Figure 36 : Représentation « sausage » par MOLMOL [171] des 22 structures modélisées retenues de PP13.

L'analyse des structures secondaires des structures modélisées sous contraintes RMN révèle un caractère majoritaire en hélice  $\alpha$  pour les résidus 3 à 11, coil pour les résidus 1 et 13 et turn pour le résidu 2. Cette structuration est identique à celle observée lors de la modélisation *de novo* générée par le webserveur PEPFOLD, impliquant que les deux types de structures se superposent parfaitement (Figure 37).

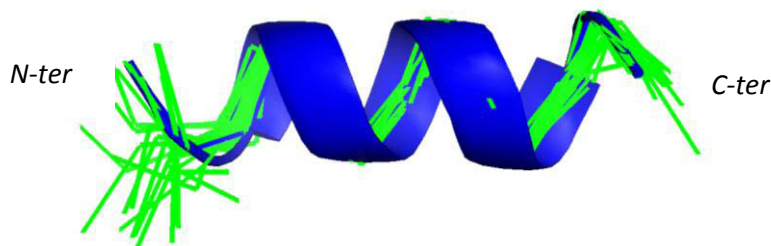


Figure 37 : Superposition de la structure de PP13 obtenue par modélisation *de novo* (cartoon, en bleu), avec les 22 structures générées par modélisation sous contraintes RMN (ribbon, en vert).

### 3.2.6 Poursuite de l'étude

---

Dans le cadre d'un projet collaboratif initié par Régis Gougeon à l'Université de Bourgogne, une étude de l'adsorption des protéines sur une argile utilisée pour la clarification des vins est actuellement en cours. Suite aux travaux précédemment réalisés [172, 173], l'interaction entre le peptide PP13 et la bentonite (montmorillonite) est évaluée par RMN du solide. Les spectres  $^{13}\text{C}$  montrant le peptide avant et après son adsorption (Figure 38) permettent d'observer une influence certaine de l'interface sur la structuration du squelette peptidique (zone des  $\text{C}_\alpha$  et des  $\text{C}=\text{O}$ ). L'étude de l'impact de l'adsorption sur la structure du peptide doit être maintenant approfondie mais les résultats préliminaires apparaissent déjà très prometteurs et pourraient permettre, à terme, d'établir un modèle et d'élucider les mécanismes d'interaction entre les protéines du raisin et cette argile. La poursuite de ce travail permettrait également de caractériser la structure spontanée de PP13 à l'état solide [174], en corrélation avec sa structuration observée en milieu liquide.

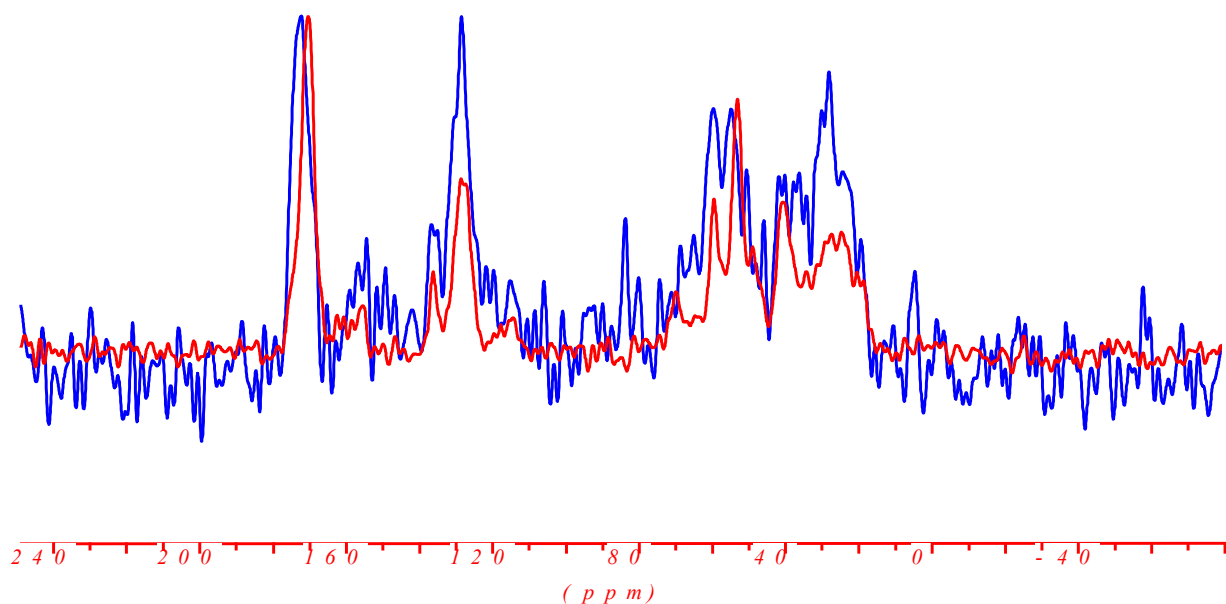


Figure 38 : Spectres  $^{13}\text{C}$  du peptide PP13 avant (en rouge) et après (en bleu) son adsorption sur l'argile, réalisés sur un spectromètre Bruker DSX 300. L'expérience  $^1\text{H}$ - $^{13}\text{C}$  CP-MAS (Cross-Polarization - Magic Angle Spinning) a été réalisée avec un temps de contact de 1 ms et un temps de recyclage de 2s pour le peptide et de 1s pour le mélange peptide/argile. 1120 scans ont été accumulés pour le peptide seul et 53120 dans le cas peptide/argile.

### 3.3 Conclusion

La structure de la protéine VVTL1 a été déterminée par modélisation par homologie et l'analyse de ses modes normaux a permis de révéler un mécanisme de type mâchoire autour d'une cavité acide, site potentiel de l'activité enzymatique de la protéine. Un feuillet  $\beta$  en épingle à cheveux (domaine III) isolé dans la structure s'est révélé être très conservé et absolument spécifique à la superfamille des protéines TL, et pourrait ainsi mener à l'élucidation complète du rôle biologique de ces protéines.

Cette étude mériterait certainement un complément expérimental au dichroïsme circulaire par cristallographie ou RMN mais les conditions pré-requises, l'obtention de cristaux d'une part et d'une protéine avec marquage isotopique d'autre part, ont limité la faisabilité de ces techniques dans le cadre de notre projet.

La structure du peptide PP13, modélisée sous contrainte RMN et *de novo*, confirme bien la structure observée de ce peptide intégré dans l'ensemble de la structure de la protéine VVTL1. Bien que l'exploitation de ce résultat seul soit relativement limitée, il ouvre la voie vers des projets de plus grande envergure impliquant l'utilisation de la RMN afin de caractériser les interactions entre la protéine VVTL1 et son environnement comme l'adsorption protéine/argile.





## Conclusion générale et perspectives

*On ne fait jamais attention à ce qui a été fait ;  
on ne voit que ce qui reste à faire.*

*Marie Curie, prix Nobel en 1903 (physique) et 1911 (chimie)*



L'objectif initial de la thèse était d'apporter une contribution à l'étude des protéines de raisin thaumatin-like et chitinase dans le cadre de l'œnologie. Nous avons ainsi réussi à obtenir des fractions pures de chacune des deux protéines isolées à partir de jus de raisin pouvant servir de standards pour leur quantification dans différents jus et Champagnes. L'élément essentiel à retenir est que la concentration de l'une ou l'autre diminue de l'ordre d'un facteur 100 au cours du procédé de vinification. Une question se pose alors : comment expliquer cette diminution visiblement non-sélective ? Les traitements de clarification des vins sur bentonite apportent certainement la majeure partie de la réponse en adsorbant *toutes* les protéines, mais des pistes en cours [175] suggèreraient la présence de protéases de levures pouvant attaquer les protéines du jus. L'étude de la double contribution élimination/dégradation sur un panel d'échantillons bien plus grand représente ainsi une piste de recherche qui permettrait de contrôler au mieux la quantité finale des protéines dans un Champagne. La quantification des autres protéines du raisin permettrait également d'obtenir le suivi du pool protéique total au cours du procédé de vinification. Le même suivi au cours des étapes de maturation des baies de raisin des trois cépages champenois Chardonnay, Pinot Noir et Pinot Meunier sous différents stress biotiques (agents pathogènes comme *Botrytis cinerea*) ou abiotiques (climat, traitements chimiques de la vigne) compléterait parfaitement l'étude et permettrait de comprendre les différences de concentration observées sur les jus de raisin.

Grâce aux protéines purifiées, nous avons pu également mener une étude de leurs propriétés tensioactives. La conclusion apparaît sans appel : la thaumatin-like, quand bien même un bon agent tensioactif, n'est pas en concentration suffisante dans un Champagne pour contribuer aux interfaces gaz/liquide ; la chitinase, quant à elle, ne montre strictement aucune propriété tensioactive. Ce résultat indique que ces deux protéines ne peuvent en aucun cas expliquer à elles seules la formation et la stabilisation des bulles et de la mousse d'un Champagne. Toute la compréhension des phénomènes de l'effervescence reste ainsi en suspens : la fraction macromoléculaire joue un rôle crucial mais quelles sont les protéines réellement impliquées ? Quels sont les effets synergiques ou antagonistes protéine/protéine ou protéine/autres composés du Champagne ? Pouvoir établir le lien entre la composition d'un Champagne – voire encore mieux celle d'un jus de raisin – et la qualité finale de la mousse et des bulles lors de la dégustation est un challenge pour les scientifiques mais un enjeu essentiel pour les professionnels de l'œnologie.

Les questions sur les mécanismes chimiques, biologiques ou physiques posées par les interactions des protéines de raisin avec leur environnement dans les baies, au cours de la vinification et dans le Champagne final pourraient trouver des réponses grâce à la biologie structurale. Cet outil, jusqu'alors non-exploité en œnologie, présente un formidable potentiel que nous avons appliqué à l'étude de la protéine thaumatin-like VVTL1. Après avoir proposé un modèle de sa structure tertiaire par homologie, l'analyse de ses modes normaux nous a permis d'identifier trois domaines conformationnels très flexibles. Les deux premiers sont impliqués dans un mécanisme de mâchoire autour d'une cavité acide, site potentiel pour l'action enzymatique de la protéine. Il serait nécessaire de poursuivre cette approche par une étude de l'interaction enzyme/substrat par docking dynamique basée sur les travaux réalisés chez d'autres espèces végétales que la vigne [81, 83]. Le troisième domaine flexible, un feuillet  $\beta$  en épingle à cheveux isolé dans la structure, s'est révélé être très conservé et absolument spécifique à la superfamille des protéines thaumatin-like. Gageons que ce motif pourrait représenter une clé majeure pour ouvrir la voie de l'élucidation complète du rôle biologique de cette famille.

La décomposition de la protéine VVTL1 en différents domaines peptidiques nous est ainsi apparue comme une approche prometteuse. Différents peptides sélectionnés dans chacun des trois domaines de VVTL1 ont été analysés par RMN, avec plus ou moins de succès. Nous avons retenu et présenté un peptide replié en hélice  $\alpha$ , situé sur la partie extérieure du domaine II. La modélisation sous contraintes RMN de ce peptide permet d'établir un résultat préliminaire à l'étude de son adsorption sur la bentonite par RMN du solide. La suite de ce projet est actuellement en cours, menée par l'IUVV à Dijon en collaboration avec l'URCA. La compréhension du mécanisme d'adsorption de ce peptide à l'interface liquide/solide pourra servir de premier modèle, avant l'étude plus ambitieuse sur des protéines de raisin entières.

L'étude structurale de VVTL1 semble ainsi bien entreprise et pleine de promesses. Cependant un point crucial subsiste, jusqu'à présent – discrètement – passé sous silence : *quid* des modifications post-traductionnelles ? En effet, une protéine acquiert sa structure finale *in vitro*, stabilisée par des interactions intramoléculaires et avec son environnement. Des modifications post-traductionnelles dans la cellule peuvent intervenir à ce moment-là et modifier le repliement général de la protéine. Les protéines de raisin sont connues pour être soumises à ce type de modification, notamment l'apparition de glycosylations (addition d'une chaîne de sucres sur un acide aminé de la séquence) [176]. Ces glycosylations, en modifiant potentiellement la structure tertiaire, peuvent modifier les propriétés biologiques et physico-chimiques des protéines. Il apparaît donc nécessaire de prendre en compte ce nouvel aspect dans la suite de l'étude structurale de VVTL1 par une approche combinée spectrométrie de masse/RMN/bioinformatique. Il serait également intéressant d'envisager l'assemblage de plusieurs chaînes protéiques qui formerait une structure quaternaire, phénomène observable sur plusieurs homologues de protéines thaumatin-like. De manière assez triviale, toute l'étude autour de la structure de VVTL1 pourrait servir de modèle à l'étude des autres protéines de raisin, à commencer par la chitinase.

La poursuite de l'ensemble des aspects de biologie structurale passera inévitablement par une étape de biologie moléculaire et de production de protéines recombinantes. La RMN pourrait ainsi développer tout son potentiel : de la caractérisation structurale sur des protéines avec marquage isotopique à l'étude des interactions protéine/autre composé sur des protéines totalement ou partiellement deutérées. De plus, la séquence en acides aminés d'une protéine recombinante est déterminée aisément alors que le séquençage total d'une protéine simplement purifiée représente un challenge et n'a pas été réalisé dans le cadre de notre étude. La relation séquence-structure-fonction des protéines pourrait être ainsi explorée par une étude comparative menée entre une protéine produite en système hétérologue et celle issue de jus de raisin. La caractérisation d'éventuelles différences d'activité enzymatique pourrait être corrélée à une différence de séquence ou de structure (modifications post-traductionnelles ?) ou à la nécessité de la présence d'un co-facteur dans les baies. L'obtention d'une grande quantité de protéine pure pourrait également orienter l'étude vers la cristallographie, sous réserve de l'obtention de cristaux.

Il est ainsi certain que le monde du Champagne et des protéines de raisin nous réserve encore bien des surprises... et quoi de plus encourageant que de savoir que les fantastiques découvertes à venir seront célébrées autour de quelques bulles dorées !

