

## Table des matières

Introduction.....	7
1. Polysaccharides issus de l'amidon et lipides végétaux.....	10
2. Polysaccharides et lipides bactériens.....	34
3. Objectifs de la thèse.....	38
Références.....	41
Méthodes.....	63
Chapitre 1. La Résonance Magnétique Nucléaire.....	65
1. Historique.....	65
2. Principe et fonctionnement.....	67
3. Formalisme mathématique.....	72
4. Techniques expérimentales utilisées.....	82
Chapitre 2. Outils théoriques utilisés.....	86
1. Calculs d'écrantage chimique.....	86
2. Génération des structures moléculaires utilisées.....	96
Références.....	107
Travaux réalisés.....	109
Chapitre 1. Etudes théoriques de complexes d'amylose.....	111
1. Contexte et objectifs du chapitre.....	111
2. Article n°1.....	115
3. Article n°2.....	140
4. Conclusion : Validité et limites de notre approche.....	173
Chapitre 2. Etude théorique et expérimentale de complexes d'amylose en présence de lipides divers.....	175
1. Avant-propos : méthodes d'études et matières premières.....	177
2. Complexation de lipides « classiques ».....	186

3.	Utilisation de lipides extraits de mycobactéries et corynébactéries.....	206
4.	Conclusion .....	224
	Références.....	226
	Chapitre 3. Etude théorique et expérimentale préliminaire des interactions entre polysaccharides branchés en présence de lipides divers. ....	228
1.	Résultats expérimentaux.....	229
2.	Résultats théoriques.....	233
3.	Conclusion, expériences futures. ....	244
	Références.....	247
	Conclusion et perspectives .....	249
	Résumé des travaux réalisés.....	252
	Etablissement de la stratégie d'étude théorique.....	252
	Etude de complexes d'amylose en présence de lipides contenant plusieurs chaînes hydrophobes. ....	253
	Etude d'interactions entre polysaccharides ramifiés et lipides divers. ....	254
	Perspectives d'études. ....	255
	Références.....	260
	Annexes.....	263
	Annexe 1. Outils de mécanique quantique .....	265
1.	Equation de Schrödinger.....	265
2.	Méthode de résolution de l'équation.....	268
	Références.....	287
	Annexe 2. Article n°1 : Informations supplémentaires .....	290
	Annexe 3. Article n°2 : Informations supplémentaires .....	300

## **Abréviations :**

AFM : Microscopie à force atomique

AsDg : Amylosucrase issue de *Deinococcus geothermalis*

BET : acide  $\beta$ -cétomycolique

CC : Coupled Cluster

CGF : Contracted Gaussian Functions

DSC : Calorimétrie à balayage différentiel

DFT : Théorie de la Fonctionnelle de la Densité

DMSO : Dimethylsulfoxyde

DP : degré de polymérisation

DS : Déterminant de Slater

DIM : dimycocérosate de phtiocérol

DLS : diffusion dynamique de la lumière

DPPC : 1,2-Dipalmitoylphosphatidylcholine

DOPC : 1,2-Dioleoyl-sn-Glycero-3-Phosphocholine

DMT : Dimycolates de tréhalose

FID : Free Induction Decay

GGA : Generalized Gradient Approximation

GAFF : General Amber Force Field

GTO : Gaussian-type orbitals

GIPAW : Gauge Including Projector Augmented Wave

HF : Hartree-Fock

IRM : imagerie par résonance magnétique

IGLO : Individual Gauge for Localized Orbital

IPBS : Institut de Pharmacologie et Biologie Structurale

LDA : Local Density Approximation

LAPW : Linearized Augmented Plane-Wave

LSDA : Local Spin-Density Approximation

LBS3 : laboratoire de niveau de sécurité 3

MD : Dynamique moléculaire

MTB : *Mycobacterium tuberculosis*

MYC : Acides corynomycoliques

MMT : Monomycolates de tréhalose

NVT : ensemble canonique

NPT : ensemble Isotherme-isobare

OLE : Acide oléique

POPE : 1-palmitoyl-2-oleoyl phosphatidyléthanolamine

PP : Pseudopotential Planewave

PAW : Projector Augmented Wave

PAL : Acide palmitique

POPC : 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

QM : Mécanique quantique

RMN : Résonance Magnétique Nucléaire

SAXS : Diffraction rayon X aux petits angles

SCF : Self Consistent Field

SEM : Microscopie à balayage électronique

STO : Slater-type orbitals

SEP : surface d'énergie potentielle

TEM : Microscopie à transmission électronique

TB MDR : Multidrug resistant

TB XDR : Extensively drug resistant

TDM : dimycolates de tréhalose

XRD : diffraction aux rayons X



Rapport-Gratuit.com

## Introduction



Les interactions entre polysaccharides et lipides forment un champ de recherche aujourd'hui extrêmement vaste. En son sein, nous pouvons notamment recenser de très nombreuses études concernant les polysaccharides issus de l'amidon que sont l'amylose et l'amylopectine. Ces deux molécules présentent des intérêts particuliers pour les industries pharmaceutiques et alimentaires en raison de leurs propriétés structurales particulières. En effet, elles sont connues pour adopter des formes hélicoïdales<sup>1</sup>, permettant l'encapsulation de divers molécules hydrophobes, telles que des lipides<sup>2</sup>, des composés aromatiques<sup>3,4</sup> ou des molécules bioactives, telles que des médicaments<sup>5</sup>.

Ces molécules sont aussi présentes au sein d'autres systèmes biologiques, tels que la capsule de *Mycobacterium tuberculosis*<sup>6</sup>. Cette partie de l'enveloppe mycobactérienne est composée principalement d'un polysaccharide appelé  $\alpha$ -glucane<sup>7,8</sup>, analogue de l'amylose et l'amylopectine. On retrouve aussi certains lipides de la mycobactérie connus pour en être des facteurs de virulence, tels que les dimycocérosates de phthiocérol et les monomycolates et dimycolates de trehalose<sup>9</sup>. Cependant, les interactions entre ces molécules ainsi que la structure intrinsèque de la capsule restent aujourd'hui méconnues.

L'étude de ces interactions au sein de ces deux environnements est donc fondamentale et pourrait amener des avancées dans des domaines divers, allant de la formulation d'aliments présentant des propriétés de biodisponibilité particulières à la création de nouveaux adjuvants vaccinaux. Afin de contextualiser le travail réalisé lors de cette thèse, l'état de l'art concernant les interactions entre ces molécules va maintenant être décrit. Nous commencerons par rappeler les travaux réalisés sur les composants de l'amidon, avant de présenter les connaissances actuelles concernant la capsule de *Mycobacterium tuberculosis*.

# 1. Polysaccharides issus de l'amidon et lipides végétaux.

## I. Amidon

### a. *Provenance et utilisation*

L'amidon est un mélange de polysaccharides intégré dans le régime alimentaire humain et servant de réserve d'énergie à la plupart des plantes. Les sources principales d'amidon consommées par les humains sont les céréales, telles que le blé, le riz ou le maïs et les tubercules, telles que les pommes de terre, les patates douces ou le tapioca<sup>10,11</sup>. Ce matériau est donc largement étudié dans le cadre de l'industrie agro-alimentaire, mais nous allons voir qu'il présente aussi un intérêt particulier dans l'industrie pharmaceutique actuelle.

Dans le cadre de l'industrie alimentaire, des études sont intensément menées sur les amidons résistants à la digestion enzymatique. Tout d'abord, il a été prouvé que l'amylose pouvait encapsuler divers types de composants, dont des arômes<sup>3,4</sup> ou des lipides<sup>12-18</sup>, permettant ainsi de mieux appréhender l'impact de la cuisson sur certains aliments.

De plus, les amidons comprenant ce type de complexes sont actuellement étudiés pour des applications pharmaceutiques, afin de formuler des aliments à index glycémique réduit<sup>19</sup>, utiles dans le cadre du traitement du diabète, ou pour prévenir des maladies digestives telles que les cancer colorectaux<sup>20-22</sup>. En outre, des études menées récemment semblent indiquer que les granules d'amidon pourraient servir d'adjuvant dans le cadre du vaccin BCG<sup>23,24</sup>, actuellement standard dans la prévention de la tuberculose.

### b. *Composition et structure*

Comme cité ci-dessus, l'amidon est un mélange de deux polysaccharides : l'amylopectine et l'amylose. Ces deux molécules sont principalement composées de résidus glycosyl branchés en  $[\alpha-1\rightarrow4]$ , et présentent des ramifications en  $[\alpha-1\rightarrow6]$

(Figure 1). L'amylopectine contient 5-6% de résidus branchés<sup>25</sup>, et il est connu que l'amylose présente un très faible pourcentage de résidus branchés<sup>26,27</sup>. La plupart des amidons connus sont majoritairement composés d'amylopectine, à hauteur de 65-85%<sup>28</sup>. Il est cependant aujourd'hui possible de trouver des amidons composés en majorité d'amylose, que l'on appelle « high-amylose starches »<sup>29</sup>, amidon à haut taux d'amylose, et des amidons composés d'un pourcentage encore supérieur d'amylopectine, pouvant atteindre 100%, appelés « waxy starches »<sup>30,31</sup>, ou amidon cireux.

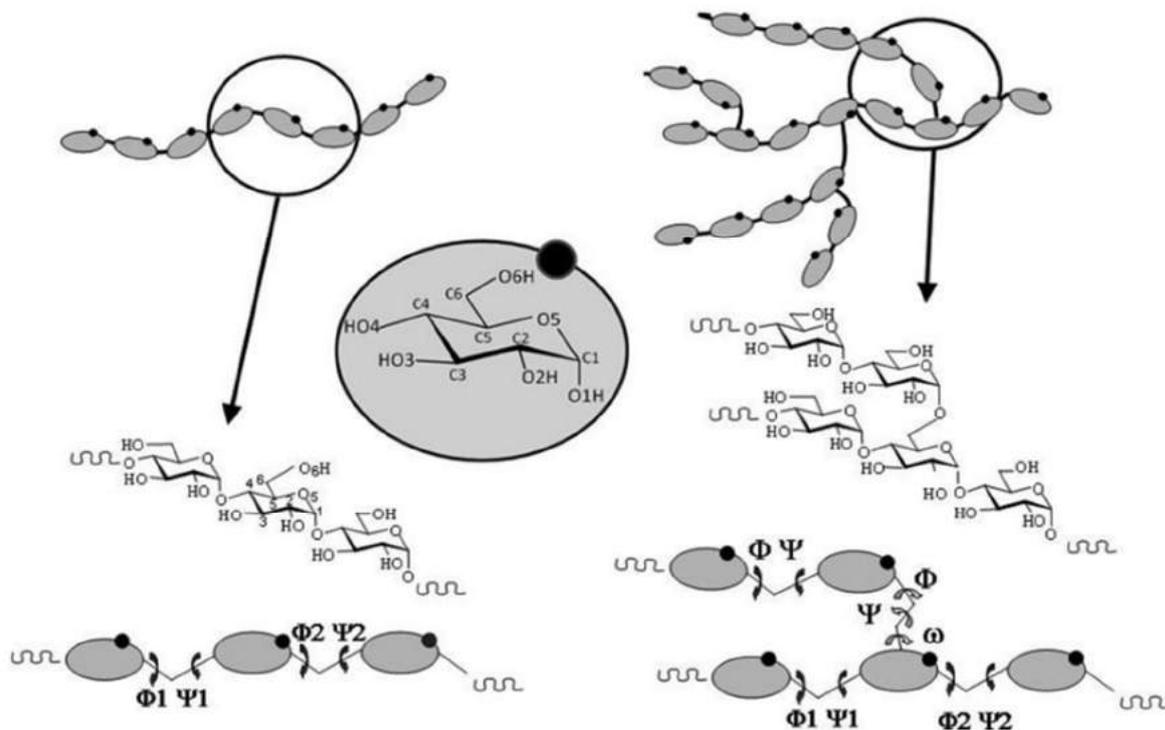


Figure 1: Structures primaires de l'amylose (gauche) et de l'amylopectine (droite). Représentation des différents angles dièdres d'intérêt :  $\phi$ ,  $\psi$  et  $\omega$ . Extrait de Perez (2010).