# Annexes



# 1 Données supplémentaires relatives à l'article « Quantification of chitinase and thaumatin-like proteins in grape juices and wines »

## Analytical and Bioanalytical Chemistry

### Electronic Supplementary Material

## Quantification of chitinase and thaumatin-like proteins in grape juices and wines

D. Le Bourse, A. Conreux, S. Villaume, P. Lameiras, J.-M. Nuzillard and P. Jeandet

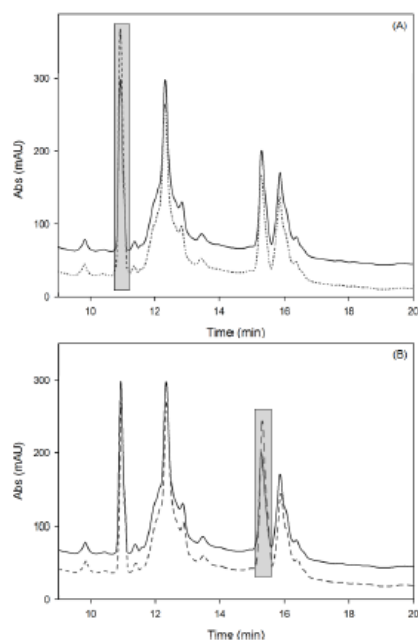
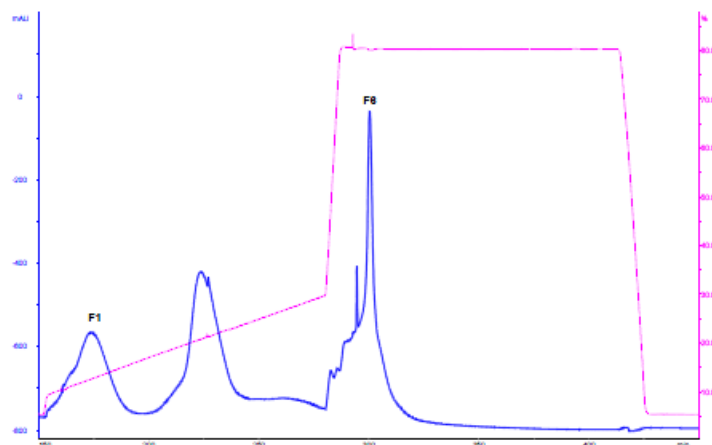
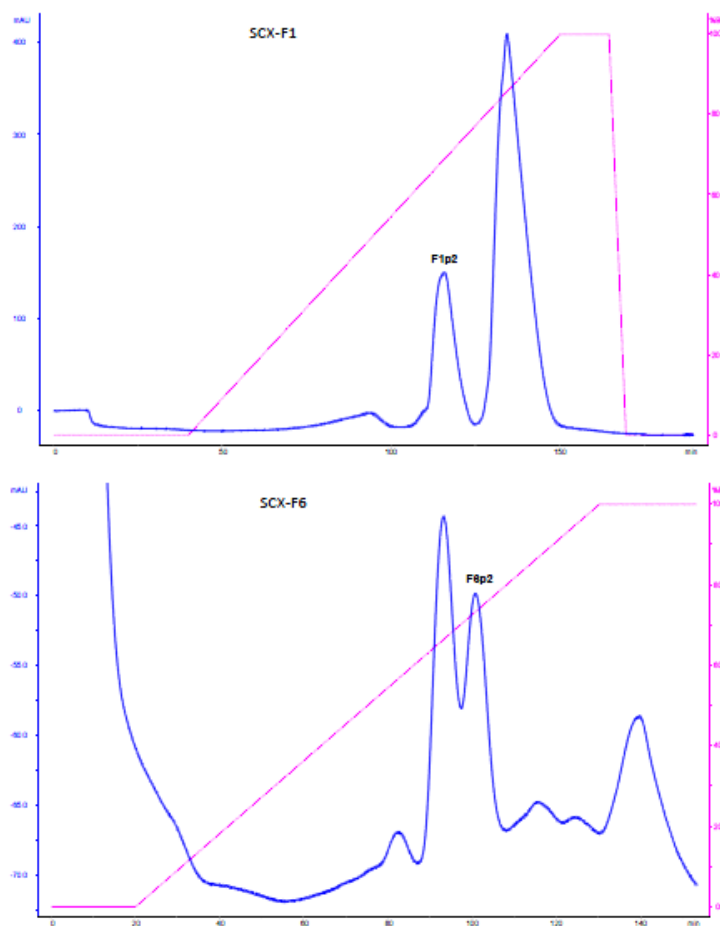


Fig. S1 UHPLC (C18) profiles for the grape juice Chardonnay 2009 with extra amounts of pure TL (chromatogram A) and chitinase (chromatogram B) to assess recovery and specificity of the quantification method. Load and wash are omitted.



**Fig. S2** Cation exchange (SCX) chromatogram (280 nm) for the grape juice Chardonnay 2009 (19.4 L loaded). The dotted line indicates the salt/pH gradient. Collected fractions (F1 and F6) were used for HIC.



**Fig. S3** HIC chromatograms (280 nm) for SCX fractions F1 and F6. Collected fractions (F1p2 and F6p2) were used for HPLC.



## 2 Modélisation et simulations moléculaires pour l'étude de la structure de protéines et de peptides

La modélisation moléculaire regroupe les méthodes servant à décrire et à prédire le comportement lors de simulations dynamiques de molécules et de systèmes macromoléculaires comme les peptides et les protéines en évaluant leur énergie. Nous détaillerons ici les principes de chaque technique utilisée dans le cadre de l'étude de la protéine VVTL1 et du peptide PP13.

### 2.1 Principe général de la mécanique moléculaire

L'énergie d'une molécule dépend de la position dans l'espace des atomes qui la composent, plus elle est basse plus la configuration spatiale de ladite molécule a de chance d'exister et d'être stable.

En mécanique quantique, un système moléculaire (enchaînement d'atomes constitués de noyaux et d'électrons) peut être totalement caractérisé par la résolution de l'équation de Schrödinger :

$$H \Psi = E \Psi$$

où  $\Psi$  désigne la fonction d'onde,  $H$  l'opérateur hamiltonien qui s'applique à  $\Psi$  et  $E$  l'énergie du système.

Cependant, cette équation ne peut être résolue analytiquement dans le cas de systèmes moléculaires complexes. On a alors recours à un ensemble d'approximations plus ou moins importantes. Si on considère que les électrons s'adaptent instantanément aux mouvements des noyaux (protons+neutrons) l'énergie totale du système peut s'écrire comme la somme de l'énergie nucléaire et de l'énergie électronique, ce qui définit l'approximation de Born-Oppenheimer.

Quand les noyaux des atomes sont considérés comme fixes, les calculs sont concentrés sur le mouvement des électrons. Ce principe est à la base de la chimie quantique, mise en œuvre dans le cadre des méthodes *ab initio* permettant de calculer la structure d'une molécule uniquement à partir de sa formule chimique.

Une autre approche considère l'étude des positions des atomes entiers (et non plus divisés en système noyaux/électrons), selon un principe de mécanique classique (et non plus quantique). La mécanique moléculaire est ainsi basée sur la description de ce qu'est une molécule : les atomes sont assimilés à des points massiques potentiellement chargés reliés entre eux par des liaisons sous forme de ressorts.

#### 2.1.1 Champ de forces

L'énergie totale d'un système  $E_{tot}$ , représentée par la somme de son énergie potentielle  $E_{pot}$  et de son énergie cinétique  $E_{cin}$  reste constante au cours du temps :

$$E_{tot} = E_{pot} + E_{cin} = \text{constante}$$

Lorsqu'un système est à l'équilibre, traduisant une conformation stable, son énergie cinétique est nulle et la minimisation de son énergie totale revient à la minimisation de son énergie potentielle. La minimisation du terme d'énergie potentielle seule induit donc l'augmentation de la stabilité du système total. L'ingrédient essentiel de la prédiction de structure d'une protéine ou d'un peptide est ainsi la définition d'une fonction d'énergie permettant d'évaluer la stabilité de la conformation.

La mécanique moléculaire s'appuie sur une représentation simplifiée de la surface de l'énergie. Selon ce principe, à chaque état conformationnel correspond une énergie potentielle empirique décrivant l'ensemble des interactions attractives et répulsives entre les atomes du système. L'énergie potentielle empirique est alors calculée comme la somme des différents termes énergétiques d'interactions plus ou moins approximatifs. La fonction mathématique décrivant ce calcul est appelé « champ de forces ». Sous ce terme sont regroupés deux éléments : d'une part l'expression des différentes fonctions contribuant au calcul énergétique et d'autre part les valeurs des différentes constantes paramétrant ces fonctions. Différents champs de forces peuvent être utilisés dans les différents types de modélisation des protéines et des peptides, les plus connus étant CHARMM (« Chemistry at Harvard Molecular Mechanics ») [177] et AMBER (« Assisted Model Building with Energy Refinement ») [178].

### 2.1.2 Minimisation de l'énergie

En mécanique moléculaire, la minimisation de l'énergie d'un système implique l'utilisation d'un algorithme mettant en place une succession de calculs itératifs pour une molécule dans une conformation initiale donnée qui optimisent la géométrie du système en le relaxant. Tous les paramètres (degrés de liberté) définissant la géométrie du système sont ainsi systématiquement modifiés par de petits incréments jusqu'à ce que l'énergie atteigne un minimum (Figure 39).

Parmi les algorithmes de minimisation existants, les deux plus utilisés reposent sur la dérivée de l'énergie potentielle selon la méthode de la plus grande pente « Steep Descent » [179] ou la méthode du gradient conjugué [180].

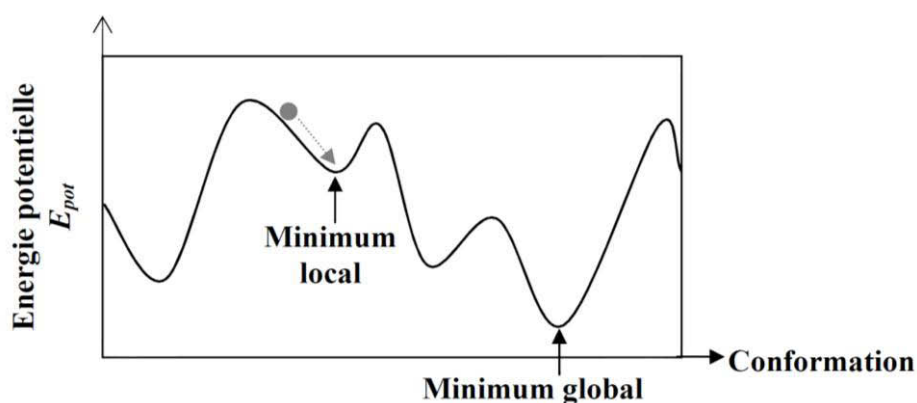


Figure 39 : Variation de l'énergie potentielle d'un système selon sa conformation. La boule grise indique l'énergie potentielle de la conformation initiale du système. La minimisation de l'énergie selon un algorithme basé sur les gradients de surface Steep Descent ou Powell permet de déterminer un minimum local d'énergie proche de la conformation initiale. Reproduit avec l'aimable autorisation de Pedro Lameiras.

Pour pouvoir se dégager des minima énergétiques locaux, il est nécessaire d'utiliser une méthode permettant d'explorer tout l'espace conformationnel avant l'application d'un algorithme de minimisation : la dynamique moléculaire.

### 2.1.3 Dynamique moléculaire

La dynamique moléculaire a pour objectif de simuler les mouvements internes d'un système en fonction du temps par le calcul du déplacement des atomes qui le composent. Elle repose sur les principes de la mécanique classique Newtonienne, selon laquelle la trajectoire de chaque atome est définie par la relation fondamentale de la dynamique :

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d\vec{v}_i}{dt} = m_i \frac{d^2 \vec{r}_i}{dt^2}$$

où  $\vec{F}_i$  est la force s'appliquant sur l'atome  $i$ ,  $m_i$  sa masse,  $\vec{a}_i$  son accélération,  $\vec{v}_i$  sa vitesse et  $\vec{r}_i$  sa position en fonction du temps  $t$ .

La force s'appliquant sur un atome dépendant de l'énergie potentielle selon :

$$\vec{F}_i = - \frac{dE_{pot}}{d\vec{r}_i}$$

il est possible de calculer l'accélération s'exerçant sur chaque atome à tout instant. Plusieurs algorithmes sont proposés pour résoudre ces équations, notamment celui de Verlet [181] dans lequel est utilisé un développement en série de Taylor du second ordre pour le vecteur position :

$$\vec{r}(t+\delta t) = 2\vec{r}(t) - \vec{r}(t-\delta t) + \delta t^2 \vec{a}(t)$$

$$\vec{v}(t+\delta t) = \vec{v}(t) + 0.5\delta t[\vec{a}(t) + \vec{a}(t+\delta t)]$$

En connaissant la position à l'instant  $t$ , il est alors possible de calculer la position et la vitesse à  $t+\delta t$ . Cet algorithme implique l'application d'un pas d'intégration  $\delta t$  très court (de l'ordre de la femtoseconde) afin de maintenir une trajectoire stable.

Dans une simulation de dynamique moléculaire, l'énergie cinétique totale d'un système peut être exprimée selon :

$$E_{cin} = \sum_i \frac{1}{2} m_i v_i^2 = \frac{3}{2} N k_B T$$

où  $N$  représente le nombre d'atomes dans le système,  $k_B$  la constante de Boltzman et  $T$  la température. L'énergie totale d'un système isolé étant constante, il sera ainsi possible d'intervenir sur l'énergie potentielle à partir de l'énergie cinétique.

Cette équation introduit donc la notion de température au sein des méthodes de dynamique moléculaire. Plus cette température est élevée, plus la vitesse communiquée aux atomes sera importante (phénomène d'agitation thermique). La méthode de recuit simulé [169] exploite ce phénomène par l'intermédiaire d'une hausse « fictive » de température, fournissant ainsi suffisamment d'énergie cinétique au système pour lui permettre de franchir les barrières d'énergie potentielle et donc d'explorer un espace conformationnel beaucoup plus large.

## 2.2 Modélisation d'une structure

### 2.2.1 Modélisation sous contraintes RMN

#### ▪ Description du champ de forces

Le calcul de l'énergie potentielle d'un système dans le cas de la modélisation sous contraintes RMN est réalisé par l'intermédiaire d'un champ de forces classique  $E_{pot}$  équivalent à celui utilisé pour les simulations de dynamique moléculaire, auquel est ajouté un terme qui permet de contrôler le respect de ces contraintes  $E_{cont}$  :

$$E = E_{pot} + E_{cont}$$

#### Fonction d'énergie potentielle $E_{pot}$

L'énergie potentielle empirique  $E_{pot}$  est exprimée comme la somme des termes représentant les interactions entre les atomes liés  $E_{liés}$  d'une part, et les atomes non-liés  $E_{non-liés}$  d'autre part :

$$E_{pot} = E_{liés} + E_{non-liés}$$

Ces deux termes sont décrits par les composantes du champ de forces CHARMM [177]:

$$E_{liés} = E_{liaisons} + E_{angles} + E_{dièdres} + E_{impropres}$$

$$E_{non-liés} = E_{van\_der\_Waals} + E_{électrostatique}$$

Les fonctions d'énergie associées à chacun de ces termes sont décrites dans le Tableau 13.

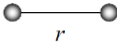
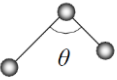
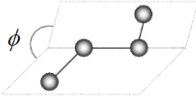

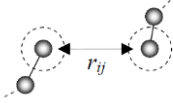
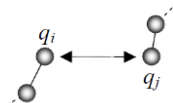
Terme du champ de forces	Expression générale	Description	Illustration
E <sub>liés</sub>	Energie de liaisons $E_{liaisons} = \sum k_r(r-r_0)^2$  $k_r$ : constante de force $r_0$ : longueur de la liaison à l'équilibre	Définit l'énergie mise en jeu par les variations de la longueur $r$ des liaisons après extension et compression à partir de leurs valeurs d'équilibre	
	Energie de déformation des angles de valence $E_{angles} = \sum k_\theta(\theta-\theta_0)^2$  $k_\theta$ : constante de force $\theta_0$ : angle à l'équilibre	Définit l'énergie impliquée dans les variations d'angle $\theta$ entre deux liaisons contiguës autour de sa valeur à l'équilibre	
	Energie de torsion $E_{dièdres} = \sum k_\phi(1+\cos n\phi)$  $k_\phi$ : constante de force $n$ : périodicité	Définit l'énergie impliquée dans la rotation d'une partie de la molécule autour d'une liaison	
	Energie des angles impropres $E_{impropres} = \sum k_\omega(\omega-\omega_0)^2$  $k_\omega$ : constante de force $\omega_0$ : valeur idéale définie dans le champ de forces	Assure le maintien de la planéité ou de la chiralité d'un groupe d'atomes	
E <sub>non-liés</sub>	Energie de van der Waals $E_{van\_der\_Waals} = \sum_{ij} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right)$ $A_{ij}, B_{ij}$ : paramètres dépendant de la nature des atomes $i$ et $j$ , distants de $r_{ij}$	Traduit l'attraction et la répulsion entre deux atomes constituant un dipôle.	
	Energie des interactions électrostatiques $E_{électrostatiques} = \sum_{ij} \left( \frac{q_i q_j}{4\pi\epsilon_{ij} r_{ij}} \right)$  $q_i, q_j$ : charge partielle des atomes $i$ et $j$ , distants de $r_{ij}$ $\epsilon_{ij}$ : constante diélectrique (solvant/molécule)	Potentiel coulombien entre deux atomes porteurs d'une charge nette.	

Tableau 13 : Expression du champ de forces.

Les termes associés aux interactions entre atomes liés permettent en premier lieu de maintenir la géométrie covalente de la molécule. En second lieu, les termes associés aux interactions entre atomes non-liés rendent compte des interactions de type van der Waals et électrostatiques qui peuvent être observées entre deux atomes très éloignés en distance. La prise en compte de l'ensemble des combinaisons possibles entre paires d'atomes d'une macromolécule impliquerait un temps de calcul considérable. Pour réduire la durée des calculs, il est possible de limiter la portée des interactions en définissant « un seuil de coupure ». Selon ce principe, seules les interactions entre atomes séparés par une distance inférieure à ce seuil seront prises en compte par le champ de forces. Aussi, pour éviter des discontinuités d'énergie induites par l'introduction de ces seuils, la fonction d'énergie  $E_{non-liés}$  est multipliée par une fonction d'amortissement ou de « switch » [182]. Ce type de fonction permet de faire tendre progressivement les énergies de van der Waals et électrostatiques vers zéro, et donc d'éviter des coupures brusques pouvant générer des erreurs de calcul.

### Intégration des contraintes RMN : $E_{cont}$

Les contraintes expérimentales  $E_{NOE}$  et  $E_{dih}$  sont introduites dans le calcul du champ de forces sous forme d'intervalles de valeurs permises pour les distances interatomiques et les angles dièdres :

$$E_{cont} = E_{NOE} + E_{dih}$$

Le non-respect d'une contrainte conduit à une pénalité énergétique, appelée violation, qui informe de l'écart entre une valeur de distance ou d'angle dièdre dans une structure calculée et l'intervalle de valeurs autorisées correspondant.

#### ➤ Terme des contraintes de distance

Le logiciel CNS propose plusieurs types de fonctions potentielles pour exprimer le terme énergétique des contraintes de distance  $E_{NOE}$ . Dans le cadre de notre étude nous avons utilisé la fonction :

$$E_{NOE} = k_{NOE} \delta^n \quad \text{avec} \quad \left| \begin{array}{l} \delta = r_{min} - d \text{ pour } d < r_{min} \\ \delta = d - r_{max} \text{ pour } d > r_{max} \end{array} \right.$$

où  $r_{min}$  et  $r_{max}$  représentent les limites de l'intervalle de distances permises pour une contrainte donnée,  $d$  la distance correspondante dans le modèle,  $k_{NOE}$  la constante de force, et  $n$  est généralement égal à 2. Notons par ailleurs que le terme  $E_{NOE}$  est systématiquement nul lorsque la valeur de distance  $d$  est encadrée par les valeurs  $r_{min}$  et  $r_{max}$ .

#### ➤ Terme des contraintes angulaires

Le terme énergétique  $E_{dih}$  lié aux contraintes angulaires  $\phi$  est représenté par une fonction potentielle carrée comparable à celle définie pour les contraintes de distance :

$$E_{dih} = k_{dih} (\phi - \phi_x)^2$$

avec  $\phi_x$  la limite minimum ou maximum de l'intervalle de valeurs autorisées pour une contrainte donnée,  $\phi$  la valeur de l'angle dans le modèle et  $k_{dih}$  la constante de force associée.

## ▪ Protocole de recuit simulé

### Génération des structures initiales aléatoires

L'étape préalable aux calculs de dynamique moléculaire consiste à générer des structures initiales par tirage aléatoire des angles  $\phi$  et  $\psi$ . La topologie de la protéine ou du peptide est obtenue à partir d'un champ de forces simplifié tel que :

$$E_{pot} = E_{liaisons} + E_{angles} + E_{impropres} + E_{repele}$$

Cette fonction exclut les termes  $E_{dièdres}$  et  $E_{électrostatiques}$  uniquement pris en compte dans le champ de forces CHARMM22 utilisé pour le raffinement des structures. Par ailleurs, le terme des interactions de van der Waals est remplacé par un potentiel continu purement répulsif ( $E_{repele}$ ) similaire à la composante répulsive du potentiel de Lennard-Jones. Ce terme est associé à une constante de force  $k_{van\_der\_Waals}$  et à un facteur multiplicateur du rayon minimum de van der Waals  $k_{repele}$ .

### Recuit simulé

Le protocole de recuit simulé s'inspire de la technique métallurgique de « recuit » qui consiste à chauffer puis refroidir très lentement un matériau afin de lui permettre d'organiser au mieux sa structure cristalline.

La méthode du recuit simulé consiste dans un premier temps à mener une dynamique à haute température (de 300 à 50000 K) à partir de la structure initiale préalablement générée et minimisée. Cet apport d'énergie cinétique permet d'explorer tout l'espace conformationnel et de se dégager des minima locaux proches des structures initiales. Après une période d'équilibration, la température est lentement diminuée jusqu'à 300 K, par paliers de 250 K, afin de réduire l'énergie cinétique et de permettre au système d'occuper les états de plus basse énergie. Finalement, les dernières étapes du recuit simulé consistent à minimiser puis à collecter les structures ainsi calculées. Il est cependant important de noter que cette méthode ne garantit pas l'identification du minimum énergétique global, mais s'en approchera probablement dès lors que le jeu de contraintes expérimentales est optimisé (diminution cohérente des degrés de liberté), et qu'un ensemble suffisant de conformations est collecté. Cette procédure est illustrée dans la Figure 40.

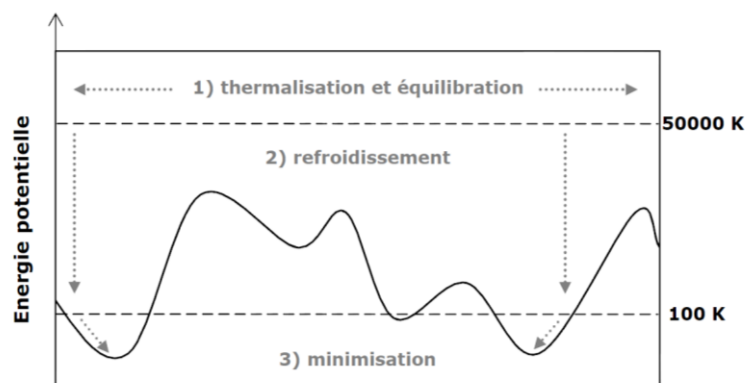


Figure 40 : Illustration de la procédure de recuit simulé par la représentation schématique d'une surface d'énergie potentielle. La dynamique à haute température permet d'explorer l'espace conformationnel et de s'éloigner de la structure initiale (non-représentée). Les structures générées sont lentement refroidies de manière à occuper les zones stables de la surface d'énergie. Une dernière étape de minimisation permet d'optimiser les conformations et d'obtenir finalement des minima énergétiques plus profonds.

## 2.2.2 Modélisation par homologie

Le logiciel Modeller permet de construire des modèles de structures de protéines par satisfaction des contraintes spatiales issues de structures de protéines homologues connues. Il utilise pour cela le champ de forces CHARMM (présenté dans la section 1.2.1), auquel sont ajoutés des termes énergétiques de contraintes de distance, d'angles et d'angles dièdres, ainsi que des contraintes définies pour des atomes ou des pseudo-atomes particuliers. Les représentations mathématiques des différentes contraintes calculées par Modeller sont représentées sur la Figure 41 [130, 183]. Le modèle 3D est obtenu par l'optimisation de la « molecular Probability Density Function » (pdf). La fonction pdf est définie et optimisée selon une procédure dans l'espace cartésien qui utilise la méthode des gradients conjugués et une dynamique moléculaire par recuit simulé.

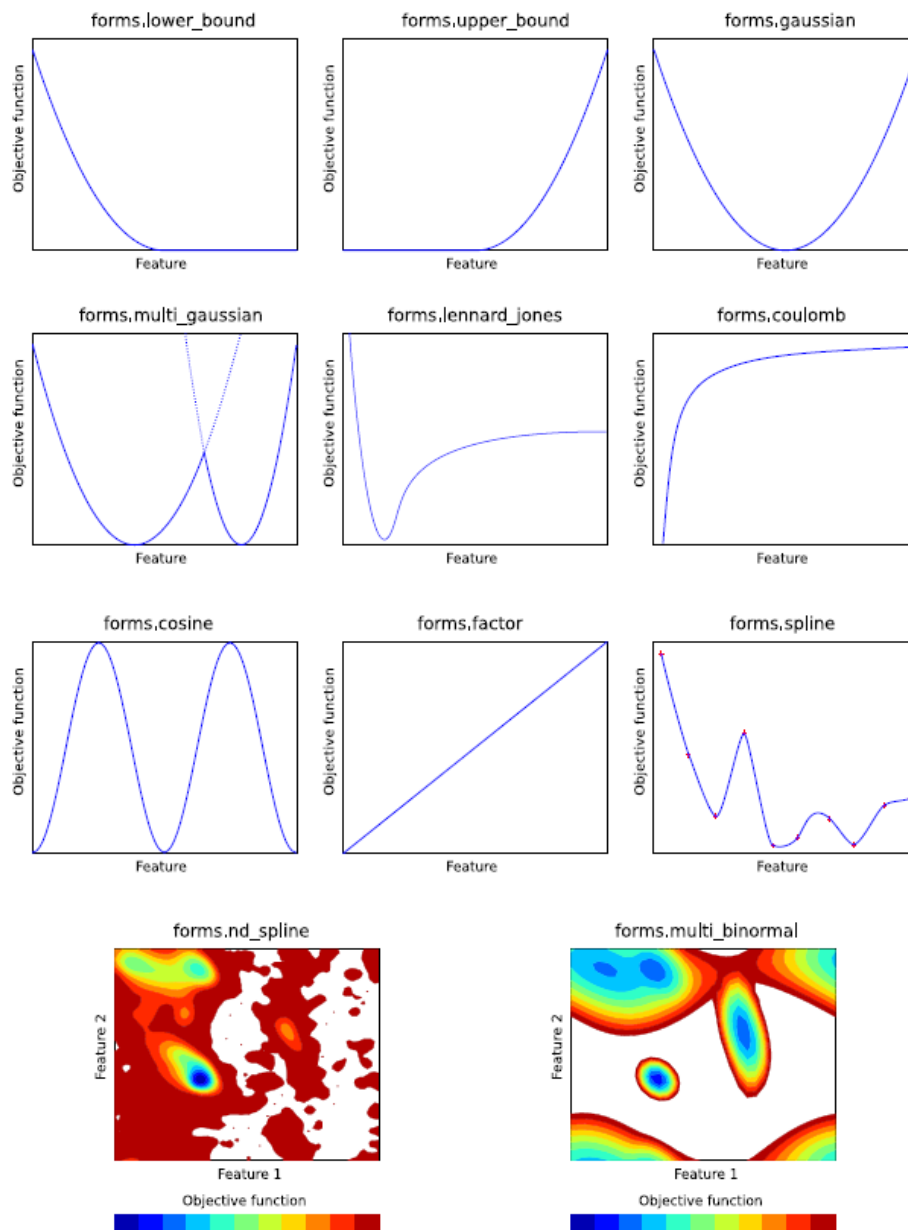


Figure 41 : Représentation mathématique des contraintes utilisées par Modeller. Chaque contrainte apporte une contribution à la fonction objective comme fonction d'un ou plusieurs paramètres. Reproduit du manuel d'utilisateur <http://salilab.org/modeller/>.