La Figure 2 résume les structures connues à différentes échelles qu'on trouve dans l'amidon. Il se présente sous forme de granules dont l'ultrastructure se résume à des sphères concentriques alternant entre phases semi-cristallines et cristallines<sup>32</sup>. Chacune de ces phases est composée de sous-unités que l'on appelle des « blocklets ». Ce terme a été utilisé pour la première fois en 1934<sup>33</sup>, et depuis leur redécouverte en 1997<sup>34</sup>, les structures qu'il définit sont encore étudiées aujourd'hui<sup>35</sup>. Les blocklets sont des édifices moléculaires considérés approximativement comme sphériques, dont le diamètre varie de 20 à 500 nm selon les sources de l'amidon

étudié. Le lien entre les blocklets et les phases cristallines et semi-cristallines n'est toujours pas univoquement établi, cependant il est supposé que les propriétés moléculaires des blocklets influencent celles des phases les contenant.

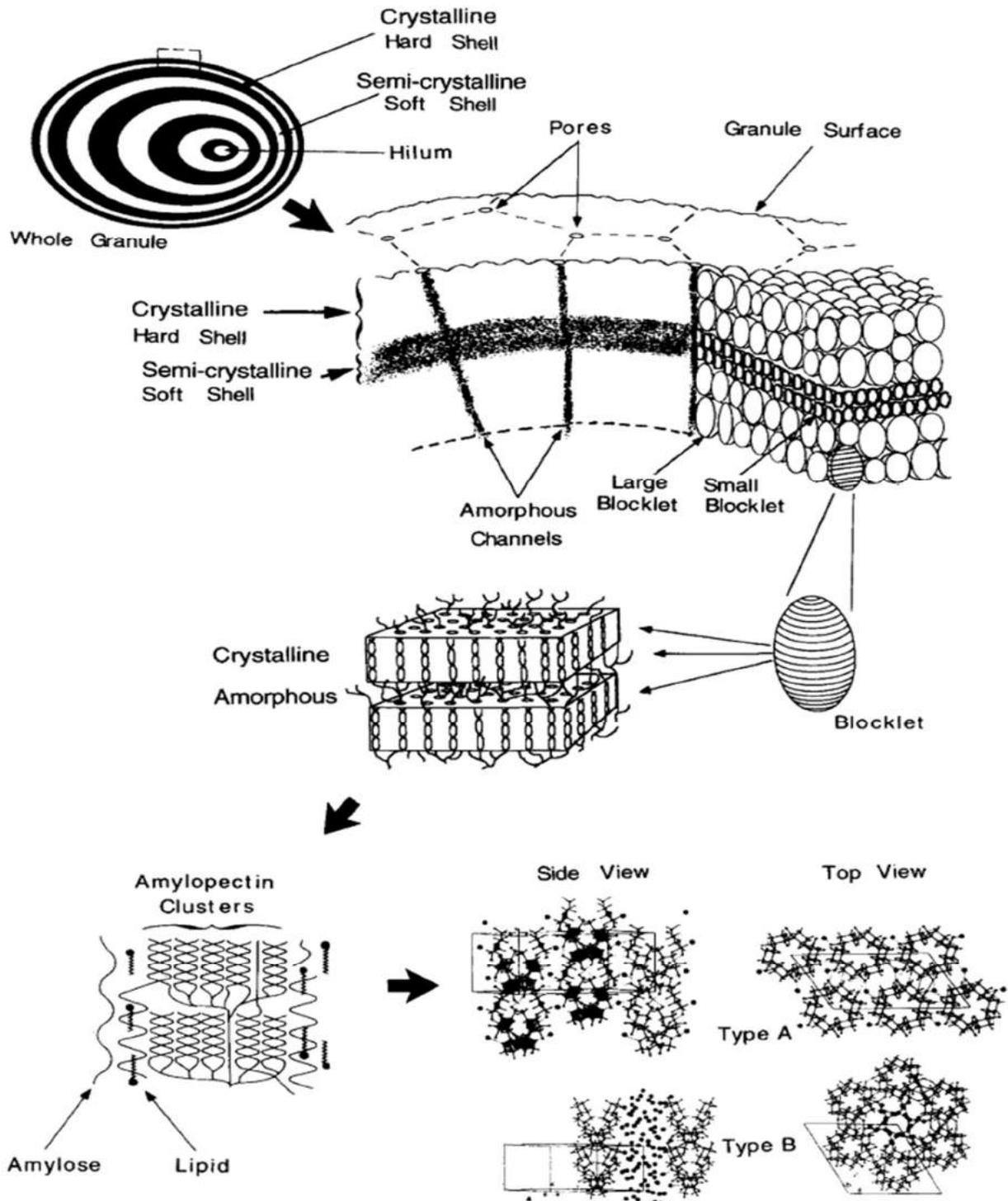


Figure 2: Structures à différentes échelles de l'amidon. Représentation du granule complet (haut, gauche), schéma des phases semi-cristallines et cristallines sous forme d'empilement de « blocklets » (haut, droite), description des blocklets en tant qu'alternance de lamelles cristallines et amorphes (milieu), description de l'organisation spatiale des molécules d'amylose, d'amylopectine et de lipides dans les lamelles (bas, gauche) et description des polymorphes A et B formant les phases cristallines (bas, droite). D'après Galland (1997).

Ainsi, d'après Gallant et co-auteurs<sup>34</sup>, les phases semi-cristallines seraient composées de blocklets plus petits en taille ceux des phases cristallines. Plus récemment, Tang et co-auteurs proposèrent un modèle dans lequel les blocklets présents en phase semi-cristalline seraient en fait imparfaits (« defective »), résultant en une cristallinité amoindrie<sup>36</sup>. Selon ces derniers, les blocklets seraient eux-mêmes composés de structures lamellaires établies par l'amylopectine, qui alternent entre des phases cristallines et amorphes.

Le lien entre blocklets et phases lamellaires reste cependant non-établi, malgré la caractérisation des phases lamellaires depuis 1962<sup>37</sup> par diffraction aux petits angles. Les lamelles cristallines d'amylopectine sont ainsi formées par les parties linéaires du polysaccharide<sup>38-41</sup>. Celles-ci forment des doubles hélices, qui selon leurs caractéristiques intrinsèques seront attribuées au polymorphe de type A ou B, dont nous discuterons dans une prochaine partie. Les points de branchements forment alors les parties amorphes et leur taux d'hydratation est connu pour influencer l'arrangement des phases cristallines<sup>42</sup>.

Le rôle et la structure de l'amylose et des lipides présents dans l'amidon sont encore flous à ce jour. Il a été postulé que l'amylose pourrait se situer principalement dans les phases amorphes des structures lamellaires, où elle pourrait notamment former des complexes d'inclusion amylose-lipides. Ces complexes pourraient cependant être aussi situés au sein des phases cristallines et ainsi former des défauts. Ces derniers viendraient s'ajouter à ceux déjà présents et dus à l'inhomogénéité des chaînes terminales de l'amylopectine<sup>43,44</sup>. Les structures de l'amylose en dehors de l'amidon seront discutées dans la partie II de cette introduction.

### *c. Transformations de l'amidon*

On retrouve deux termes fréquemment utilisés dans la littérature concernant les transformations de l'amidon : La gélatinisation et la rétrogradation (Figure 3).

La gélatinisation décrit tous les changements de textures et de structures qui apparaissent lorsque l'amidon est chauffé en présence d'eau<sup>45,46</sup> (Figure 3, phase (II)). Ce processus était à l'origine considéré comme le passage de l'état solide à l'état

liquide de l'amidon, dû au pic endotherme observé en calorimétrie à balayage différentiel (DSC). C'est en 1988 que Zobel et Co-auteurs ont pu montrer par une étude de diffraction aux rayons X qu'il s'agissait de 2 processus différents<sup>46</sup>. Lors de la gélatinisation, les différents réseaux de liaisons hydrogène présents dans les doubles hélices se cassent. Ceci engendre la perte de structure cristalline locale et la disruption des granules d'amidon<sup>47-50</sup>. Les effets des différents composants de l'amidon sur la gélatinisation ainsi que ceux de ce processus sont étudiés encore aujourd'hui<sup>16,51-53</sup>.

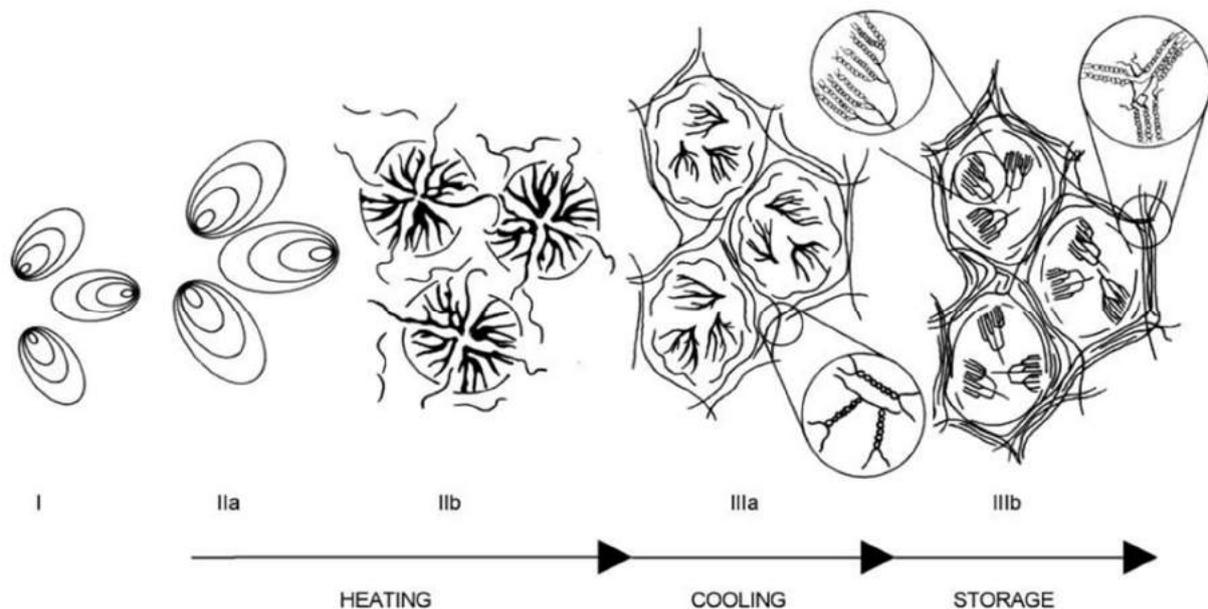


Figure 3: (I) Granules sans altérations. Lors de la montée en température on assiste dans un premier temps au gonflement des granules (II), puis à l'extraction aqueuse des contenus solubles, tels que l'amylose (III) qui aboutissent en une phase gélatinée. Lors de la redescende en température, on assiste dans un premier temps à la rétrogradation de l'amylose (IIIa) suivi de celle de l'amylopectine (IIIb). Extrait de Goesaert, (2005)<sup>192</sup>.