### 2.2.3 Modélisation de novo

Le webserveur PEPFOLD propose une génération *ab initio* de modèles protéiques à partir de représentations discrètes des protéines et de critères d'énergie simplifiés [138, 139].

Cette méthode se base sur la notion d'alphabet structural, *i.e.* un ensemble restreint de conformations prototypes permettant de décrire l'ensemble des structures protéiques. L'alphabet structural HMM-SA (*Hidden Markov Model Structural Alphabet*) a été appris par l'intermédiaire de chaînes de Markov cachées [140-142]. L'apprentissage de la dépendance entre les lettres de l'alphabet structural et les séquences en acides aminés qui les émettent a permis de prédire l'ensemble des lettres de l'alphabet structural pouvant décrire une séquence protéique donnée (Figure 42).

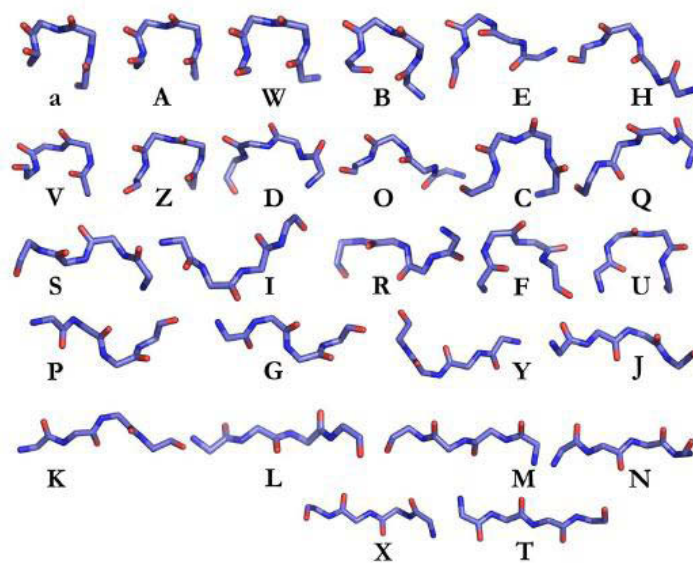


Figure 42 : Les 27 lettres de l'alphabet structural HMM-SA. Ne sont représentés dans cette figure que les prototypes moyens correspondant à chaque état.

Ainsi, la première étape de la méthode est de prédire l'ensemble des fragments à assembler pour pouvoir générer un modèle protéique complet. Pour cette étape, l'algorithme SAFrAN (*Structural Alphabet candidate Fragments from AmiNo-acid sequence*) est utilisé. Il permet la recherche de fragments compatibles avec une séquence HMM-SA prédite sous contrainte de la prédiction des structures secondaires par le logiciel PSIPRED [184].

Lors de la deuxième étape, l'assemblage des fragments prédits est envisagé par un algorithme glouton [143, 144], guidé par une version simplifiée du potentiel énergétique OPEP (*Optimized Potential for Efficient protein structure Prediction*) [185].

Le champ de forces OPEP est un potentiel gros grains, exprimé comme la somme de l'énergie locale  $E_{locale}$ , non-liée  $E_{non-liés}$  et associée aux liaisons hydrogènes  $E_{liaisons-H}$  :

$$E = E_{locale} + E_{non-liés} + E_{liaisons-H}$$

Le calcul de l'énergie locale s'exprime comme :

$$E_{locale} = w_l \sum_{liaisons} k_b (r - r_{eq})^2 + w_a \sum_{angles} k_\alpha (\alpha - \alpha_{eq})^2 \\ + w_\Omega \sum_{torsions-imp} k_\Omega (\Omega - \Omega_{eq})^2 + w_{\phi,\psi} \left( \sum_{\phi} E_\phi + \sum_{\psi} E_\psi \right)$$

Le terme  $E_{locale}$  contient des constantes de forces associées à un changement des longueurs et angles de liaison de toutes les particules ainsi que des constantes de forces en relation avec un changement des torsions impropres des chaînes latérales et de la liaison peptidique. Les constantes de forces et les valeurs d'équilibre associées aux atomes de la chaîne principale sont issues d'AMBER [186] ; les constantes de forces associées aux chaînes latérales sont très similaires à celles d'AMBER.

Les termes  $E_\phi$  et  $E_\psi$  reproduisent le diagramme de Ramachandran dans le modèle gros grain, en accord avec les structures protéiques complètes ; ils sont exprimés par les polynômes quadratiques suivants :

$$E_\phi = k_{\phi\psi} (\phi - \phi_0)^2 \\ E_\psi = k_{\phi\psi} (\psi - \psi_0)^2$$

L'énergie associée aux interactions non liantes s'exprime par :

$$E_{non-liés} = w_{1,4} \sum_{1,4} E_{vdW} + w_{C_\alpha C_\alpha} \sum_{C_\alpha, C_\alpha} E_{vdW} + w_{1>4} \sum_{CP', CP'} E_{vdW} + w_{1>4} \sum_{CP', C_\alpha} E_{vdW} \\ + w_{1>4} \sum_{CL, CL} E_{vdW} + \sum_{CL, CL} w_{CL, CL} E_{vdW}$$

avec 1,4 les interactions de type 1-4 pour chaque degré de liberté le long des angles de torsion,  $CP'$  les atomes  $N$ ,  $HN$ ,  $C''$  et  $O$  le long de la chaîne principale ( $CP = CP' + C_\alpha$ ), et  $CL$  les chaînes latérales. Les interactions de courte portée sont séparées des interactions de longue portée ( $j > i+4$ ), et le  $C_\alpha$  est traité séparément des autres atomes de la chaîne principale.

Les interactions de type van der Waals ( $E_{vdW}$ ) sont définies comme :

$$E_{vdW} = \begin{cases} -\epsilon_{ij} \left( \frac{r_{ij}^0}{r_{ij}} \right)^6, & \text{pour } \epsilon_{ij} < 0, \\ \epsilon_{ij} \left( \left( \frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{ij}^0}{r_{ij}} \right)^6 \right) & \text{sinon.} \end{cases}$$

avec  $r_{ij}$  la distance entre les particules  $i$  et  $j$ ,  $r_{ij}^0 = r_i^0 + r_j^0$ , où  $r_i^0$  est le rayon de van der Waals de la particule  $i$ . Toutes les interactions non liantes sont considérées, sans distance seuil (cut-off).

L'énergie associée aux liaisons hydrogène ( $E_{liaisonsH}$ ) est composée de deux termes : un terme classique à deux corps ( $E_{LH1}$ ), et un terme plus original à quatre corps ( $E_{LH2}$ ).

Le terme à deux corps est défini par :

$$E_{LH1} = w_{lh1=4} \sum_{ij, j=i+4} \epsilon_{lh1=4} \mu(r_{ij}) v(\alpha_{ij}) + w_{lh1>4} \sum_{ij, j>i+4} \epsilon_{lh1>4} \mu(r_{ij}) v(\alpha_{ij})$$

où :

$$\mu(r_{ij}) = 5 \left( \frac{\sigma}{r_{ij}} \right)^{12} - 6 \left( \frac{\sigma}{r_{ij}} \right)^{10}$$

$$v(\alpha_{ij}) = \begin{cases} \cos^2 \alpha_{ij}, & \alpha_{ij} > 90^\circ \\ 0, & \text{sinon.} \end{cases}$$

La somme est effectuée sur tous les résidus  $i$  et  $j$  séparés par  $j = i + 4$  et  $j > i + 4$  (les hélices  $3_{10}$  sont donc exclues),  $r_{ij}$  est la distance  $O \cdots H$  (l'oxygène du groupement carbonyle et l'hydrogène du groupement amide),  $\alpha_{ij}$  l'angle  $\overline{NHO}$  et  $\sigma$ , la valeur d'équilibre de la distance  $O \cdots H$ .

L'effet à quatre corps mimant l'effet coopératif entre les liaisons hydrogène  $ij$  et  $kl$  est défini par :

$$E_{LH2} = \sum \epsilon_{\alpha}^{coop} \exp\left(-\frac{(r_{ij} - \sigma)^2}{2}\right) \exp\left(-\frac{(r_{kl} - \sigma)^2}{2}\right) \Delta(ijkl)$$

$$+ \sum \epsilon_{\beta}^{coop} \exp\left(-\frac{(r_{ij} - \sigma)^2}{2}\right) \exp\left(-\frac{(r_{kl} - \sigma)^2}{2}\right) \Delta'(ijkl)$$

Les paramètres  $\epsilon_{\alpha}^{coop}$  et  $\epsilon_{\beta}^{coop}$  sont exprimés par :

$$\epsilon_{\alpha}^{coop} = w_{\alpha}^{coop} E_{\alpha}^{coop} + \sum_R w_{\alpha}^R E_{\alpha}^R$$

$$\epsilon_{\beta}^{coop} = w_{\beta}^{coop} E_{\beta}^{coop} + \sum_R w_{\beta}^R E_{\beta}^R$$

$E_{\alpha}^{coop}$  et  $E_{\beta}^{coop}$  sont les énergies coopératives, indépendamment des acides aminés impliqués, et  $E_{\alpha}^R$  et  $E_{\beta}^R$  sont les potentiels associés aux propensités du résidu  $R$  pour les hélices  $\alpha$  et les feuillets  $\beta$  respectivement.

## 2.3 Evolution dynamique d'une structure dans une boîte d'eau

Afin de solvater le modèle de la protéine VVTL1, nous avons utilisé le programme NAMD (module implémenté dans VMD [133]) et le champ de forces CHARMM associé. La minimisation de l'énergie a été effectuée selon la méthode du gradient conjugué.

### 2.3.1 Traitement de l'eau

Le modèle d'eau utilisé a été le modèle TIP3P [187]. Ce modèle rigide prend en compte les interactions coulombiennes et les expressions de Lennard-Jones, l'angle entre les atomes d'hydrogène est maintenu constant. Il comporte trois sites d'interactions électrostatiques : charge négative de l'oxygène, charges partielles positives sur des hydrogènes. L'interaction de van der Waals entre deux molécules d'eau sera calculée d'après la fonction de Lennard-Jones avec seulement un simple point d'interaction par molécule centré sur l'atome d'oxygène.

Pour éviter non seulement que la densité du solvant au voisinage du soluté ne varie, mais aussi que la zone de solvant ne soit au contact direct du vide (effets de bord), des conditions périodiques aux limites sont imposées (Figure 43). Le système soluté-solvant constitue la boîte centrale qui est répliquée sous forme d'images dans toutes les directions de l'espace pour un total de  $[(3 \times 9) - 1] = 26$  images.

La fonction d'énergie potentielle n'est calculée que pour le système central et ses interactions éventuelles avec les atomes des boîtes images. Les éventuels mouvements de chacun des atomes sont ensuite répercutés dans toutes les images. Ainsi, si une molécule de solvant sort du système principal, par le jeu des conditions périodiques aux limites, l'image de cette même molécule est réintroduite dans la boîte primaire ce qui permet d'avoir toujours une densité constante du solvant.

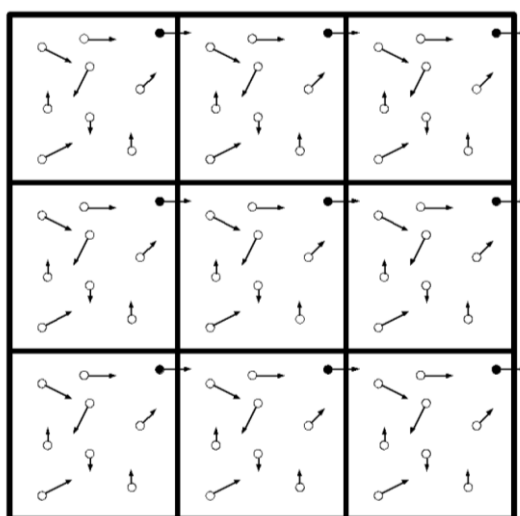


Figure 43 : Principe des conditions périodiques aux limites.

### 2.3.2 Protocole de simulation

Toutes les simulations ont été réalisées dans l'ensemble isotherme-isobare (NPT : 298 K, 1 atmosphère).

La protéine est immergée dans une boîte d'eau étendue d'au moins 13 Å dans les trois directions de l'espace par rapport aux dimensions minimales de la protéine.

Le système est ensuite soumis à un cycle de 5000 pas de minimisation et de 25000 pas de simulation afin de relâcher les contraintes de la protéine, puis à 250000 pas d'équilibre en

réinitialisant les vitesses à 298 K avec un coefficient de friction de  $5 \text{ ps}^{-1}$ . Le système est enfin équilibré pendant 250000 pas à 298 K en diminuant la constante de friction à  $1 \text{ ps}^{-1}$ . La production proprement dite consiste à laisser le système évoluer à 298 K, 1 atmosphère sur une durée totale de 150 ns. Les coordonnées, les vitesses et les énergies sont sauvegardées toutes les 5 ns. A la fin des 150 ns de simulation, 30000 structures sont générées.

## 2.4 Analyse des modes normaux

L'analyse des modes normaux consiste à considérer l'ensemble des fluctuations atomiques d'une molécule comme la superposition de modes de vibration harmoniques indépendants. Chacun de ces modes de vibration est caractérisé par une direction de déplacement (appelé mode normal, car les modes de vibration obtenus par cette méthode forment une base orthogonale de l'espace) et une fréquence.