Suivant la gélatinisation, qui intervient donc à haute température, la rétrogradation décrit les différents changements de textures et de structures lors du refroidissement subséquent ainsi que ceux se déroulant lors du stockage de l'amidon<sup>54,55</sup> (Figure 3, phase (III)). En résulte une réorganisation moléculaire plus structurée notamment grâce à la formation de structures cristallines de type polymorphe B<sup>56</sup>. Cependant, en présence de lipides, on peut observer la formation de complexes d'inclusion, de type polymorphe V. Il a d'abord été postulé que la formation de ces complexes se déroulait soit pendant la phase à haute température, soit pendant la gélatinisation. Cependant les études les plus récentes semblent indiquer que c'est pendant la redescende en température que ce type de complexes cristallisent, laissant apparaître un pic de

viscosité lors de cette dernière<sup>57-59</sup>. Enfin, il est connu que le taux d'hydratation et la température forment deux facteurs influant sur cette étape<sup>60</sup>.

## II. L'amylose et ses structures

### a. Historique et production d'amylose

Bien que le terme d'amylose fut connu avant, c'est seulement à partir de 1940<sup>61</sup> qu'il a commencé à être utilisé par Meyer et co-auteurs afin de caractériser la partie linéaire, soluble dans l'eau à haute température, de l'amidon. Suite à cela, il a été démontré en 1943<sup>62</sup> que l'amylose était responsable de la complexation de l'iode, lorsque l'on en ajoute dans de l'amidon. Ceci engendra par ailleurs le développement d'un test colorimétrique qui encore aujourd'hui permet de caractériser la quantité d'amylose présente dans un amidon considéré<sup>63</sup>. Cette quantité est d'importance car connue pour influencer certaines propriétés de l'amidon, telle que sa rétrogradation<sup>55,64</sup> ou sa capacité à être digérée par voie enzymatique.

Il existe principalement deux façons d'obtenir de l'amylose : la purification d'amylose à partir d'un amidon donné et sa synthèse enzymatique.

La méthode utilisée par Meyer et co-auteurs afin de séparer l'amylose de l'amylopectine fait partie des méthodes dites de fractionnement de l'amidon. Isao Hanachiro divise cette catégorie en 4 sous-catégories que sont l'extraction aqueuse<sup>61</sup>, la précipitation de l'amylose à l'aide d'un agent complexant, la chromatographie de perméation de gel et la formation d'un complexe insoluble entre l'amylopectine et une lectin A appelée concanavalin A. Les avantages et inconvénients de chaque méthode sont résumés dans le chapitre 2 du livre *Starch : Metabolism and Structure* de Yasunori Nakamura et ne forment pas le propos de ce manuscrit.

En parallèle de ces méthodes, il est possible d'effectuer des synthèses enzymatiques en utilisant différentes enzymes à l'instar de celles listées dans le Tableau 1<sup>65</sup>. L'avantage principal de ces méthodes réside dans l'homogénéité plus grande de l'amylose obtenue et dans le fait qu'elles évitent la contamination par des molécules d'amylopectine. Parmi ces enzymes nous retiendrons la glucane phosphorylase et l'amylosucrase, car ce sont les deux qui seront rencontrées directement ou indirectement lors de ce travail.

Tableau 1: Liste des différentes enzymes synthétisant de l'amylose ainsi que le substrat utilisé. ADP : Adénosine diphosphate. UDP : Uridine diphosphate. G-1-P : Glucose-1-phosphate.

Starch (Glycogen) synthase	ADP-glucose (UDP-glucose)
Isoamylase	Starch
CGTase	Cyclodextrin
D-enzyme	Maltodextrin
Amylosucrase	Sucrose
Glucan phosphorylase	G-1-P
Sucrose phosphorylase + Glucan phosphorylase	Sucrose
Cellobiose phosphorylase + Glucan phosphorylase	Cellobiose

La glucane phosphorylase est pour le moment le meilleur choix en termes de maîtrise de la taille de l'amylose produite<sup>66</sup>, cette dernière pouvant atteindre un degré de polymérisation (DP) de 1400, soit 1400 résidus glucose, en moyenne. Cependant, le substrat requis est le Glucose-1-phosphate, connu pour son prix élevé, à contrario du sucrose, utilisé par les amylosucrases<sup>67-71</sup>. Ces dernières permettent de générer des chaînes beaucoup plus courtes, pouvant aller jusqu'à un DP moyen de 100, qui ont notamment été étudiées en raison de la haute cristallinité des polymorphes qu'elles génèrent après rétrogradation<sup>72,73</sup>.

#### b. Amylose en solution.

Entre les phases de gélatinisation et de rétrogradation, à température élevée (~100 degrés), l'amylose est donc soluble<sup>61</sup>, mais il n'y a pas encore de consensus actuel sur sa structure en solution. Pour commencer, ses caractéristiques dépendent du solvant utilisé, et il a été démontré que le diméthylsulfoxyde (DMSO), l'eau alcaline et le formamide formaient de bons solvants pour l'amylose<sup>74</sup>. Cependant, dans une logique de formulation pour des produits alimentaires ou pharmaceutiques, l'utilisation d'eau à pH neutre devrait être privilégiée. En comparaison avec les trois solvants précédemment cités, l'amylose est généralement moins soluble dans l'eau neutre et cette solubilité dépend de différents facteurs (DP, source d'amylose ...). La température joue un rôle important, comme décrit dans les étapes de gélatinisation. Si l'amylose que l'on trouve naturellement dans l'amidon ne se dissout qu'à chaud<sup>61</sup>, il a été montré qu'elle peut être dissoute à froid dans de l'eau neutre<sup>75</sup> après avoir subi un

traitement particulier, passant notamment par du chauffage à haute pression dans un autoclave.

Le traitement réalisé dans un autoclave est par ailleurs un des rares permettant de dissoudre à nouveau l'amylose rétrogradée. Dans ces cas-là, il est alors nécessaire de faire passer un flux d'azote dans la solution avant d'effectuer le chauffage à haute pression, afin d'éviter l'oxydation et la dégradation du soluté.

La structuration de l'amylose en solution a fait l'objet de nombreuses études, expérimentales<sup>74,76,77</sup> et ,depuis peu, théoriques<sup>78,79</sup>. Plusieurs modèles ont été postulés concernant sa conformation : on y retrouve des structures sous forme de « random coil », des structures hélicoïdales ou un mélange des deux<sup>74</sup> (Figure 4).

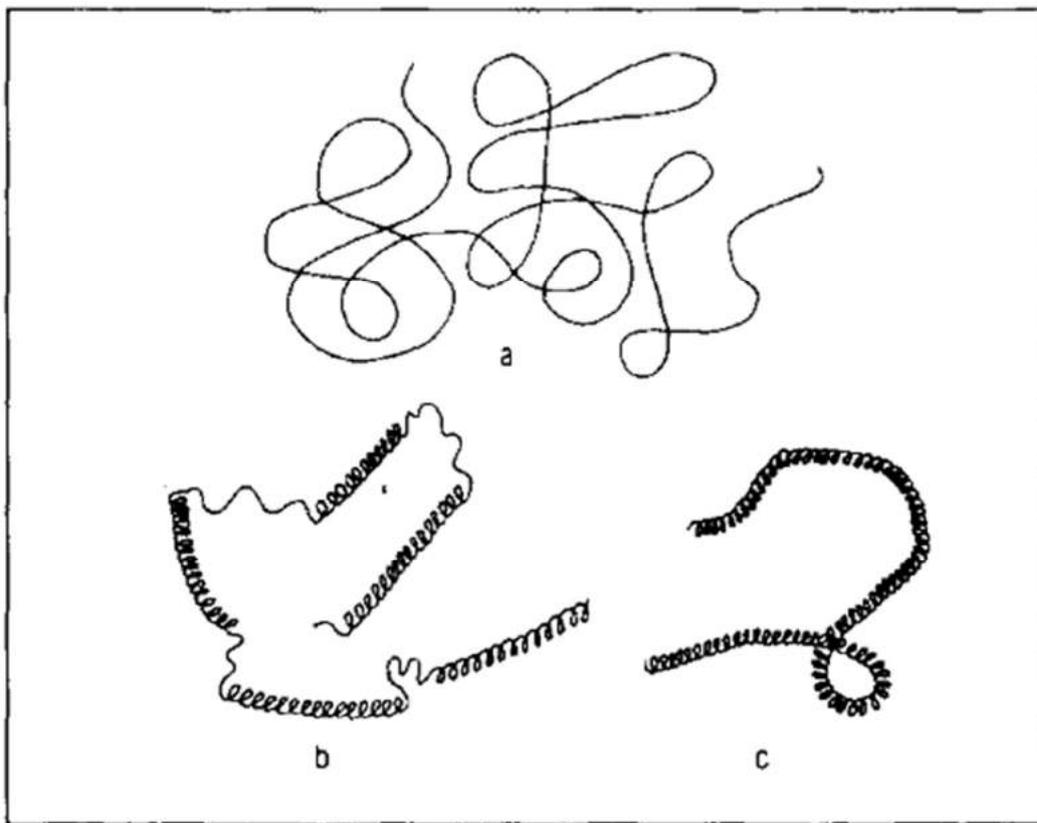


Figure 4: Structure random coil (a), partiellement hélicoïdale (b) et hélicoïdale (c). Extrait de Banks et al. (1971).

Dans la littérature, nous rencontrons bon nombre d'études semblant confirmer que l'amylose adopte une structure plutôt hélicoïdale en solution. En 1952, Foster et Zucker concluèrent que l'amylose adoptait la même conformation en solution que complexée avec de l'iode<sup>80</sup>. Pour rappel, il a par la suite été démontré que l'amylose s'arrange sous forme hélicoïdale quand complexée avec de l'iode<sup>81,82</sup>. De plus, Hollo et Szejtli

remarquèrent aussi que l'ajout d'iode n'affectait que très peu la viscosité d'une solution d'amylose, suggérant de ce fait que la conformation de l'amylose ne changeait pas<sup>83</sup>. Une observation similaire avait été faite par Rao et Foster sur des échantillons contenant de l'amylose et du dodecyl sulfate de sodium<sup>84</sup>. De plus les auteurs démontrèrent que la viscosité à pH neutre et à pH alcalin était différente, suggérant des conformations différentes de l'amylose dans ces conditions.

A contrario, Banks et Greenwood démontrèrent que la viscosité de l'amylose en solution pouvait décroître en présence d'agents complexant, dans un article appelé « Amylose : a non-helical biopolymer in solution »<sup>85</sup>. De plus, en 1985 Jane et co-auteurs montrèrent, par le biais d'expériences de RMN à l'état liquide, que l'ajout d'agents complexant en solution affectait les déplacements chimiques des carbones C1 et C4 (tels que notés Figure 1). Ces déplacements chimiques se déplaçaient vers les champs plus faibles, indiquant le rôle important des agents complexant dans la transition d'une conformation hélicoïdale à une structure de type random coil.

Le consensus expérimental n'étant pas établi, une possible solution à cette situation pourrait venir des méthodes théoriques d'étude de conformation de l'amylose en solution. En 2011, Tusch et co-auteurs purent démontrer que la conformation en hélice V (conformation qui sera abordée plus tard) sans présence d'agent complexant était instable en solution aqueuse et dans le DMSO<sup>78</sup>. La conformation en hélice V est cependant maintenue plus longtemps dans le DMSO que dans l'eau en raison d'un nombre de liaisons hydrogène établies avec le solvant inférieur à celui observé en solution aqueuse. Il a d'ailleurs été démontré récemment par dynamique moléculaire que, lors d'une augmentation de la pression à laquelle on réalise la gélification, les liaisons hydrogène intra moléculaires sont détruites au profit de l'établissement de liaisons hydrogène avec le solvant<sup>53</sup>. Plus la pression est élevée plus ce phénomène est important, impliquant ainsi la déformation de structures hélicoïdales de l'amylose.

De plus, une étude de Koneru et co-auteurs sortie très récemment a pu démontrer que les deux types de conformations existaient en solution grâce à l'utilisation de dynamique moléculaire<sup>79</sup>. S'agissant d'une étude visant à comparer la robustesse des champs de forces actuels, le poids respectif de chaque conformation a été démontré comme étant dépendant du champ de force utilisé (Figure 5). Il semblerait donc que les deux types de structures soient adoptées par l'amylose en solution, du moins avant que la rétrogradation ait lieu.

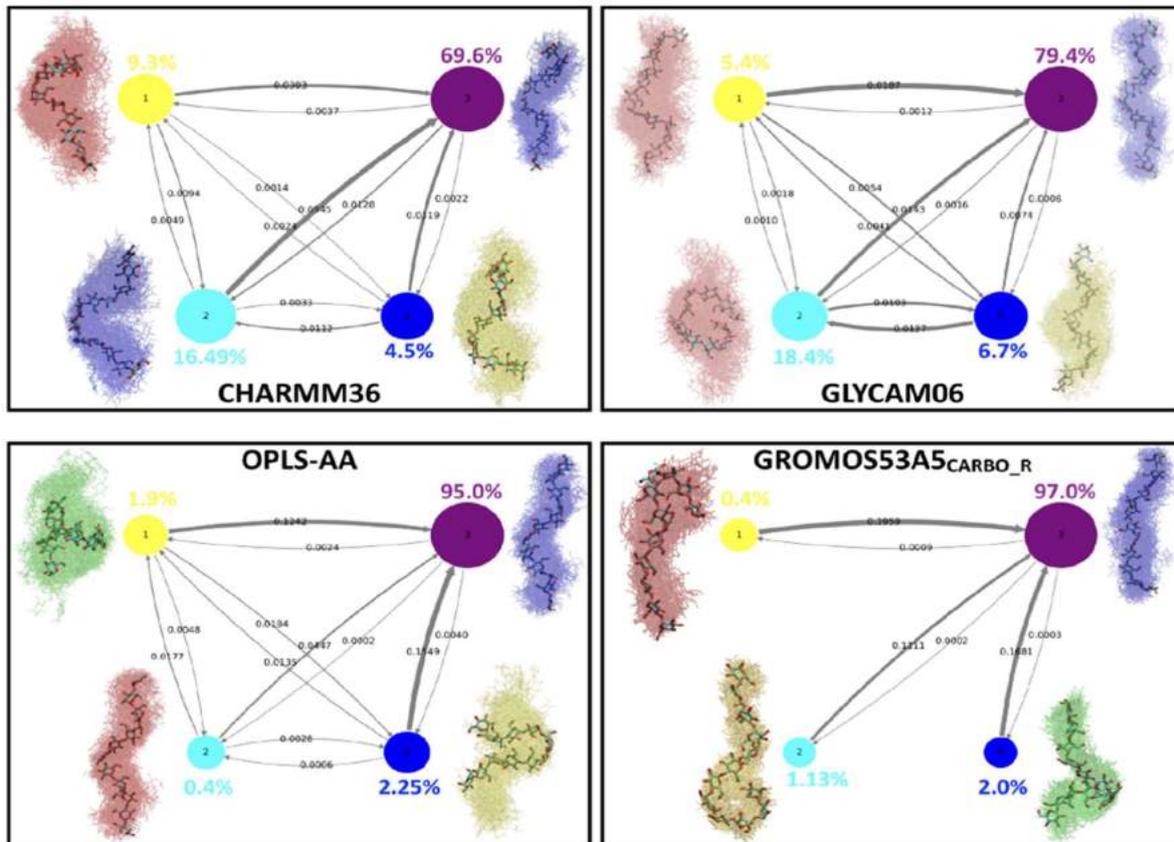


Figure 5: Poids des différentes conformations prises par l'amylose en solution aqueuse pour 4 champs de forces différents. Extrait de Koneru et al. (2019).

Enfin, il est important de rappeler que l'amylose ne reste pas soluble à température ambiante et que la rapidité de la rétrogradation dépend de plusieurs facteurs mais surtout de la taille du polymère. L'amylose de DP relativement bas (~80-100) engendre rapidement des précipités hautement cristallins, et l'amylose de haut DP (>1000) peut prendre plusieurs semaines avant de commencer à précipiter. Ce comportement a par ailleurs été démontré par Pfannemüller dans des solutions aqueuses d'amylose contenant 5% de DMSO<sup>86</sup>.

### c. Amylose cristalline : A et B

Comme nous l'avons mentionné dans la partie I de cette introduction, l'amylopectine peut prendre la forme de doubles hélices appelées polymorphe A ou B selon leurs caractéristiques. A l'état naturel, le polymorphe A est plutôt présent dans les céréales, tandis que le B se retrouve dans les tubercules et amidons à haute

concentration en amylose. Pendant longtemps il n'était pas possible d'étudier en détails la structure de l'amidon, car sa cristallinité n'est pas totale, comme décrit plus haut. De ce fait l'amylose pouvant être recristallisée *in vitro* sous différentes formes servait de modèle d'étude pour comprendre la structure moléculaire des grains d'amidon<sup>87</sup>. Il est généralement établi que la synthèse *in vitro* de chaque polymorphe nécessite différentes conditions de préparation. La formation de polymorphe B est engendrée par des chaînes plus longues (DP supérieurs à 30 résidus) et des températures de recristallisation plus basses (de l'ordre de 4°C) alors que le polymorphe A est formé dans le cadre d'amylose de chaîne plus courte (DP de l'ordre de 10-15 résidus), à haute concentration et température (60°C)<sup>88</sup>.

Les deux polymorphes sont composés de doubles hélices gauches d'amylose présentant 6 résidus glucose par tour, dont le pas varie entre 2.08 et 2.13 nm<sup>89-91</sup>. Si des modèles structuraux ont tout d'abord proposé des hélices droites<sup>92</sup>, les modèles les plus récents postulent des hélices gauches, ces dernières étant plus stables<sup>89-91</sup>. Les différences structurales intrinsèques entre le polymorphe A et le B résident dans leur empilement cristallin (Figure 6). Le premier cristallise dans le groupe d'espace B2 avec un système réticulaire monoclinique. Il présente quatre molécules d'eau par cellule unitaire<sup>89</sup>. Le second cristallise dans le groupe d'espace P6<sub>1</sub> avec une cellule unitaire hexagonale contenant 12 molécules d'eau. Le

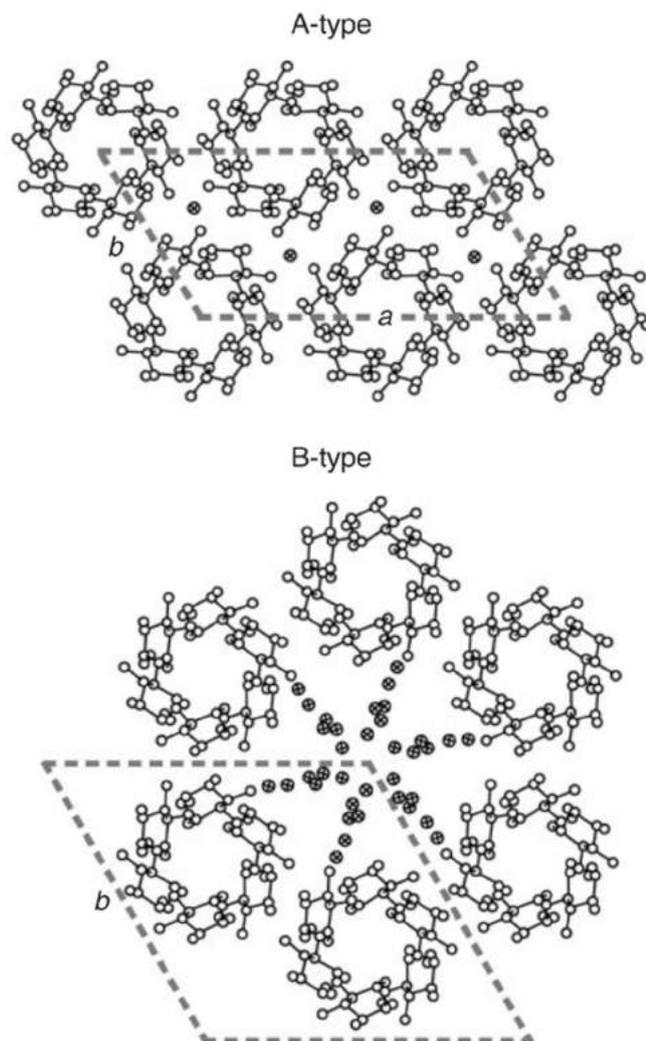


Figure 6: Empilement cristallin dans le plan (a,b) des polymorphes A et B de l'amylose. Les structures sont vues depuis l'axe hélicoïdale et les ⊕ représentent les molécules d'eau. Extrait de Buléon et al. (2007).

profil de diffraction aux rayons X de ces deux polymorphes permet de les identifier

formellement (Figure 7 (a)), et il en est de même pour leur signal RMN  $^{13}\text{C}$ . Il est par ailleurs admis que la différence de packing cristallin entre les deux polymorphes est responsable de leurs signaux RMN caractéristiques.

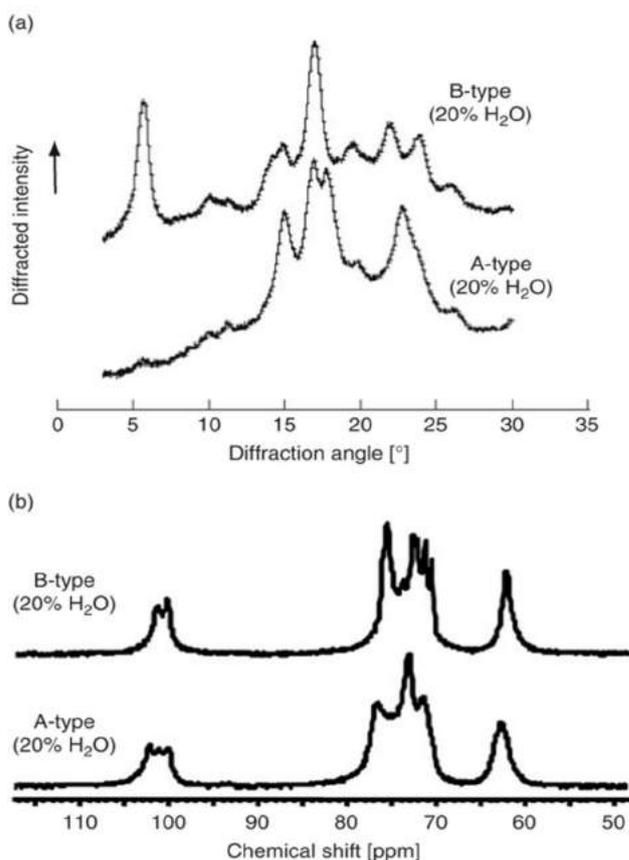


Figure 7: Profil de diffraction aux rayons X (a) et spectres RMN  $^{13}\text{C}$  à l'état solide (b) des polymorphes A et B. Extrait de Buléon et al. (2007).