Si l'on note  $x_i$  les  $3N$  coordonnées des atomes d'une molécule, et  $x_i^0$  les coordonnées de référence, l'énergie potentielle  $V$  peut être décomposée en série de Taylor :

$$V = V_0 + \sum_{i=1}^{3N} \left( \frac{\partial V}{\partial x_i} \right)_0 (x_i - x_i^0) + \frac{1}{2} \sum_{i,j=1}^{3N} \left( \frac{\partial^2 V}{\partial x_i \partial x_j} \right)_0 (x_i - x_i^0)(x_j - x_j^0) + \dots$$

$V$  étant définie à une constante additive près, on peut poser :

$$V_0 = 0$$

De plus, si la structure se trouve être à un minimum d'énergie, on a :

$$\forall i \in \{1, 2, \dots, N\}, \left( \frac{\partial V}{\partial x_i} \right)_0 = 0$$

Dans le cadre de petits déplacements autour de la position d'équilibre, on peut négliger les termes d'ordre supérieur ou égal à 3. Au final, on obtient  $V$  sous la forme d'une somme de potentiels harmoniques :

$$V \approx \frac{1}{2} \sum_{i,j=1}^{3N} \left( \frac{\partial^2 V}{\partial x_i \partial x_j} \right)_0 (x_i - x_i^0)(x_j - x_j^0)$$

L'équation du mouvement s'écrit donc, sous forme matricielle :

$$M\ddot{X} = -K(X - X_0)$$

$M$  est la matrice ( $3N \times 3N$ ) diagonale contenant les masses des atomes,  $K$  est la matrice des dérivées secondes du potentiel dont les éléments sont donnés par  $K_{ij} = \frac{\partial^2 V}{\partial x_i \partial x_j}$ ,  $X$  est le vecteur des coordonnées  $x_i$  et  $X_0$  est ce même vecteur dans la structure de référence.

En utilisant des coordonnées pondérées par les masses, cette équation devient :

$$\ddot{\tilde{X}} = -\tilde{K}(\tilde{X} - \tilde{X}_0)$$

avec :  $\tilde{X} = \sqrt{M}X$  ,  $\tilde{X}_0 = \sqrt{M}X_0$  ,  $\tilde{K} = \sqrt{M}^{-1}K\sqrt{M}^{-1}$

Cette équation classique d'oscillateurs harmoniques à  $3N$  dimensions se résout en recherchant des solutions de la forme :

$$x_i(t) = \sum_{k=1}^{3N} C_k a_{ik} \cos(\omega_k t + \phi_k)$$

c'est-à-dire comme une superposition de  $3N$  modes normaux de vibration. Chaque mode normal de vibration est défini par :

- une direction de déplacement donnée par le vecteur  $A_k = \{a_{ik}\}_{i=1,3N}$ .
- une fréquence  $\omega_k$ .

Ces grandeurs sont obtenues en diagonalisant la matrice  $\tilde{K}$  : les vecteurs  $A_k$  sont les vecteurs propres de cette matrice et les valeurs propres associées sont égales au carré des fréquences. Il est à noter qu'il existe 6 modes de fréquence nulle (5 pour une molécule linéaire) qui correspondent aux mouvements d'ensemble de la protéine (3 translations et 2 ou 3 rotations).

## 3 Fichiers de contraintes RMN

### 3.1 Contraintes de distances

Le fichier suivant présente les 129 contraintes de distances inter-protons (en Å) issues du spectre NOESY du peptide PP13 utilisé pour la modélisation sous contraintes RMN. Les protons géminés ambigus sont désignés par un « # ». La valeur en commentaire (après le « ! ») indique la valeur mesurée sur le modèle du peptide généré par la prédiction *de novo*.

```

assign (resid 5 and name Ha) (resid 7 and name HN) 5.8 2.0 2.0 !4.83
assign (resid 4 and name HN) (resid 5 and name Ha) 5.5 2.0 2.0 !5.50
assign (resid 10 and name HN) (resid 9 and name Hd#) 5.3 2.0 2.0 !5.02
assign (resid 3 and name Hb) (resid 5 and name HN) 5.2 2.0 2.0 !4.93
assign (resid 7 and name HN) (resid 8 and name Hb2) 5.2 2.0 2.0 !5.18
assign (resid 11 and name HN) (resid 12 and name Hb1) 5.1 2.0 2.0 !4.32
assign (resid 3 and name Hg#) (resid 4 and name HN) 5.0 2.0 2.0 !4.21
assign (resid 10 and name Hb1) (resid 8 and name HN) 5.0 2.0 2.0 !4.99
assign (resid 10 and name HN) (resid 11 and name Hb2) 5.0 2.0 2.0 !4.95
assign (resid 12 and name HN) (resid 13 and name Hg#) 5.0 2.0 2.0 !4.36
assign (resid 2 and name Hg#) (resid 3 and name HN) 4.9 2.0 2.0 !4.48
assign (resid 2 and name HN) (resid 5 and name Hb1) 4.9 2.0 2.0 !5.07
assign (resid 4 and name Ha) (resid 8 and name HN) 4.9 2.0 2.0 !3.97
assign (resid 5 and name Hb2) (resid 8 and name HN) 4.9 2.0 2.0 !5.63
assign (resid 2 and name HN) (resid 5 and name Hb2) 4.8 2.0 2.0 !4.34
assign (resid 7 and name Ha) (resid 9 and name HN) 4.8 2.0 2.0 !4.85
assign (resid 8 and name He#) (resid 8 and name Hb1) 4.8 2.0 2.0 !4.65
assign (resid 2 and name HN) (resid 2 and name Hb) 4.7 2.0 2.0 !3.99
assign (resid 5 and name Hb1) (resid 6 and name HN) 4.6 2.0 2.0 !3.8
assign (resid 11 and name Hb1) (resid 12 and name HN) 4.5 2.0 2.0 !2.72
assign (resid 11 and name Hb2) (resid 12 and name HN) 4.5 2.0 2.0 !3.95
assign (resid 2 and name Hb) (resid 4 and name HN) 4.4 2.0 2.0 !5.38
assign (resid 6 and name Hg#) (resid 7 and name HN) 4.4 2.0 2.0 !4.19
assign (resid 10 and name HN) (resid 8 and name Hb2) 4.4 2.0 2.0 !4.89
assign (resid 4 and name Hb2) (resid 5 and name HN) 4.3 2.0 2.0 !3.57
assign (resid 10 and name Hb1) (resid 7 and name HN) 4.3 2.0 2.0 !4.65
assign (resid 12 and name HN) (resid 12 and name Hb2) 4.3 2.0 2.0 !3.66
assign (resid 3 and name Hb) (resid 7 and name HN) 4.2 2.0 2.0 !5.55
assign (resid 4 and name HN) (resid 6 and name Hb#) 4.2 2.0 2.0 !5.65
assign (resid 3 and name HN) (resid 3 and name Hg#) 4.1 2.0 2.0 !4.23
assign (resid 5 and name Ha) (resid 8 and name Hd#) 4.1 2.0 2.0 !4.49
assign (resid 6 and name Ha) (resid 9 and name HN) 4.1 2.0 2.0 !3.67
assign (resid 7 and name Hd#) (resid 7 and name Ha) 4.1 2.0 2.0 !2.61
assign (resid 10 and name HN) (resid 7 and name Hb2) 4.1 2.0 2.0 !5.67
assign (resid 11 and name HN) (resid 11 and name Hb2) 4.1 2.0 2.0 !4.76
assign (resid 1 and name Hd1) (resid 1 and name Hb2) 4.0 2.0 2.0 !3.95
assign (resid 8 and name Hb1) (resid 9 and name HN) 4.0 2.0 2.0 !3.70
assign (resid 4 and name Ha) (resid 7 and name HN) 3.9 2.0 2.0 !3.74
assign (resid 7 and name Hd#) (resid 7 and name Hb1) 3.9 2.0 2.0 !2.28
assign (resid 7 and name Ha) (resid 8 and name HN) 3.9 2.0 2.0 !3.6
assign (resid 10 and name HN) (resid 10 and name Hb2) 3.9 2.0 2.0 !3.56
assign (resid 13 and name Hd#) (resid 13 and name Hb2) 3.9 2.0 2.0 !3.90
assign (resid 4 and name Ha) (resid 5 and name HN) 3.8 2.0 2.0 !3.63
assign (resid 6 and name Ha) (resid 6 and name Hd#) 3.8 2.0 2.0 !4.41
assign (resid 7 and name Hd#) (resid 7 and name Hb2) 3.8 2.0 2.0 !2.28
assign (resid 9 and name HN) (resid 9 and name Hg#) 3.8 2.0 2.0 !4.26
assign (resid 11 and name HN) (resid 11 and name Hg#) 3.8 2.0 2.0 !4.42
assign (resid 3 and name Ha) (resid 5 and name HN) 3.7 2.0 2.0 !4.90
assign (resid 5 and name HN) (resid 5 and name Hb1) 3.7 2.0 2.0 !3.57
assign (resid 6 and name HN) (resid 6 and name Hg#) 3.7 2.0 2.0 !4.31
assign (resid 9 and name Ha) (resid 9 and name Hd#) 3.7 2.0 2.0 !4.44
assign (resid 10 and name Ha) (resid 11 and name HN) 3.7 2.0 2.0 !3.49
assign (resid 2 and name Ha) (resid 3 and name HN) 3.6 2.0 2.0 !3.58
assign (resid 4 and name Hd#) (resid 7 and name Hb2) 3.6 2.0 2.0 !5.42
assign (resid 11 and name HN) (resid 8 and name Ha) 3.6 2.0 2.0 !4.02
assign (resid 11 and name Ha) (resid 12 and name HN) 3.6 2.0 2.0 !3.55
assign (resid 3 and name Ha) (resid 6 and name HN) 3.5 2.0 2.0 !3.99

```

## Annexes

assign	(resid 5 and name Ha)	(resid 6 and name HN)	3.5	2.0	2.0	!3.60
assign	(resid 7 and name HN)	(resid 7 and name Hb2)	3.5	2.0	2.0	!3.51
assign	(resid 8 and name Ha)	(resid 9 and name HN)	3.5	2.0	2.0	!3.63
assign	(resid 1 and name Ha)	(resid 2 and name HN)	3.4	1.2	1.2	!3.39
assign	(resid 2 and name HN)	(resid 2 and name Hg#)	3.4	1.2	1.2	!3.44
assign	(resid 6 and name Ha)	(resid 7 and name HN)	3.4	1.2	1.2	!3.66
assign	(resid 7 and name Hb2)	(resid 8 and name HN)	3.4	1.2	1.2	!3.40
assign	(resid 8 and name HN)	(resid 8 and name Hb2)	3.4	1.2	1.2	!2.59
assign	(resid 4 and name Ha)	(resid 4 and name Hb1)	3.3	1.2	1.2	!3.04
assign	(resid 7 and name Ha)	(resid 7 and name Hb1)	3.3	1.2	1.2	!2.98
assign	(resid 7 and name HN)	(resid 8 and name HN)	3.3	1.2	1.2	!3.06
assign	(resid 9 and name HN)	(resid 9 and name Ha)	3.3	1.2	1.2	!2.82
assign	(resid 11 and name Ha)	(resid 11 and name Hb2)	3.3	1.2	1.2	!3.04
assign	(resid 12 and name HN)	(resid 9 and name Ha)	3.3	1.2	1.2	!3.31
assign	(resid 13 and name Hd#)	(resid 13 and name Hb1)	3.3	1.2	1.2	!3.31
assign	(resid 3 and name HN)	(resid 4 and name HN)	3.2	1.2	1.2	!3.24
assign	(resid 4 and name HN)	(resid 5 and name HN)	3.1	1.2	1.2	!2.98
assign	(resid 8 and name Ha)	(resid 8 and name Hb2)	3.1	1.2	1.2	!3.03
assign	(resid 1 and name Hd1)	(resid 1 and name Hb1)	3.0	1.2	1.2	!2.98
assign	(resid 1 and name Ha)	(resid 1 and name Hb1)	3.0	1.2	1.2	!3.04
assign	(resid 5 and name Ha)	(resid 5 and name Hb2)	3.0	1.2	1.2	!2.38
assign	(resid 11 and name Ha)	(resid 11 and name Hb2)	3.0	1.2	1.2	!3.04
assign	(resid 12 and name HN)	(resid 12 and name Ha)	3.0	1.2	1.2	!2.96
assign	(resid 3 and name HN)	(resid 3 and name Hb)	2.9	0.8	0.8	!2.88
assign	(resid 5 and name Ha)	(resid 5 and name Hb1)	2.9	0.8	0.8	!2.51
assign	(resid 5 and name HN)	(resid 5 and name Ha)	2.9	0.8	0.8	!2.84
assign	(resid 6 and name Ha)	(resid 6 and name Hb#)	2.9	0.8	0.8	!2.65
assign	(resid 10 and name HN)	(resid 11 and name HN)	2.9	0.8	0.8	!2.88
assign	(resid 2 and name Ha)	(resid 2 and name Hg#)	2.8	0.8	0.8	!2.74
assign	(resid 4 and name Hd#)	(resid 4 and name Ha)	2.8	0.8	0.8	!2.83
assign	(resid 4 and name HN)	(resid 4 and name Ha)	2.8	0.8	0.8	!2.77
assign	(resid 1 and name Hb2)	(resid 2 and name HN)	2.7	0.8	0.8	!2.77
assign	(resid 3 and name Hb)	(resid 3 and name Ha)	2.7	0.8	0.8	!3.02
assign	(resid 3 and name HN)	(resid 3 and name Ha)	2.7	0.8	0.8	!2.65
assign	(resid 6 and name Hb#)	(resid 6 and name Hg#)	2.7	0.8	0.8	!2.48
assign	(resid 7 and name Ha)	(resid 7 and name Hb2)	2.7	0.8	0.8	!2.66
assign	(resid 7 and name HN)	(resid 7 and name Ha)	2.7	0.8	0.8	!2.73
assign	(resid 9 and name He1)	(resid 9 and name Hd#)	2.7	0.8	0.8	!2.66
assign	(resid 10 and name Ha)	(resid 10 and name Hb2)	2.7	0.8	0.8	!2.49
assign	(resid 10 and name HN)	(resid 10 and name Ha)	2.7	0.8	0.8	!2.73
assign	(resid 11 and name Ha)	(resid 11 and name Hg#)	2.7	0.8	0.8	!2.74
assign	(resid 1 and name Hb1)	(resid 1 and name Hb2)	2.6	0.8	0.8	!1.8
assign	(resid 2 and name Ha)	(resid 2 and name Hb)	2.6	0.8	0.8	!2.56
assign	(resid 2 and name Hb)	(resid 2 and name Hg#)	2.5	0.8	0.8	!2.45
assign	(resid 2 and name Hb)	(resid 3 and name HN)	2.5	0.8	0.8	!2.47
assign	(resid 3 and name Hb)	(resid 3 and name Hg#)	2.5	0.8	0.8	!2.48
assign	(resid 4 and name Ha)	(resid 4 and name Hb2)	2.5	0.8	0.8	!2.50
assign	(resid 4 and name Hd#)	(resid 4 and name He#)	2.5	0.8	0.8	!2.48
assign	(resid 6 and name Ha)	(resid 6 and name Hg#)	2.5	0.8	0.8	!2.44
assign	(resid 8 and name Ha)	(resid 8 and name Hb1)	2.5	0.8	0.8	!2.48
assign	(resid 11 and name Hd#)	(resid 11 and name Hg#)	2.5	0.8	0.8	!2.48
assign	(resid 13 and name Hb1)	(resid 13 and name Hb2)	2.5	0.8	0.8	!1.8
assign	(resid 6 and name Ha)	(resid 6 and name Hg#)	2.5	0.8	0.8	!2.44
assign	(resid 6 and name Hd#)	(resid 6 and name Hg#)	2.4	0.8	0.8	!2.46
assign	(resid 6 and name HN)	(resid 6 and name Hb#)	2.4	0.8	0.8	!2.40
assign	(resid 7 and name He#)	(resid 7 and name Hd#)	2.4	0.8	0.8	!2.48
assign	(resid 9 and name Hb#)	(resid 9 and name Hg#)	2.4	0.8	0.8	!2.46
assign	(resid 12 and name Ha)	(resid 12 and name Hb2)	2.4	0.8	0.8	!2.43
assign	(resid 12 and name Hb1)	(resid 12 and name Hb2)	2.4	0.8	0.8	!1.8
assign	(resid 6 and name He#)	(resid 6 and name Hd#)	2.3	0.8	0.8	!2.32
assign	(resid 8 and name Hb1)	(resid 8 and name Hb2)	2.3	0.8	0.8	!1.8
assign	(resid 8 and name Hd#)	(resid 8 and name Hb1)	2.3	0.8	0.8	!2.30
assign	(resid 8 and name Hd#)	(resid 8 and name Hb2)	2.3	0.8	0.8	!2.30
assign	(resid 13 and name Ha)	(resid 13 and name Hb1)	2.3	0.8	0.8	!2.32
assign	(resid 13 and name Ha)	(resid 13 and name Hb2)	2.3	0.8	0.8	!2.32
assign	(resid 2 and name HN)	(resid 2 and name Ha)	2.2	0.8	0.8	!2.15
assign	(resid 5 and name Hb1)	(resid 5 and name Hb2)	2.2	0.8	0.8	!1.8
assign	(resid 9 and name Hb#)	(resid 9 and name Hd#)	2.2	0.8	0.8	!2.57
assign	(resid 1 and name Hb1)	(resid 2 and name HN)	2.1	0.8	0.8	!2.05
assign	(resid 9 and name Hd#)	(resid 9 and name Hg#)	2.1	0.8	0.8	!2.47
assign	(resid 12 and name Ha)	(resid 13 and name Hd#)	2.1	0.8	0.8	!2.12
assign	(resid 11 and name HN)	(resid 12 and name HN)	2.0	0.8	0.8	!2.03