La différence majeure entre ces deux signaux se situe au niveau du  $\text{C}_1$ , celui du polymorphe A se présentant sous la forme d'un triplet, et celui du polymorphe B sous la forme d'un doublet qui attestent de la présence de 3 et 2 résidus différents respectivement au sein de ces polymorphes (Figure 7 (b)). L'attribution complète du spectre RMN d'un échantillon hautement cristallin d'amylose sous la forme d'un polymorphe B a par ailleurs été réalisé en 2006 par Rondeau-Mouro et co-auteurs<sup>72</sup>, prouvant la présence de deux résidus différents. Les auteurs utilisèrent

une séquence d'impulsion en 2D appelée INADEQUATE qui leur a permis de caractériser les déplacements chimiques des carbones  $\text{C}_2$  à  $\text{C}_5$ . Cette séquence fut réutilisée récemment afin de caractériser des échantillons d'amidons riche en amylose, riche en amylopectine et natif provenant de céréales<sup>93</sup>. Ces expériences permirent aux auteurs de caractériser la cristallinité de l'amidon produit par une algue appelée *C. reinhardtii* en utilisant de la RMN en cellule.

#### d. Complexes d'inclusion : Polymorphe V

Le polymorphe V a été caractérisé par Katz pour la première fois en 1930 en utilisant la diffraction aux rayons X<sup>94,95</sup>. Ce profil de diffraction étant observé lors de la

gélatinisation, « Verkleisterung » en allemand, il fut nommé « V-spectrum » dans cet article, ce qui engendra le nom de polymorphe V par la suite<sup>96</sup>. Aujourd'hui cette dénomination représente les complexes d'inclusion de l'amylose en présence de molécules hydrophobes, représentés dans la Figure 8<sup>81,97</sup>.

Il s'agit donc d'une hélice gauche simple, dont la cavité interne est hydrophobe, à l'intérieur de laquelle se situe l'agent complexant. Nous avons nous-mêmes observé que dans le cadre du polymorphe-V l'hélice gauche était plus stable que l'hélice droite, comme cela sera discuté dans la partie 3 de cette thèse. Le complexe le plus populaire dans cette catégorie est le complexe amylose-iode<sup>81,98-100</sup>, car, comme déjà cité ci-dessus, il permet de connaître la quantité d'amylose native dans un amidon donné<sup>62,63</sup>.

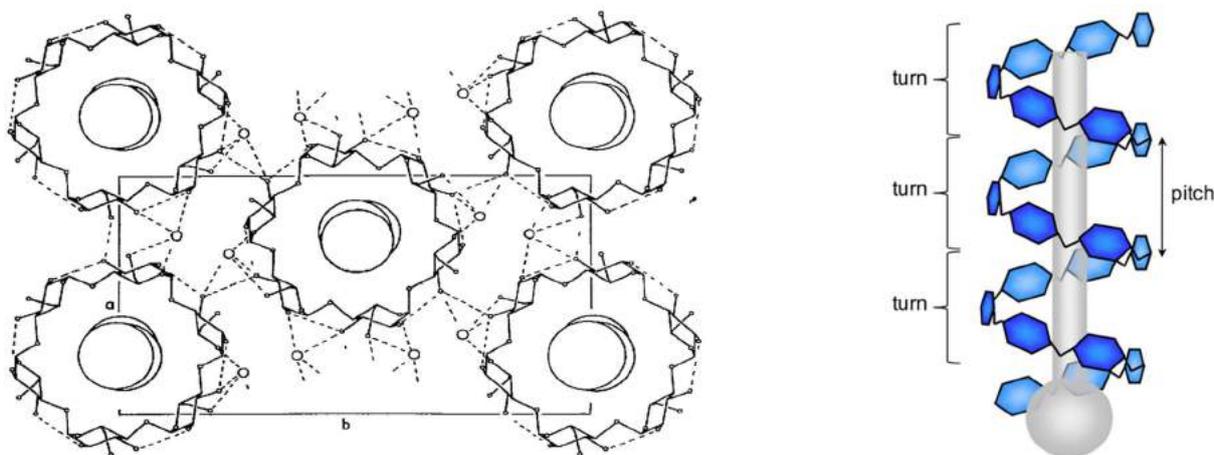


Figure 8: Complexe amylose-iode (gauche), par Bluhm et Zugenmaier (1981). Modèle de complexe d'inclusion amylose-lipide (droite), extrait de Putseys et co-auteurs (2011).

#### e. Polymorphe V : une famille d'allomorphes

Si le polymorphe V désigne une famille de structures se présentant sous la forme d'une hélice simple, il existe en son sein différentes familles d'allomorphes décrivant des empilements cristallins divers et un nombre de résidus glucose par tour d'hélice pouvant varier de 6 à 8. Généralement les différents allomorphes sont désignés comme tels :  $V_{Xi}$ , avec X désignant le nombre de résidus par tour (6, 7 ou 8), et i désignant un profil de diffraction aux rayons X ou électronique particulier.

Les premiers types d'hélices ayant été caractérisées en présence d'iode comprenaient 6 résidus par tour<sup>61,81,98-100</sup>, il s'agissait donc de polymorphes  $V_6$ . En 1944, Bear fut le premier à proposer qu'un nombre supérieur de résidus par tour serait nécessaire pour

former ce type de complexe autour d'alcools branchés et non linéaires<sup>101</sup>, cette observation fut confirmée en 1966 par Yamashita et Hirai<sup>102</sup>. Ces derniers furent aussi parmi les premiers à mentionner qu'une interconversion  $V_7 \rightarrow V_6$  pouvait être observée en plongeant les cristaux dans du méthanol, sans dissolution apparente. Enfin, c'est seulement en 1998 que la présence d'un polymorphe  $V_8$  fut caractérisé en présence d'acide salicylique<sup>103</sup>, avant d'être confirmé en présence d' $\alpha$ -naphthol par Cardoso et co-auteurs en 2007<sup>104</sup>.

A ce jour il semblerait qu'il existe dix familles d'allomorphes différentes, dont cinq ont été caractérisées par Cong Anh Khanh Le durant son travail de thèse au CERMAV<sup>105</sup> (Tableau 2). Après avoir testé 121 agents complexant, 4 allomorphes de type  $V_6$ , 4 de type  $V_7$  et 2 de type  $V_8$  ont été finalement identifiés, et les polymorphes  $V_{6I}$ ,  $V_{6II}$  et  $V_{7II}$  sont ceux qui sont le plus souvent retrouvés. L'eau présente dans chaque type de complexe est très importante pour la stabilité et la cristallinité de ces différents allomorphes.

Tableau 2: Dix familles d'allomorphes telles que définies par Cong Anh Khanh Le (2018).

Allomorph	Name	Space group	Crystal system	Unit cell parameters			Number of helices per unit cell	Interhelical space per helix (nm <sup>3</sup> ) <sup>c</sup>
				<i>a</i> (nm)	<i>b</i> (nm)	<i>c</i> (nm)		
V6 <sub>I</sub>	V6 <sub>I</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	orthorhombic	1.372 ± 0.006	2.376 ± 0.006	0.809 ± 0.008	2	0.126
		<i>P</i> 6 <sub>5</sub> 22	hexagonal	1.372 ± 0.006	1.372 ± 0.006	0.809 ± 0.008	1	
V6 <sub>II</sub>	V6 <sub>II</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	orthorhombic	2.646 ± 0.009	2.705 ± 0.013	0.807 ± 0.009	4	0.254
V6 <sub>III</sub>	V6 <sub>III</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	orthorhombic	1.913 ± 0.006	1.913 ± 0.010	0.814 ± 0.002	2	0.290
V <sub>4-hydroxybenzoic acid</sub> <sup>a</sup>	V6 <sub>IV</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	orthorhombic	1.550 ± 0.001	2.836 ± 0.001	0.790 ± 0.001	2	0.572
V <sub>(-)-borneol</sub> <sup>a</sup>	V7 <sub>I</sub>	<i>P</i> 1	orthorhombic	1.526 ± 0.017	2.643 ± 0.029	0.803 ± 0.008	2	0.200
V7	V7 <sub>II</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	orthorhombic	2.813 ± 0.013	2.973 ± 0.011	0.797 ± 0.001	4	0.258
V <sub>1-naphthol</sub> <sup>a</sup>	V7 <sub>III</sub>	<i>P</i> 1	orthorhombic	1.663 ± 0.001	2.518 ± 0.001	0.856 ± 0.001	2	0.280
V <sub>quinoline</sub> <sup>a</sup>	V7 <sub>IV</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	orthorhombic	2.702 ± 0.001	3.291 ± 0.001	0.786 ± 0.001	4	0.358
V <sub>salicylic acid</sub> <sup>a</sup>	V8 <sub>I</sub>	<i>P</i> 2 <sub>1</sub>	monoclinic <sup>b</sup>	3.245 ± 0.006	3.246 ± 0.005	0.793 ± 0.001	4	0.232
V8	V8 <sub>II</sub>	<i>P</i> 4 <sub>3</sub> 2 <sub>1</sub> 2	tetragonal	2.313 ± 0.006	2.313 ± 0.006	0.790 ± 0.001	2	0.485

<sup>a</sup> new allomorphs

<sup>b</sup>  $\gamma = 116.62 \pm 0.13^\circ$

<sup>c</sup> interhelical space per helix = [unit cell volume -  $N \cdot (\pi \cdot d^2 / 4)$ ]/ $N$ , where  $N$  is the number of helix per unit cell;  $d$  is the external helix diameter and is 1.37, 1.50 and 1.62 nm for V6, V7, and V8 helix, respectively.

Avant de parler des méthodes d'étude de ces polymorphes et des paramètres importants pour leur fabrication, il semble pertinent de discuter d'un dernier aspect : leur classification selon leurs propriétés thermiques. On parle de complexes de type I ou II si la température à laquelle on observe un pic endothermique en DSC est inférieure ou supérieure à 100°C respectivement<sup>1,106,107</sup>. Les complexes de type I

seraient attribués à des phases relativement peu ordonnées, et ceux de type II à un mélange entre phases amorphes et phases cristallines très structurées. Les complexes de type IIa et IIb ont été proposés en 1995 par Karkalas<sup>106</sup>, les seconds présentant la température de dissociation des complexes la plus élevée (121°C).

Il a été proposé récemment la division entre type Ia et Ib, car une température de dissociation en dessous de 90°C a été observée par Kong et co-auteurs<sup>15</sup>. Selon ces derniers, il s'agirait de complexes V ne présentant aucune structure ordonnée dans l'échantillon.

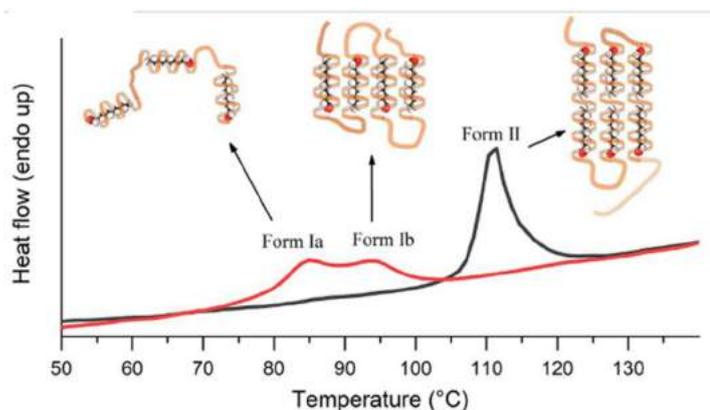


Figure 9: Pics Endomorphes récemment observés par Kong et co-auteurs, et modèles d'organisation leur correspondant.

S'il est important de discuter de cette nomenclature, c'est pour éviter la confusion entre les polymorphes de type I ou II et la numérotation des différents allomorphes (Ex : V<sub>6I</sub> et V<sub>6II</sub> ...). Le type désigne donc une information thermique sur les complexes étudiés, information décorrélée de l'allomorphe considéré.

#### f. Protocoles de fabrication des complexes d'inclusion

Les protocoles de fabrication de ce type de complexes peuvent être séparés en 3 catégories<sup>1,2</sup> : les protocoles classiques, thermomécaniques et enzymatiques.

Les protocoles classiques se déroulent généralement en 3 étapes : solubilisation de l'amylose et de l'agent complexant dans un solvant adapté, montée en température, descente en température contrôlée<sup>2</sup> (Figure 10). Les différents types de solvants rencontrés dans la littérature sont des solutions aqueuses alcalines<sup>106,108</sup>, des mélanges eau/DMSO<sup>17,18,60,107</sup> ou des solutions aqueuses à pH neutre<sup>4,109,110</sup>. Comme précisé précédemment, l'amylose se dissout difficilement dans les solutions aqueuses à PH neutre, rendant les autres types de protocoles classiques plus simples à mettre en place, malgré l'utilisation de solvants toxiques.

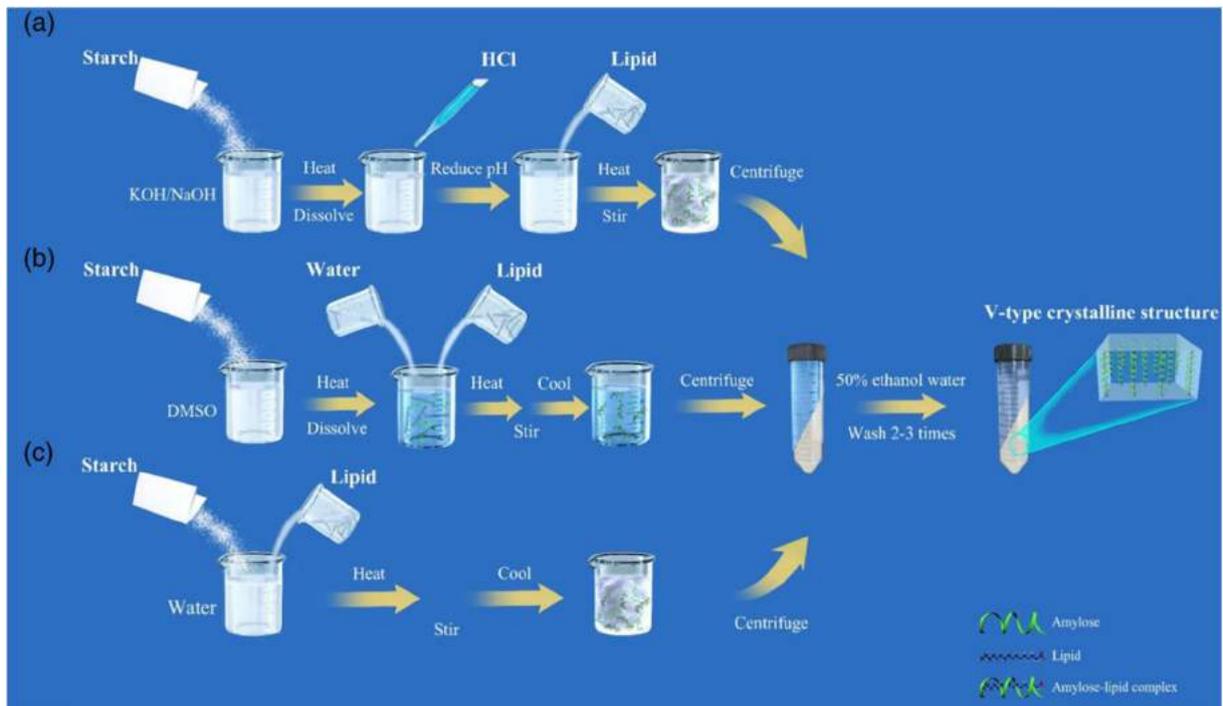


Figure 10: Description générale des protocoles classiques utilisant de l'eau alcaline (a), des mélanges eau/DMSO (b) ou de l'eau neutre (c). Extrait de Wang et co-auteurs (2020).

Les méthodes thermomécaniques permettent de s'abstenir d'utiliser des solvants tels que le DMSO. Depuis 1979 et les travaux de Mercier et co-auteurs, la méthode dite d'extrusion d'amidon est connue pour être capable de générer des polymorphes V<sup>111,112</sup>. Le principe réside dans la destruction des granules d'amidon à l'aide d'une ou plusieurs vis, à haute température, induisant une rétrogradation mécanique et une destruction des structures préétablies au sein de l'amidon (Figure 11). Les paramètres importants pour la formation de polymorphes V par ces méthodes ont été étudiés dans les années 1990 par Bhatnagar et co-auteurs<sup>113-115</sup>. Il en a notamment résulté que la longueur de la vis utilisée influence grandement l'avancée de la gélatinisation et la formation de complexes<sup>115</sup>. Récemment, les travaux de De Pilli et co-auteur permirent de confirmer que lors de l'utilisation de cette méthode, la quantité d'eau et la quantité de lipides présents dans l'amidon constituaient aussi des facteurs importants pour la fabrication de ces complexes<sup>116,117</sup>. Outre la méthode d'extrusion, les méthodes d'homogénéisation<sup>118-120</sup> et de cuisson à la vapeur<sup>121-125</sup> ont été intensément étudiées ces dernières années, car ces méthodes pourraient permettre non seulement de former ces complexes, mais aussi d'être appliquées sur des échelles industrielles. Enfin, il a aussi été démontré qu'il était possible de former ces complexes dans des appareils tels que les Rapid Visco Analyzer<sup>126,127</sup> (RVA) ou des DSC<sup>128,129</sup>, permettant

simultanément leur formation et l'étude de leurs propriétés visqueuses et calorimétriques respectivement.

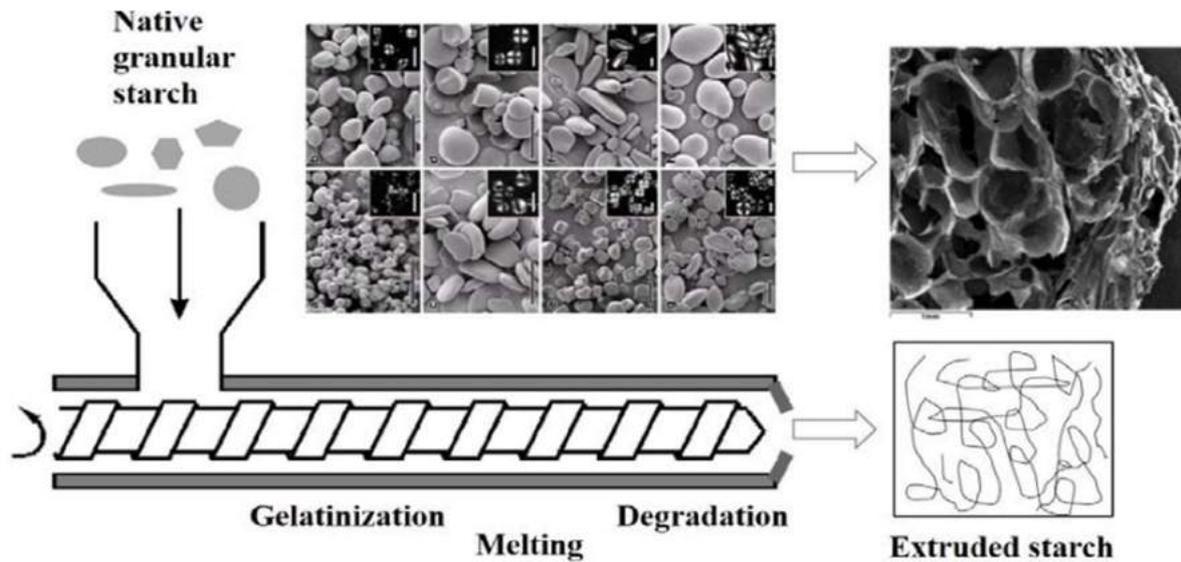


Figure 11: Principe de l'extrusion d'amidon et structures des granules observées avant et après extrusion. Extrait de Ye (2017).

Finalement, les méthodes dites enzymatiques de fabrication de ces complexes se basent sur la fabrication enzymatique d'amylose, décrite précédemment, en présence d'un agent complexant. Kaneko et co-auteurs définirent ainsi le terme de « Vine-twining polymerisation » pour désigner la fabrication de l'hélice V autour de différents polymères<sup>130-132</sup>, grâce à l'utilisation de phosphorylase. Ces méthodes ont pour avantage principal de permettre le contrôle de la taille de l'amylose présente autour des lipides<sup>133,134</sup>. On retrouve aussi le terme de synthèse semi-enzymatique<sup>1,2</sup>, afin de désigner des méthodes similaires, qui diffèrent par les enzymes utilisées. On retrouve ainsi dans la littérature l'utilisation d'amylases<sup>19,135-137</sup> et de pullulanase<sup>138,139</sup> afin de débrancher de l'amidon en vue de la fabrication de ces complexes. On peut aussi citer l'étude de Carciofi et co-auteurs, qui réussirent à produire un amidon contenant exclusivement de l'amylose, et qui prouvèrent qu'il était possible de former des complexes d'inclusion en l'utilisant<sup>29</sup>.