### 3.2 Contraintes d'angles dièdres

Le fichier suivant présente les 11 contraintes d'angles dièdres  $\phi$  sous la forme des valeurs de constantes de couplage  $^3J_{\text{HN-H}\alpha}$  (en Hz) issues du spectre 1D  $^1\text{H}$  du peptide PP13 utilisé pour la modélisation sous contraintes RMN. La conversion est effectuée directement par CNS en utilisant la relation de Karplus avec les coefficients de Pardi.

```
!2T
assign (resid 1 and name c ) (resid 2 and name n)
      (resid 2 and name ca) (resid 2 and name c) 8.2 0.5
!3T
assign (resid 2 and name c ) (resid 3 and name n)
      (resid 3 and name ca) (resid 3 and name c) 5.1 0.5
!4Y
assign (resid 3 and name c ) (resid 4 and name n)
      (resid 4 and name ca) (resid 4 and name c) 5.2 0.5
!5S
assign (resid 4 and name c ) (resid 5 and name n)
      (resid 5 and name ca) (resid 5 and name c) 5.3 0.5
!6K
assign (resid 5 and name c ) (resid 6 and name n)
      (resid 6 and name ca) (resid 6 and name c) 4.9 0.5
!7F
assign (resid 6 and name c ) (resid 7 and name n)
      (resid 7 and name ca) (resid 7 and name c) 5.7 0.5
!8F
assign (resid 7 and name c ) (resid 8 and name n)
      (resid 8 and name ca) (resid 8 and name c) 6.2 0.5
!9K
assign (resid 8 and name c ) (resid 9 and name n)
      (resid 9 and name ca) (resid 9 and name c) 6.4 0.5
!10D
assign (resid 9 and name c ) (resid 10 and name n)
      (resid 10 and name ca) (resid 10 and name c) 6.4 0.5
!11R
assign (resid 10 and name c ) (resid 11 and name n)
      (resid 11 and name ca) (resid 11 and name c) 8.3 0.5
!12C
assign (resid 11 and name c ) (resid 12 and name n)
      (resid 12 and name ca) (resid 12 and name c) 8.7 0.5
```



## Bibliographie



1. Hsu, J.C., et al., *Heat-unstable proteins in grape juice and wine. II. Characterization and removal by ultrafiltration*. American Journal of Enology and Viticulture, 1987. **38**(1): p. 17-22.
2. Brissonnet, F. and A. Maujean, *Characterization of foaming proteins in a champagne base wine*. American Journal of Enology and Viticulture, 1993. **44**(3): p. 297-301.
3. Lamikanra, O. and I.D. Inyang, *Temperature influence on muscadine wine protein characteristics*. American Journal of Enology and Viticulture, 1988. **39**(2): p. 113-116.
4. Robinson, S.P. and C. Davies, *Molecular biology of grape berry ripening*. Australian Journal of Grape and Wine Research, 2000. **6**(2): p. 175-188.
5. Jégou, S., et al., *One step purification of the grape vacuolar invertase*. Analytica Chimica Acta, 2009. **638**(1): p. 75-78.
6. Lee, T.H., K.F. Pocock, and E.J. Waters, *The effect of mechanical harvesting and transport of grapes, and juice oxidation, on the protein stability of wines*. Physical Stability of Wine, 1986. **4**(3): p. 23-39.
7. Pocock, K.F. and E.J. Waters, *The effect of mechanical harvesting and transport of grapes, and juice oxidation, on the protein stability of wines*. Australian Journal of Grape and Wine Research, 1998. **4**(3): p. 136-139.
8. Hsu, J.C. and D.A. Heatherbell, *Isolation and characterization of soluble proteins in grapes, grape juice, and wine*. American Journal of Enology and Viticulture, 1987. **38**(1): p. 6-10.
9. Bayly, F.C. and H.W. Berg, *Grape and wine proteins of white wine varieties*. American Journal of Enology and Viticulture, 1967. **18**(1): p. 18-32.
10. Somers, T.C. and G. Ziemelis, *Direct determination of wine proteins*. American Journal of Enology and Viticulture, 1973. **24**: p. 47-50.
11. Murphey, J.M., S.E. Spayd, and J.R. Powers, *Effect of grape maturation on soluble protein characteristics of Gewurztraminer and White Riesling juice and wine*. American Journal of Enology and Viticulture, 1989. **40**(3): p. 199-207.
12. Dorrestein, E., et al., *Electrophoretic and FPLC analysis of soluble proteins in four Portuguese wines*. American Journal of Enology and Viticulture, 1995. **46**(2): p. 235-242.
13. Santoro, M., *Fractionation and characterization of must and wine proteins*. American Journal of Enology and Viticulture, 1995. **46**(2): p. 250-254.
14. Fukui, M., et al., *Characteristics of proteins in juice and wine produced from Semillon grapes*. Journal Brewing Society of Japan, 2003. **98**: p. 582-588.
15. Paetzold M., L. Dulau, and D. Dubourdiou, *Fractionnement et caractérisation des glycoprotéines dans les moûts de raisins blancs*. Journal International des Sciences de la Vigne et du Vin, 1990. **24**: p. 13-28.
16. Dambrouck, T., et al., *Immunodetection of proteins from grapes and yeast in a white wine*. Journal of Agricultural and Food Chemistry, 2003. **51**(9): p. 2727-2732.
17. Marchal, R., S. Bouquelet, and A. Maujean, *Purification and partial biochemical characterization of glycoproteins in a champenois chardonnay wine*. Journal of Agricultural and Food Chemistry, 1996. **44**(7): p. 1716-1722.
18. Feuillat, M., G. Brilliant, and J. Rochard, *Mise en évidence d'une production de protéases exocellulaires par les levures au cours de la fermentation alcoolique du moût de raisin*. Conn Vigne Vin, 1980. **14**: p. 37-52.
19. Cilindre, C., et al., *Influence of Botrytis cinerea infection on Champagne wine proteins (characterized by two-dimensional electrophoresis/immunodetection) and wine foaming properties*. Food Chemistry, 2007. **103**(1): p. 139-149.
20. Cilindre, C., et al., *Proteomic approach to identify champagne wine proteins as modified by Botrytis cinerea infection*. Journal of Proteome Research, 2008. **7**(3): p. 1199-1208.
21. Antoniw, J.F., et al., *Comparison of three pathogenesis-related proteins from plants of two cultivars of tobacco infected with TMV*. Journal of General Virology, 1980. **47**(1): p. 79-87.
22. van Loon, L.C., *The nomenclature of pathogenesis-related proteins*. Physiological and Molecular Plant Pathology, 1990. **37**(3): p. 229-230.

23. Derckel, J.P., et al., *Chitinases of the grapevine (Vitis vinifera L.): Five isoforms induced in leaves by salicylic acid are constitutively expressed in other tissues*. Plant Science, 1996. **119**(1-2): p. 31-37.
24. Robinson, S.P., A.K. Jacobs, and I.B. Dry, *A class IV chitinase is highly expressed in grape berries during ripening*. Plant Physiology, 1997. **114**(3): p. 771-778.
25. Giribaldi, M. and M.G. Giuffrida, *Heard it through the grapevine: Proteomic perspective on grape and wine*. Journal of Proteomics. **73**(9): p. 1647-1655.
26. Tattersall, D.B., R. Van Heeswijck, and P.B. Hoj, *Identification and characterization of a fruit-specific, thaumatin-like protein that accumulates at very high levels in conjunction with the onset of sugar accumulation and berry softening in grapes*. Plant Physiology, 1997. **114**(3): p. 759-769.
27. Pocock, K.F., et al., *Thaumatin-like proteins and chitinases, the haze-forming proteins of wine, accumulate during ripening of grape (Vitis vinifera) berries and drought stress does not affect the final levels per berry at maturity*. Journal of Agricultural and Food Chemistry, 2000. **48**(5): p. 1637-1643.
28. Monteiro, S., et al., *The diversity of pathogenesis-related proteins decreases during grape maturation*. Phytochemistry, 2007. **68**(4): p. 416-425.
29. van Loon, L.C., M. Rep, and C.M.J. Pieterse, *Significance of Inducible Defense-related Proteins in Infected Plants*. Annual Review of Phytopathology, 2006. **44**(1): p. 135-162.
30. Christensen, A.B., et al., *The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins*. Molecular Plant Pathology, 2002. **3**(3): p. 135-144.
31. Peng, Z., et al., *Taste Properties of Grape (Vitis vinifera) Pathogenesis-Related Proteins Isolated from Wine*. Journal of Agricultural and Food Chemistry, 1997. **45**(12): p. 4639-4643.
32. Waters, E.J., et al., *Sequence Analysis of Grape (Vitis vinifera) Berry Chitinases That Cause Haze Formation in Wines*. Journal of Agricultural and Food Chemistry, 1998. **46**(12): p. 4950-4957.
33. Flach, J., P.E. Pilet, and P. Jolles, *What's new in chitinase research?* Experientia, 1992. **48**(8): p. 701-716.
34. Arakane, Y. and S. Muthukrishnan, *Insect chitinase and chitinase-like proteins*. Cellular and Molecular Life Sciences. **67**(2): p. 201-216.
35. Sharma, N., et al., *Role of chitinase in plant defense*. Asian Journal of Biochemistry. **6**(1): p. 29-37.
36. Kasprzewska, A., *Plant chitinases - Regulation and function*. Cellular and Molecular Biology Letters, 2003. **8**(3): p. 809-824.
37. Ferreira, R.B., et al., *The role of plant defence proteins in fungal pathogenesis*. Molecular Plant Pathology, 2007. **8**(5): p. 677-700.
38. Pocock, K.F., et al., *Sulfate - A candidate for the missing essential factor that is required for the formation of protein haze in white wine*. Journal of Agricultural and Food Chemistry, 2007. **55**(5): p. 1799-1807.
39. Ferreira, R.B., et al., *The wine proteins*. Trends in Food Science and Technology, 2001. **12**(7): p. 230-239.
40. Blade, W.H. and R. Boulton, *Adsorption of protein by bentonite in a model wine solution*. American Journal of Enology and Viticulture, 1988. **39**(3): p. 193-199.
41. Hoj, P.B., et al., *The "haze-proteins" of wine - A summary of properties, factors affecting their accumulation in grapes, and the amount of bentonite required for their removal from wine*. American Journal of Enology and Viticulture, 2001. **51**: p. 147-154.
42. Desportes, C., et al., *Isolation, identification, and organoleptic characterization of low-molecular-weight Peptides from white wine*. American Journal of Enology and Viticulture, 2001. **52**(4): p. 376-380.