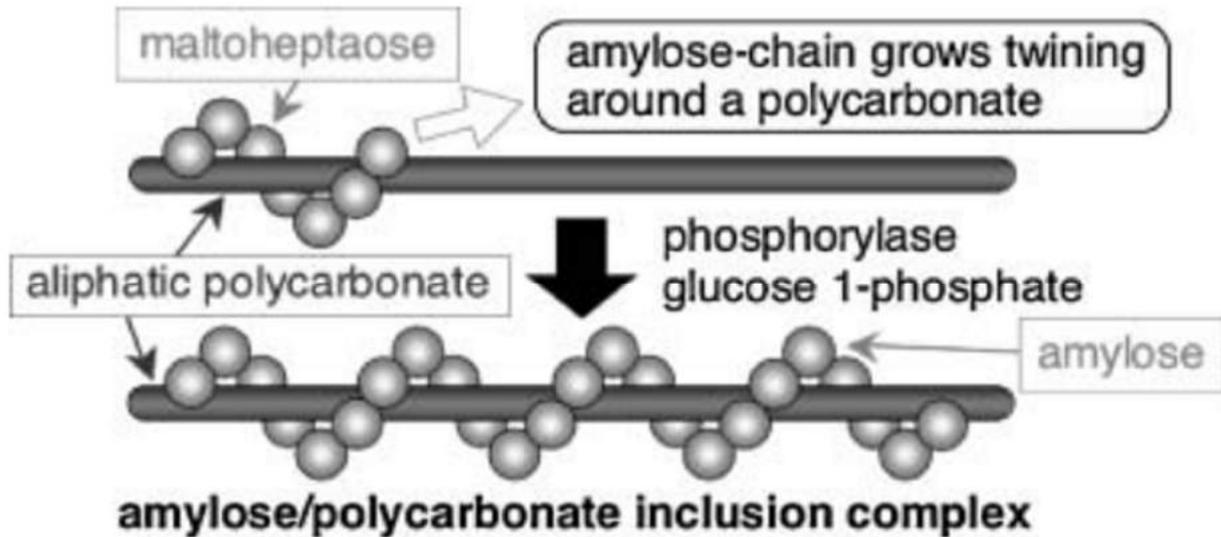


Figure 12: Vine-twining polymerisation tel que défini par Kaneko et co-auteurs (2008).

#### g. Méthodes d'études

Les polymorphes V ont été étudiés en utilisant un grand nombre de méthodes expérimentales, et relativement peu de méthodes théoriques.

Expérimentalement, quatre types de méthodes peuvent être recensées<sup>2,55</sup> : les méthodes d'étude thermiques et rhéologiques, les méthodes de diffraction aux rayons X ou aux neutrons, les méthodes de microscopie et les méthodes de spectroscopie.

La DSC est peut-être la méthode d'étude thermique de ces complexes la plus utilisée à travers l'histoire. Elle est basée sur la mesure d'une différence d'échange de chaleur entre l'échantillon et une référence. Au cours du temps, on augmente la température de façon linéaire, en appliquant un flux de chaleur constant. Puis on trace un thermogramme dont l'abscisse est la température appliquée et l'ordonnée le flux de chaleur mesuré. Chaque pic observé sur le thermogramme est significatif d'un changement organisationnel de l'échantillon (Figure 13)<sup>140</sup>.

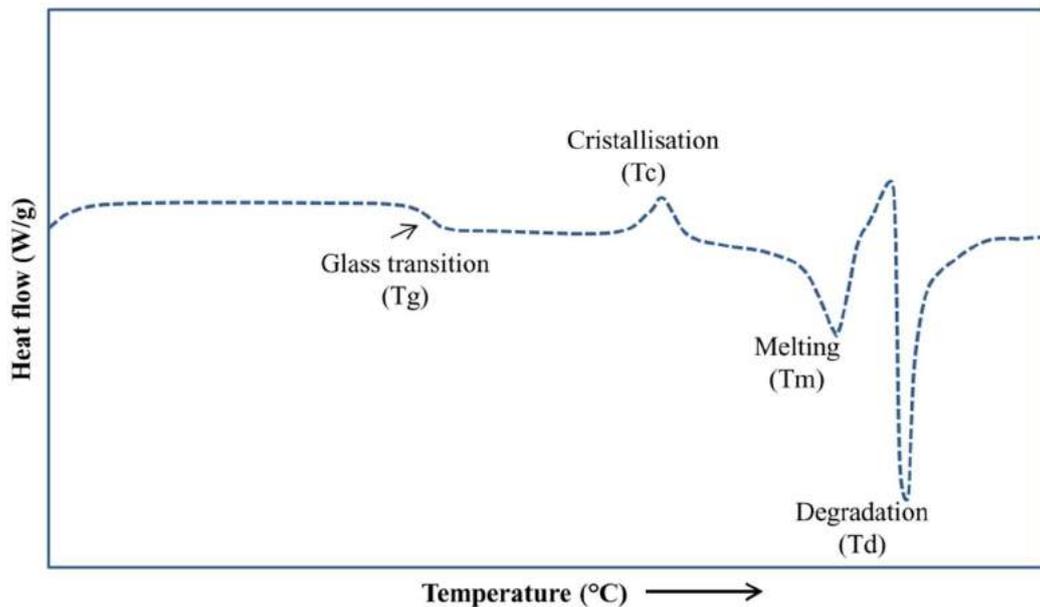


Figure 13: Représentation schématique des transitions observées sur un thermogramme pour un échantillon semi-cristallin. Extrait de Leyva-Porras et al. (2019).

Comme précisé précédemment, les polymorphes V sont classés entre type I et II en fonction de la température à laquelle apparaît le pic endothermique situé aux alentours de 100°C ; plus cette température est élevée, plus la structure de l'échantillon est ordonnée<sup>15,106,107,141</sup>.

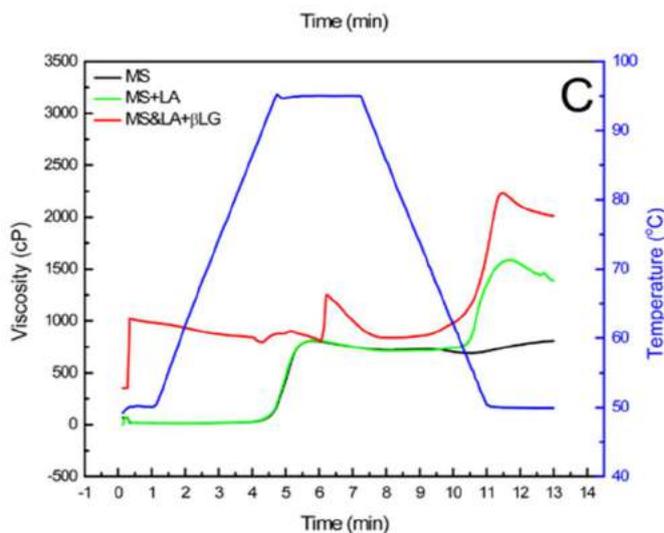


Figure 14: Profil de RVA obtenu sur des échantillons d'amidon de maïs (MS) en présence d'acide laurique (LA) et de  $\beta$ -lactoglobuline ( $\beta$ LG). Extrait de Wang (2017).

De manière similaire, les RVA font aussi subir des changements de température à l'échantillon étudié, mais cette fois-ci c'est la viscosité du système qui est mesurée (Figure 14)<sup>59</sup>. Le premier pic de viscosité durant la période à haute température est en général attribué au relargage d'amylose présent dans l'amidon induisant la gélatinisation du milieu. Le second, apparaissant lors de la descente en température est quant à lui induit

par la formation de complexes d'inclusions durant la rétrogradation de l'amidon<sup>126</sup>.

Si les méthodes thermiques et rhéologiques permettent de caractériser la texture et l'ordre des échantillons étudiés, les méthodes de diffraction et de microscopie permettent d'établir de façon complémentaire une structure en 3 dimensions de ces complexes, en apportant des informations structurales à des échelles différentes (Figure 15)<sup>142</sup>. Le profil de diffraction aux rayons X (XRD) du polymorphe V sous forme de poudre est connu depuis 1930<sup>94,95</sup>, et par la suite il fut proposé par Rundle et co-auteurs qu'il était généré par une structuration en hélice simple de l'amylose<sup>98,99</sup>. Il a été démontré que le taux d'hydratation influence le profil de diffraction, celui-ci étant différent quand l'échantillon est anhydre<sup>114,143</sup> ou hydraté<sup>144</sup>. L'utilisation de la XRD permet d'établir la structure tridimensionnelle de ces complexes et d'accéder à certaines informations structurales d'intérêt pour leur caractérisation. Ainsi, on sait que le polymorphe-V présente un pas d'hélice d'environ 8 Å environ<sup>145</sup>, et que le couple d'angles  $\{\varphi, \psi\}$  se situent aux alentours de  $\{105^\circ, 115^\circ\}$ <sup>82</sup>. La diffraction aux rayons X aux petits angles (SAXS) a aussi été utilisée récemment afin de caractériser la structure à l'échelle nanométrique. Elle permet notamment de montrer que ces complexes alternaient entre phases amorphes et cristallines<sup>146</sup> et que les complexes de type II présentaient une structure en lamelles plus fines et plus organisées que ceux de type I<sup>141</sup>. Ces méthodes ont fait l'objet d'une revue récente par Blazek et Gilbert<sup>142</sup>.

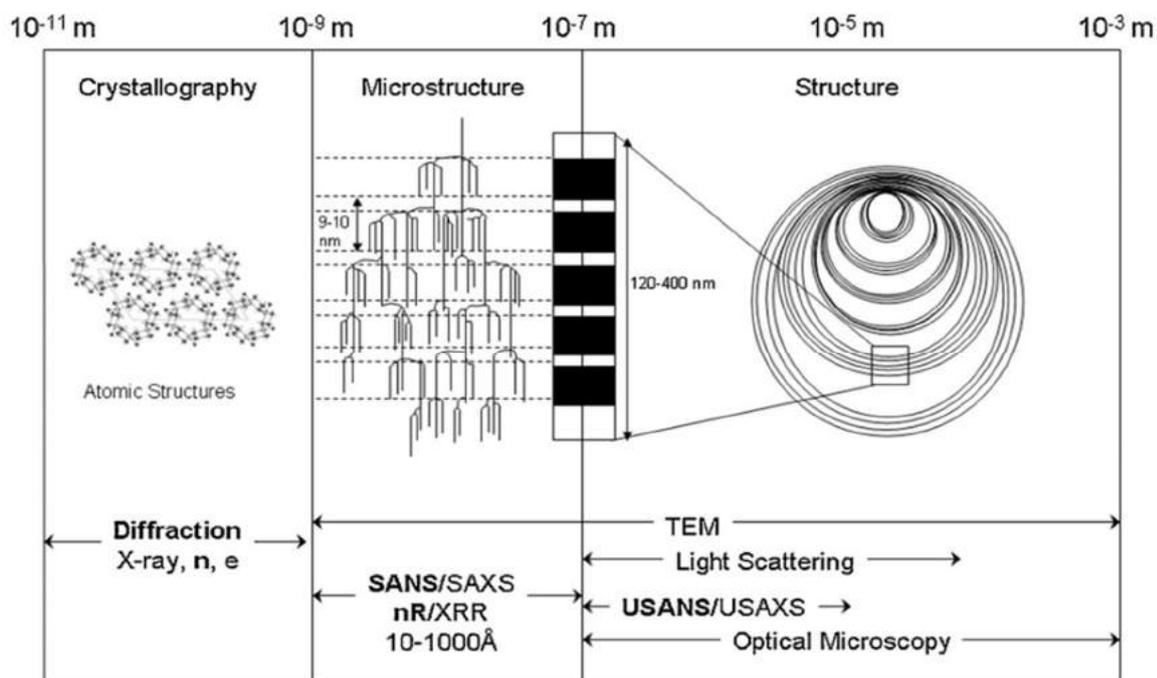


Figure 15: Différentes échelles de structures observables pour la diffraction aux rayons X, neutrons et électrons (X-ray, n, e), la diffraction aux rayons X et aux neutrons aux petits angles (SAXS, SANS) et la microscopie électronique. Extrait de Blazek (2011).