43. Lubbers, S., et al., *Influence of Mannaproteins from Yeast on the Aroma Intensity of a Model Wine*. LWT - Food Science and Technology, 1994. **27**(2): p. 108-114.
44. Jones, P.R., et al., *The influence of interactions between major white wine components on the aroma, flavour and texture of model white wine*. Food Quality and Preference, 2008. **19**(6): p. 596-607.
45. Maujean, A., et al., *Etude de la teneur et de la qualité de mousse des vins effervescents. II. Mise au point d'une technique de mesure de la moussabilité, de la tenue et de la stabilité de la mousse des vins effervescents*. Bulletin de l'O.I.V., 1990: p. 711-712.
46. Girbau-Sola, T., et al., *Foam aptitude of Trepas and Monastrell red varieties in Cava elaboration. 1. Base wine characteristics*. Journal of Agricultural and Food Chemistry, 2002. **50**(20): p. 5596-5599.
47. Liger-Belair, G., G. Polidori, and P. Jeandet, *Recent advances in the science of champagne bubbles*. Chemical Society Reviews, 2008. **37**(11): p. 2490-2511.
48. Senée, J., B. Robillard, and M. Vignes-Adler, *Films and foams of Champagne wines*. Food Hydrocolloids, 1999. **13**(1): p. 15-26.
49. Péron, N., et al., *Layers of macromolecules at the champagne/air interface and the stability of champagne bubbles*. Langmuir, 2001. **17**(3): p. 791-797.
50. Péron, N., et al., *Characterization by drop tensiometry and by ellipsometry of the adsorption layer formed at the air/champagne wine interface*. Advances in Colloid and Interface Science, 2000. **88**(1-2): p. 19-36.
51. Le Bourse, D., et al., *Review of preparative and analytical procedures for the study of proteins in grape juice and wine*. Analytica Chimica Acta. **667**(1-2): p. 33-42.
52. Le Bourse, D., et al., *Quantification of chitinase and thaumatin-like proteins in grape juices and wines*. Analytical and Bioanalytical Chemistry. **401**(5): p. 1545-1553.
53. Van Sluyter, S.C., et al., *Two-Step Purification of Pathogenesis-Related Proteins from Grape Juice and Crystallization of Thaumatin-like Proteins*. Journal of Agricultural and Food Chemistry, 2009. **57**(23): p. 11376-11382.
54. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature, 1970. **227**(5259): p. 680-685.
55. Neuhoff, V., R. Stamm, and H. Eibl, *Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: A systematic analysis*. Electrophoresis, 1985. **6**(9): p. 427-448.
56. Rabilloud, T., et al., *Silver-staining of proteins in polyacrylamide gels: a general overview*. Cellular and Molecular Biology, 1994. **40**(1): p. 57-75.
57. Towbin, H., T. Staehelin, and J. Gordon, *Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications*. Proceedings of the National Academy of Sciences of the United States of America, 1979. **76**(9): p. 4350-4354.
58. Dambrouck, T., et al., *Determination of the grape invertase content (Using PTA-ELISA) following various fining treatments versus changes in the total protein content of wine. Relationships with wine foamability*. Journal of Agricultural and Food Chemistry, 2005. **53**(22): p. 8782-8789.
59. Adamson, A.W. and A.P. Gast, Physical Chemistry of Surfaces. 6th ed., 1997.
60. Wilhelmy, L., *Ueber die Abhängigkeit der Capillaritäts-Constanten des Alkohols von Substanz und Gestalt des benetzten festen Körpers*. Annals of Physics, 1863. **119**(6): p. 177-217.
61. King, T.B., *The surface tension and structure of silicate slags*. Journal of the Society of Glass Technology, 1951. **35**(166): p. 241-259.
62. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations*. Journal of Molecular Biology, 1963. **7**: p. 95-99.
63. van der Wel, H. and K. Loeve, *Isolation and Characterization of Thaumatin I and II, the Sweet-Tasting Proteins from Thaumatooccus daniellii Benth*. European Journal of Biochemistry, 1972. **31**(2): p. 221-225.

64. Shatters Jr, R.G., et al., *Phylogenetic and structural relationships of the PR5 gene family reveal an ancient multigene family conserved in plants and select animal taxa*. Journal of Molecular Evolution, 2006. **63**(1): p. 12-29.
65. Brandazza, A., et al., *Plant stress proteins of the thaumatin-like family discovered in animals*. FEBS Letters, 2004. **572**(1-3): p. 3-7.
66. Grenier, J., C. Potvin, and A. Asselin, *Some fungi express  $\beta$ -1,3-glucanases similar to thaumatin-like proteins*. Mycologia, 2000. **92**(5): p. 841-848.
67. Sakamoto, Y., et al., *Lentinula edodes tlg1 encodes a thaumatin-like protein that is involved in lentinan degradation and fruiting body senescence*. Plant Physiology, 2006. **141**(2): p. 793-801.
68. Jami, S.K., et al., *Molecular, biochemical and structural characterization of osmotin-like protein from black nightshade (*Solanum nigrum*)*. Journal of Plant Physiology, 2007. **164**(3): p. 238-252.
69. Tachi, H., et al., *Molecular characterization of a novel soybean gene encoding a neutral PR-5 protein induced by high-salt stress*. Plant Physiology and Biochemistry, 2009. **47**(1): p. 73-79.
70. Falconer, R.J., et al., *Thermal stability of thaumatin-like protein, chitinase, and invertase isolated from sauvignon blanc and semillon juice and their role in haze formation in wine*. Journal of Agricultural and Food Chemistry, 2010. **58**(2): p. 975-980.
71. Waters, E.J., et al., *Preventing protein haze in bottled white wine*. Australian Journal of Grape and Wine Research, 2005. **11**(2): p. 215-225.
72. Fierens, E., et al., *Biochemical and structural characterization of TLXI, the *Triticum aestivum* L. thaumatin-like xylanase inhibitor*. Journal of Enzyme Inhibition and Medicinal Chemistry, 2009. **24**(3): p. 646-654.
73. Smole, U., et al., *Mal d 2, the thaumatin-like allergen from apple, is highly resistant to gastrointestinal digestion and thermal processing*. International Archives of Allergy and Immunology, 2008. **147**(4): p. 289-98.
74. Ogata, C.M., et al., *Crystal structure of a sweet tasting protein thaumatin I, at 1.65 Å resolution*. Journal of Molecular Biology, 1992. **228**(3): p. 893-908.
75. Masuda, T., et al., *Crystal structure of the sweet-tasting protein thaumatin II at 1.27 Å*. Biochemical and Biophysical Research Communications. **410**(3): p. 457-460.
76. Batalia, M.A., et al., *The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like, PR-5 protein family*. Nature Structural and Molecular Biology, 1996. **3**(1): p. 19-22.
77. Koiwa, H., et al., *Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins*. Journal of Molecular Biology, 1999. **286**(4): p. 1137-45.
78. Min, K., et al., *Crystal Structure of Osmotin, a Plant Antifungal Protein*. Proteins: Structure, Function and Genetics, 2004. **54**(1): p. 170-173.
79. Dall'Antonia, Y., et al., *Crystallization and preliminary structure determination of the plant food allergen Pru av 2*. Acta Crystallography, 2005. **61**(Pt 2): p. 186-198.
80. Leone, P., et al., *Resolution of the structure of the allergenic and antifungal banana fruit thaumatin-like protein at 1.7-Å*. Biochimie, 2006. **88**(1): p. 45-52.
81. Ghosh, R. and C. Chakrabarti, *Crystal structure analysis of NP24-I: a thaumatin-like protein*. Planta, 2008. **228**(5): p. 883-890.
82. Velazhahan, R., S.K. Datta, and S. Muthukrishnan, *The PR-5 family: Thaumatin-like proteins*. Pathogenesis-Related Proteins in Plants, 1999: p. 107-129.
83. Liu, J.J., R. Sturrock, and A.K.M. Ekramoddoullah, *The superfamily of thaumatin-like proteins: Its origin, evolution, and expression towards biological function*. Plant Cell Reports. **29**(5): p. 419-436.
84. Petre, B., et al., *Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar*. BMC Plant Biology, 2011. **11**: p. 33.



85. Abad, L.R., et al., *Antifungal activity of tobacco osmotin has specificity and involves plasma membrane permeabilization*. Plant Science, 1996. **118**(1): p. 11-23.
86. Yun, D.J., et al., *Osmotin, a plant antifungal protein, subverts signal transduction to enhance fungal cell susceptibility*. Molecular Cell, 1998. **1**(6): p. 807-817.
87. Vigers, A.J., W.K. Roberts, and C.P. Selitrennikoff, *A new family of plant antifungal proteins*. Molecular plant-microbe interactions, 1991. **4**(4): p. 315-323.
88. Fierens, E., et al., *TLXI, a novel type of xylanase inhibitor from wheat (Triticum aestivum) belonging to the thaumatin family*. Biochemical Journal, 2007. **403**(3): p. 583-591.
89. Schimoler-O'Rourke, R., M. Richardson, and C.P. Selitrennikoff, *Zeamatin Inhibits Trypsin and Amylase Activities*. Applied and Environmental Microbiology, 2001. **67**(5): p. 2365-2366.
90. Narasimhan, M.L., et al., *Osmotin is a homolog of mammalian adiponectin and controls apoptosis in yeast through a homolog of mammalian adiponectin receptor*. Molecular Cell, 2005. **17**(2): p. 171-180.
91. Kim, M.J., et al., *In vitro and in planta interaction evidence between Nicotiana tabacum thaumatin-like protein 1 (TLP1) and Cucumber mosaic virus proteins*. Plant Molecular Biology, 2005. **59**(6): p. 981-994.
92. Yu, X.M. and M. Griffith, *Antifreeze proteins in winter rye leaves form oligomeric complexes*. Plant Physiology, 1999. **119**(4): p. 1361-1369.
93. Griffith, M., et al., *Antifreeze proteins modify the freezing process in planta*. Plant Physiology, 2005. **138**(1): p. 330-340.
94. Barthakur, S., V. Babu, and K.C. Bansal, *Over-expression of osmotin induces proline accumulation and confers tolerance to osmotic stress in transgenic tobacco*. Journal of Plant Biochemistry and Biotechnology, 2001. **10**(1): p. 31-37.
95. Hulo, N., et al., *Development of transgenic strawberry (Fragaria x ananassa Duch.) plants tolerant to salt stress*. The 20 years of PROSITE. Nucleic Acids Research 36(database issue), 2008. **174**(4): p. D245-249.
96. Parkhi, V., et al., *Expression of apoplastically secreted tobacco osmotin in cotton confers drought tolerance*. Molecular Breeding, 2009. **23**(4): p. 625-639.
97. Takemoto, D., et al., *Identification of Chitinase and Osmotin-Like Protein as Actin-Binding Proteins in Suspension-Cultured Potato Cells*. Plant and Cell Physiology, 1997. **38**(4): p. 441-448.
98. Ibeas, J.I., et al., *Fungal cell wall phosphomannans facilitate the toxic activity of a plant PR-5 protein*. Plant Journal, 2000. **23**(3): p. 375-383.
99. Veronese, P., et al., *In defense against pathogens. Both plant sentinels and foot soldiers need to know the enemy*. Plant Physiology, 2003. **131**(4): p. 1580-1590.
100. Narasimhan, M.L., et al., *Overexpression of a cell wall glycoprotein in Fusarium oxysporum increases virulence and resistance to a plant PR-5 protein*. Plant Journal, 2003. **36**(3): p. 390-400.
101. Kobayashi, K., et al., *Cytokinin-binding proteins from tobacco callus share homology with osmotin-like protein and an endochitinase*. Plant and Cell Physiology, 2000. **41**(2): p. 148-157.
102. Pastorello, E.A., et al., *Identification of grape and wine allergens as an endochitinase 4, a lipid-transfer protein, and a thaumatin*. Journal of Allergy and Clinical Immunology, 2003. **111**(2): p. 350-359.
103. Vassilopoulou, E., et al., *Severe immediate allergic reactions to grapes: Part of a lipid transfer protein-associated clinical syndrome*. International Archives of Allergy and Immunology, 2007. **143**(2): p. 92-102.
104. Krebitz, M., et al., *Plant-based heterologous expression of Mal d 2, a thaumatin-like protein and allergen of apple (Malus domestica), and its characterization as an antifungal protein*. Journal of Molecular Biology, 2003. **329**(4): p. 721-730.
105. Marzban, G., et al., *Conformational changes of Mal d 2, a thaumatin-like apple allergen, induced by food processing*. Food Chemistry, 2009. **112**(4): p. 803-811.

106. Lehto, M., et al., *Thaumatococcus*-like protein and baker's respiratory allergy. *Annals of Allergy, Asthma and Immunology*, 2003. **104**(2): p. 139-146.
107. Jayasankar, S., Z. Li, and D.J. Gray, *Constitutive expression of Vitis vinifera thaumatococcus-like protein after in vitro selection and its role in anthracnose resistance*. *Functional Plant Biology*, 2003. **30**(11): p. 1105-1115.
108. Monteiro, S., et al., *Osmotin and Thaumatococcus from Grape: A Putative General Defense Mechanism Against Pathogenic Fungi*. *Phytopathology*, 2003. **93**(12): p. 1505-1512.
109. Grenier, J., et al., *Some thaumatococcus-like proteins hydrolyse polymeric  $\beta$ -1,3-glucans*. *Plant Journal*, 1999. **19**(4): p. 473-480.
110. Zareie, R., D.L. Melanson, and P.J. Murphy, *Isolation of fungal cell wall degrading proteins from barley (Hordeum vulgare L.) leaves infected with Rhynchosporium secalis*. *Molecular Plant-Microbe Interactions*, 2002. **15**(10): p. 1031-1039.
111. Osmond, R.I.W., et al., *Binding interactions between barley thaumatococcus-like proteins and (1,3)- $\beta$ -D-glucans*. *European Journal of Biochemistry*, 2001. **268**(15): p. 4190-4199.
112. Fagoaga, C., et al., *Increased tolerance to Phytophthora citrophthora in transgenic orange plants constitutively expressing a tomato pathogenesis related protein PR-5*. *Molecular Breeding*, 2001. **7**(2): p. 175-185.
113. Menu-Bouaouiche, L., et al., *A molecular basis for the endo-beta 1,3-glucanase activity of the thaumatococcus-like proteins from edible fruits*. *Biochimie*, 2003. **85**(1-2): p. 123-31.
114. Barre, A., et al., *Purification and structural analysis of an abundant thaumatococcus-like protein from ripe banana fruit*. *Planta*, 2000. **211**(6): p. 791-799.
115. Ho, V.S.M., J.H. Wong, and T.B. Ng, *A thaumatococcus-like antifungal protein from the emperor banana*. *Peptides*, 2007. **28**(4): p. 760-766.
116. Salzman, R.A., et al., *Inorganic cations mediate plant PR5 protein antifungal activity through fungal Mnn1-and Mnn4-regulated cell surface glycans*. *Molecular Plant-Microbe Interactions*, 2004. **17**(7): p. 780-788.
117. de A Campos, M., et al., *Expression in Escherichia coli, purification, refolding and antifungal activity of an osmotin from Solanum nigrum*. *Microbial Cell Factories*, 2008. **7**.
118. Kuwabara, C., et al., *Abscisic acid- and cold-induced thaumatococcus-like protein in winter wheat has an antifungal activity against snow mould, Microdochium nivale*. *Physiologia Plantarum*, 2002. **115**(1): p. 101-110.
119. Vitali, A., et al., *Purification and characterization of an antifungal thaumatococcus-like protein from Cassia didymobotrya cell culture*. *Plant Physiology and Biochemistry*, 2006. **44**(10): p. 604-610.
120. Garcia-Casado, G., et al., *Characterization of an apoplastic basic thaumatococcus-like protein from recalcitrant chestnut seeds*. *Physiologia Plantarum*, 2000. **110**(2): p. 172-180.
121. Chu, K.T. and T.B. Ng, *Isolation of a large thaumatococcus-like antifungal protein from seeds of the Kweilin chestnut Castanopsis chinensis*. *Biochemical and Biophysical Research Communications*, 2003. **301**(2): p. 364-370.
122. Emanuelsson, O., et al., *Locating proteins in the cell using TargetP, SignalP and related tools*. *Nature protocols*, 2007. **2**(4): p. 953-971.
123. Kelly, S.M., T.J. Jess, and N.C. Price, *How to study proteins by circular dichroism*. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 2005. **1751**(2): p. 119-139.
124. Sreerama, N. and R.W. Woody, *Computation and Analysis of Protein Circular Dichroism Spectra*, in *Methods in Enzymology*. 2004. p. 318-351.
125. Whitmore, L. and B.A. Wallace, *Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases*. *Biopolymers*, 2008. **89**(5): p. 392-400.
126. Whitmore, L. and B.A. Wallace, *DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data*. *Nucleic Acids Research*, 2004. **32**(suppl 2): p. W668-W673.

127. Altschul, S.F., et al., *Basic local alignment search tool*. Journal of Molecular Biology, 1990. **215**(3): p. 403-410.
128. Berman, H.M., et al., *The Protein Data Bank*. Nucleic Acids Research, 2000. **28**(1): p. 235-242.
129. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0*. Bioinformatics, 2007. **23**(21): p. 2947-2948.
130. Sali, A. and T.L. Blundell, *Comparative Protein Modelling by Satisfaction of Spatial Restraints*. Journal of Molecular Biology, 1993. **234**(3): p. 779-815.
131. Shen, M.Y. and A. Sali, *Statistical potential for assessment and prediction of protein structures*. Protein Science, 2006. **15**(11): p. 2507-2524.
132. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures*. Journal of Applied Crystallography, 1993. **26**(2): p. 283-291.
133. Humphrey, W., A. Dalke, and K. Schulten, *VMD: Visual molecular dynamics*. Journal of Molecular Graphics, 1996. **14**(1): p. 33-38.
134. Tama, F. and Y.H. Sanejouand, *Conformational change of proteins arising from normal mode calculations*. Protein Engineering, 2001. **14**(1): p. 1-6.
135. Suhre, K. and Y. Sanejouand, *ElNemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement*. Nucleic Acids Research, 2004. **32**: p. W610 - 4.
136. Hollup, S., G. Salensminde, and N. Reuter, *WEBnm@: a web application for normal mode analyses of proteins*. BMC Bioinformatics, 2005. **6**(1): p. 52.
137. De Brevern, A.G., et al., *Extension of a local backbone description using a structural alphabet: A new approach to the sequence-structure relationship*. Protein Science, 2002. **11**(12): p. 2871-2886.
138. Maupetit, J., P. Derreumaux, and P. Tufféry, *PEP-FOLD: an online resource for de novo peptide structure prediction*. Nucleic Acids Research, 2009.
139. Maupetit, J., P. Derreumaux, and P. Tufféry, *A fast method for large-scale De Novo peptide and miniprotein structure prediction*. J Comput Chem, 2009.
140. Camproux, A.C., et al., *Exploring the use of a structural alphabet for structural prediction of protein loops*. Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), 2001. **106**(1): p. 28-35.
141. Camproux, A.C., R. Gautier, and P. Tufféry, *A Hidden Markov Model Derived Structural Alphabet for Proteins*. Journal of Molecular Biology, 2004. **339**(3): p. 591-605.
142. Camproux, A.C., et al., *Hidden Markov model approach for identifying the modular framework of the protein backbone*. Protein Engineering, 1999. **12**(12): p. 1063-1073.
143. Tuffery, P. and P. Derreumaux, *Dependency between consecutive local conformations helps assemble protein structures from secondary structures using Go potential and greedy algorithm*. 2005, Wiley Subscription Services, Inc., A Wiley Company. p. 732-740.
144. Tuffery, P., F. Guyon, and P. Derreumaux, *Improved greedy algorithm for protein structure reconstruction*. 2005, Wiley Subscription Services, Inc., A Wiley Company. p. 506-513.
145. Arunkumar, A.I., et al., *Induction of helical conformation in all  $\beta$ -sheet proteins by trifluoroethanol*. Journal of Biomolecular Structure and Dynamics, 1996. **14**(3): p. 381-385.
146. Jayaraman, G., et al., *2,2,2-Trifluoroethanol induces helical conformation in an all  $\beta$ -sheet protein*. Biochemical and Biophysical Research Communications, 1996. **222**(1): p. 33-37.
147. Amodeo, P., et al., *Solution structure of nociceptin peptides*. Journal of Peptide Science, 2002. **8**(9): p. 497-509.
148. Saviano, G., et al., *Solution structure of human  $\beta$ -endorphin in helicogenic solvents: An NMR study*. Journal of Peptide Science, 1999. **5**(9): p. 410-422.
149. Sreerama, N. and R.W. Woody, *Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set*. Analytical Biochemistry, 2000. **287**(2): p. 252-260.

150. Wishart, D.S., B.D. Sykes, and F.M. Richards, *Relationship between nuclear magnetic resonance chemical shift and protein secondary structure*. Journal of Molecular Biology, 1991. **222**(2): p. 311-333.
151. Aue, W.P., E. Bartholdi, and R.R. Ernst, *Two-dimensional spectroscopy. Application to nuclear magnetic resonance*. The Journal of Chemical Physics, 1976. **64**(5): p. 2229-2246.
152. Bax, A. and D.G. Davis, *MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy*. Journal of Magnetic Resonance (1969), 1985. **65**(2): p. 355-360.
153. Kumar, A., R.R. Ernst, and K. Wüthrich, *A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules*. Biochemical and Biophysical Research Communications, 1980. **95**(1): p. 1-6.
154. Bodenhausen, G. and D.J. Ruben, *Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy*. Chemical Physics Letters, 1980. **69**(1): p. 185-189.
155. Bax, A. and M.F. Summers, *<sup>1</sup>H and <sup>13</sup>C assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR*. Journal of the American Chemical Society, 1986. **108**(8): p. 2093-2094.
156. Wüthrich, K., *Protein structure determination in solution by NMR spectroscopy*. Journal of Biological Chemistry, 1990. **265**(36): p. 22059-62.
157. Wishart, D.S., et al., *<sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shift referencing in biomolecular NMR*. Journal of Biomolecular NMR, 1995. **6**(2): p. 135-140.
158. Wishart, D.S., B.D. Sykes, and F.M. Richards, *The chemical shift index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy*. Biochemistry, 1992. **31**(6): p. 1647-1651.
159. Wishart, D.S. and B.D. Sykes, *The <sup>13</sup>C Chemical-Shift Index: A simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data*. Journal of Biomolecular NMR, 1994. **4**(2): p. 171-180.
160. Ramachandran, G.N., R. Chandrasekaran, and K.D. Kopple, *Variation of the NH-C alpha-H coupling constant with dihedral angle in the NMR spectra of peptides*. Biopolymers - Peptide Science Section, 1971. **10**(11): p. 2113-2131.
161. Cung, M.T., M. Marraud, and J. Neel, *Experimental calibration of a Karplus relationship in order to study the conformations of peptides by nuclear magnetic resonance*. Macromolecules, 1974. **7**(5): p. 606-613.
162. Bystrov, V.F., *Spin-spin coupling and the conformational states of peptide systems*. Progress in Nuclear Magnetic Resonance Spectroscopy, 1976. **10**(2): p. 41-81.
163. Pardi, A., M. Billeter, and K. Wüthrich, *Calibration of the angular dependence of the amide proton-CA proton coupling constants, <sup>3</sup>JHNHA, in a globular protein. Use of <sup>3</sup>JHNHA for identification of helical secondary structure*. Journal of Molecular Biology, 1984. **180**(3): p. 741-751.
164. Wüthrich, K., *NMR of Proteins and Nucleic Acids*. Wiley, New York., 1986.
165. Hwang, T.L. and A.J. Shaka, *Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients*. Journal of Magnetic Resonance, Series A, 1995. **112**(2): p. 275-279.
166. Thrippleton, M.J. and J. Keeler, *Elimination of zero-quantum interference in two-dimensional NMR spectra*. Angewandte Chemie - International Edition, 2003. **42**(33): p. 3938-3941.
167. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination*. Acta Crystallographica Section D: Biological Crystallography, 1998. **54**(5): p. 905-921.
168. Brunger, A.T., *Version 1.2 of the Crystallography and NMR system*. Nature protocols, 2007. **2**(11): p. 2728-2733.

- 
169. Stein, E.G., L.M. Rice, and A.T. Brunger, *Torsion-Angle Molecular Dynamics as a New Efficient Tool for NMR Structure Calculation*. Journal of Magnetic Resonance, 1997. **124**(1): p. 154-164.
170. Delano, W.L., *The PyMOL Molecular Graphics System 2002*. DeLano Scientific, San Carlos, CA, USA, 2002.
171. Koradi, R., M. Billeter, and K. Wüthrich, *MOLMOL: A program for display and analysis of macromolecular structures*. Journal of Molecular Graphics, 1996. **14**(1): p. 51-55.
172. Gougeon, R.D., et al., *Polypeptide adsorption on a synthetic montmorillonite: A combined solid-state NMR spectroscopy, X-ray diffraction, thermal analysis and N<sub>2</sub> adsorption study*. European Journal of Inorganic Chemistry, 2003(7): p. 1366-1372.
173. Gougeon, R.G.D., et al., *Solid-state NMR investigation on the interactions between a synthetic montmorillonite and two homopolypeptides*. Solid State Nuclear Magnetic Resonance, 2006. **29**(4): p. 322-329.
174. Shaw, W.J., et al., *Determination of statherin N-terminal peptide conformation on hydroxyapatite crystals*. Journal of the American Chemical Society, 2000. **122**(8): p. 1709-1716.
175. Younes, B., et al., *Evidence for an extracellular acid proteolytic activity secreted by living cells of Saccharomyces cerevisiae PIR1: Impact on grape proteins*. Journal of Agricultural and Food Chemistry. **59**(11): p. 6239-6246.
176. Palmisano, G., D. Antonacci, and M.R. Larsen, *Glycoproteomic profile in wine: A 'Sweet' molecular renaissance*. Journal of Proteome Research. **9**(12): p. 6148-6159.
177. Brooks, B.R., et al., *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations*. Journal of Computational Chemistry, 1983. **4**: p. 187-217.
178. Cornell, W., et al., *A second generation force field for the simulation of proteins and nucleic acids*. J Am Chem Soc, 1995. **117**: p. 5179 - 5197.
179. Arkenf, G.B. and H.J. Weber, *The method of steepest descents*. Mathematical Methods for Physicists, 3rd edition. Academic Press, Orlando, FL, USA, 1985.
180. Powell, M.J.D., *Restart procedures for the conjugate gradient method*. Mathematical Programming, 1977. **12**(1): p. 241-254.
181. Verlet, L., *Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules*. Physical Review, 1967. **159**(1): p. 98-103.
182. Brunger, A., *"XPLOR Manual Version 3.1"*. Yale University Press; New Haven, 1993.
183. Sali, A. and J.P. Overington, *Derivation of rules for comparative protein modeling from a database of protein structure alignments*. Protein Science, 1994. **3**(9): p. 1582-1596.
184. Jones, D.T., *Protein secondary structure prediction based on position-specific scoring matrices*. Journal of Molecular Biology, 1999. **292**(2): p. 195-202.
185. Maupetit, J., P. Tuffery, and P. Derreumaux, *A coarse-grained protein force field for folding and structure prediction*. Proteins: Structure, Function and Genetics, 2007. **69**(2): p. 394-408.
186. Case, D.A., et al., *The Amber biomolecular simulation programs*. Journal of Computational Chemistry, 2005. **26**(16): p. 1668-1688.
187. Jorgensen, W.L., et al., *Comparison of simple potential functions for simulating liquid water*. The Journal of Chemical Physics, 1983. **79**(2): p. 926-935.



## Le mot de la fin

*A poet once said, "The whole universe is in a glass of wine." We will probably never know in what sense he meant that, for poets do not write to be understood. But it is true that if we look at a glass of wine closely enough we see the entire universe. There are the things of physics : the twisting liquid which evaporates depending on the wind and weather, the reflections in the glass, and our imagination adds the atoms. The glass is a distillation of the Earth's rocks, and in its composition we see the secrets of the universe's age, and the evolution of stars. What strange arrays of chemicals are in the wine ? How did they come to be ? There are the ferments, the enzymes, the substrates, and the products. There in wine is found the great generalization : all life is fermentation. Nobody can discover the chemistry of wine without discovering, as did Louis Pasteur, the cause of much disease. How vivid is the claret, pressing its existence into the consciousness that watches it ! If our small minds, for some convenience, divide this glass of wine, this universe, into parts — physics, biology, geology, astronomy, psychology, and so on — remember that Nature does not know it ! So let us put it all back together, not forgetting ultimately what it is for. Let it give us one more final pleasure : **drink it and forget it all !***

*Richard Feynman, prix Nobel de physique 1965  
The Feynman Lectures on Physics, volume I, chapter 3, page 10 (1964)*