Les méthodes de microscopie telles que la microscopie à transmission électronique (TEM), la microscopie à balayage électronique (SEM) et la microscopie à force atomique (AFM) et, dans une moindre mesure, la microscopie optique, ont été utilisées afin de caractériser les morphologies de ces complexes à l'échelle microscopique. Les premières observations, obtenues dans les années 1940 par microscopie optique, révélaient des structures microscopiques en forme de fleurs pour le complexe formé en présence d'iode<sup>147</sup> ou de rectangles pour celui formé en présence de butanol<sup>148</sup>. C'est seulement dans les années 60 que les premières images issues de TEM sont apparues, révélant l'organisation en structures lamellaires de ces complexes<sup>149,150</sup>. Cette méthode est aujourd'hui régulièrement utilisée en combinaison avec les méthodes de diffraction afin de caractériser la morphologie des complexes obtenus<sup>12,104,105,151</sup>. Récemment, la microscopie confocale à balayage laser a été utilisée en combinaison avec la SEM afin de caractériser la distribution hétérogène de lipides au sein de ce type de complexes<sup>152</sup>. Concernant l'AFM, il s'agit de la méthode utilisée à la fin des années 90 afin de mettre en évidence l'existence des blocklets<sup>34</sup> et elle fut récemment utilisée pour discuter de l'importance de l'amylose pour la formation de gels lors du processus de rétrogradation de l'amidon<sup>57</sup>.

En combinaison avec ces méthodes, les méthodes de spectroscopie sont souvent utilisées afin de caractériser ces complexes. Nous pouvons notamment citer l'utilisation de spectroscopie infrarouge<sup>152</sup> ou Raman. Cependant, nous allons nous attarder sur la RMN, notamment celle effectuée sur les noyaux <sup>13</sup>C à l'état solide. Malgré une utilisation générale de la RMN de plus en plus fréquente dans l'étude des différentes structures de l'amidon, la RMN à l'état solide reste peu utilisée ces dernières années dans le cadre de l'étude structurale de ces polymorphes<sup>153</sup>. Le signal RMN des différentes structures de l'amylose est pourtant connu et il est possible de les discriminer en se basant sur ce dernier (Tableau 3)<sup>154</sup>.

Tableau 3: Déplacement chimique <sup>13</sup>C de différents polysaccharides composés de résidus glucoses. Les échantillons qui ne précise pas l'état de l'échantillon sont à l'état solide. Extrait de Gidley (1988).

material <sup>a</sup>	C-1	C-4	C-3	C-2,5	C-6
V <sub>6</sub> amylose (sodium palmitate complex)	103.6	82.2	75.9	72.7	61.0
V <sub>6</sub> amylose (hexanoic acid complex)	103.9	82.7	75.9	72.9	61.9
V <sub>6</sub> amylose hydrate	103.5	83.2	75.7	73.1	62.2
V <sub>6</sub> amylose (butanol complex)	104.1	83.5	75.8	73.5	62.0
V <sub>7</sub> amylose ( <i>tert</i> -butyl alcohol complex)	103.9	83.2	75.7	72.9	61.9
V <sub>8</sub> amylose (1-naphthol complex)	104.9	84.1	75.4	73.4	61.7
α-cyclodextrin hydrate	102.1 <sup>b,c</sup> (102.9) <sup>e</sup>	81.4 <sup>b,d</sup> (82.1) <sup>f</sup>	71-76	71-76	61.6 <sup>b</sup>
α-cyclodextrin (aqueous soln)	102.6	82.4	74.5	72-73	61.7
β-cyclodextrin hydrate	103.2 <sup>b,g</sup>	82.2 <sup>b,h</sup>	71-76	71-76	61.1 <sup>b</sup>
β-cyclodextrin (aqueous soln)	103.1	82.3	74.3	72-73	61.7
A-type α-(1→4) glucan	100.4 <sup>b,i</sup>	76.0	70-75	70-75	62.1
B-type α-(1→4) glucan	100.4 <sup>b,j</sup>	76.2	70-75	70-75	62.3
amylose (aqueous soln)	100.9	78.6	74.6	72-73	61.9

Les déplacements chimiques des carbones C<sub>1</sub> (entre 99 et 105 ppm) et C<sub>4</sub> (entre 76 et 82 ppm) des monomères de glucose forment une sonde particulièrement pratique pour caractériser la formation de polymorphe V, influencés principalement par les changements des angles dièdres  $\{\varphi, \psi\}$  formant la liaison glycosidique (Figure 1). Il a été montré que le déplacement chimique des lipides était lui-même affecté par la complexation<sup>18</sup>, et que le signal RMN obtenu sur des cristaux de polymorphes V était différent en fonction de l'allomorphe considéré<sup>105</sup>. Une des caractéristiques de la RMN, qui la rend attractive pour l'étude de complexes d'amylose, est qu'elle permet d'étudier des échantillons solides de composition hétérogène. Le comportement d'un échantillon d'amylose contenant différentes concentrations d'acide palmitique a notamment été étudié grâce à cette méthode par P. Le-bail et co-auteurs en 2000<sup>18</sup>. Il fut démontré que l'augmentation de la concentration initiale en lipides dans le milieu engendrait la formation d'un plus grand pourcentage de polymorphe V dans l'échantillon.

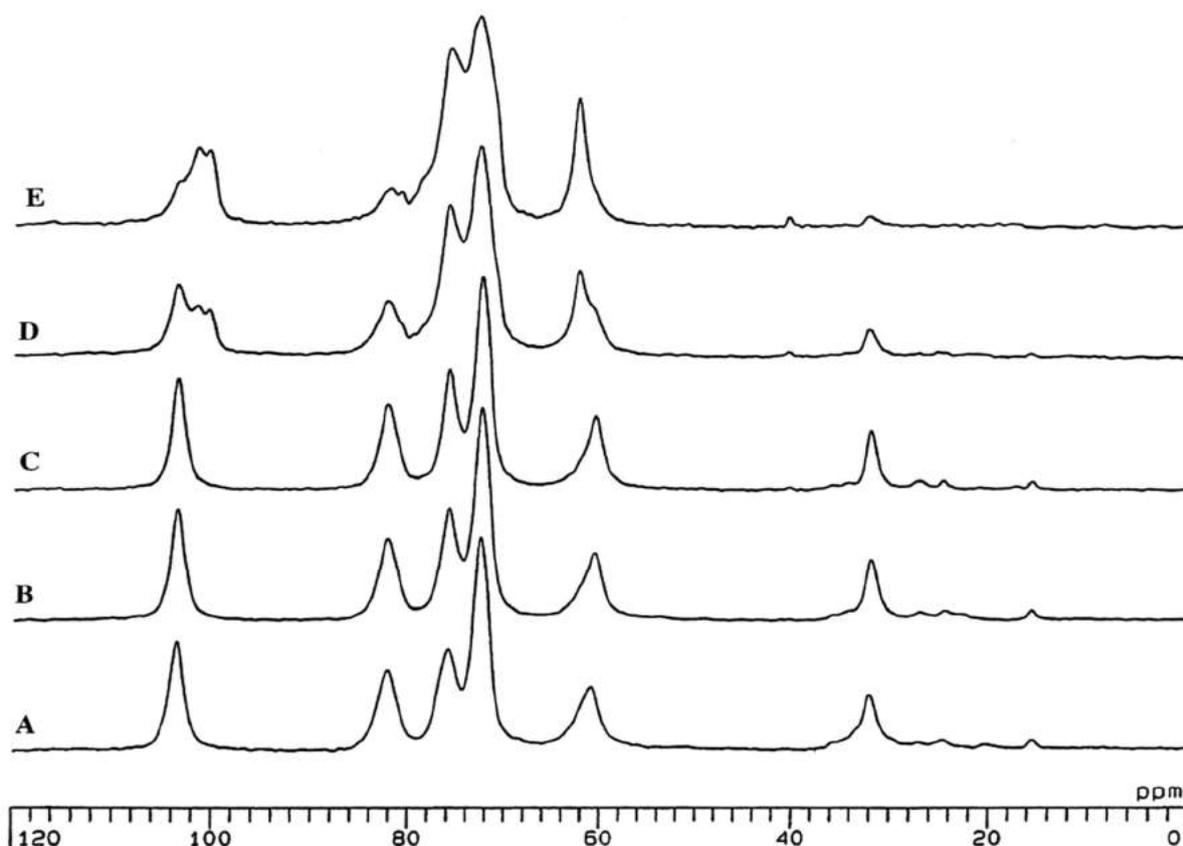


Figure 16: Evolution du spectre RMN 13C d'un échantillon d'amylose avec une concentration initiale en acide palmitique de 10%(A), 7%(B), 5%(C), 2%(D) et 1%(E). Extrait de Le-bail (2000).

La RMN est donc une méthode pratique pour suivre la quantité de polymorphe V au sein d'amidons d'origines végétales différentes<sup>155</sup>, ou d'amidons ayant subi des cuissons diverses<sup>156</sup>.

Comme nous l'avons précisé dans le sous-paragraphe a, c'est aussi grâce à l'utilisation de cette méthode qu'il a pu être démontré que le polymorphe B contenait en réalité deux types de résidus et donc qu'il existait deux environnements chimiques différents au sein des cristaux de ce polymorphe<sup>72</sup>. Nous rediscuterons de ce fait dans la troisième partie de ce manuscrit, lors de l'établissement des stratégies de calculs.

L'utilisation de ces quatre différentes méthodes expérimentales a permis de mettre en évidence les différents paramètres importants pour la fabrication de ces complexes qui sont : la source d'amidon, le type de ligand utilisé et le protocole de préparation<sup>2</sup>.

Si de nombreuses études expérimentales ont été publiées, de plus en plus de travaux théoriques sont réalisés sur ces complexes, principalement par le biais de la dynamique moléculaire (MD)<sup>78,157-161</sup>. Aujourd'hui, relativement peu de publications présentent uniquement des résultats issus de calculs théoriques sur des complexes de type polymorphe-V. Parmi ces derniers, on peut notamment citer les travaux de Lopez et co-auteurs, qui furent les premiers à observer le repliement d'une chaîne d'amylose autour d'un acide palmitique, en utilisant la dynamique moléculaire<sup>161</sup>. Dans cette publication, ils démontrèrent notamment que le champ de force utilisé (GROMOS) nécessitait une modification afin de permettre une meilleure reproduction des résultats expérimentaux obtenus par AFM. Nous rediscuterons de ces modifications dans la troisième partie de ce manuscrit, en même temps que celles que nous avons effectuées sur le champ de force que nous utilisons (GLYCAM06). Plus récemment, Cheng et co-auteurs étudièrent le comportement d'une chaîne d'amylose contenant 26 résidus glucose en présence d'une ou plusieurs molécules d'acide linoléique<sup>158,159</sup>. Ces travaux démontrèrent notamment que la conformation hélicoïdale de l'amylose était bien présente autour de la partie hydrophobe du lipide pendant que le reste de la chaîne était non structuré. De plus l'addition de molécules d'acide linoléique implique l'augmentation du nombre de résidus structurés en hélice V, et les structures observées laissent supposer qu'il est possible de complexer plusieurs chaînes lipidiques côte à côte au sein du polysaccharide (Figure 17). Nous reviendrons sur ce point dans le chapitre 2 de la partie 3 de ce manuscrit.

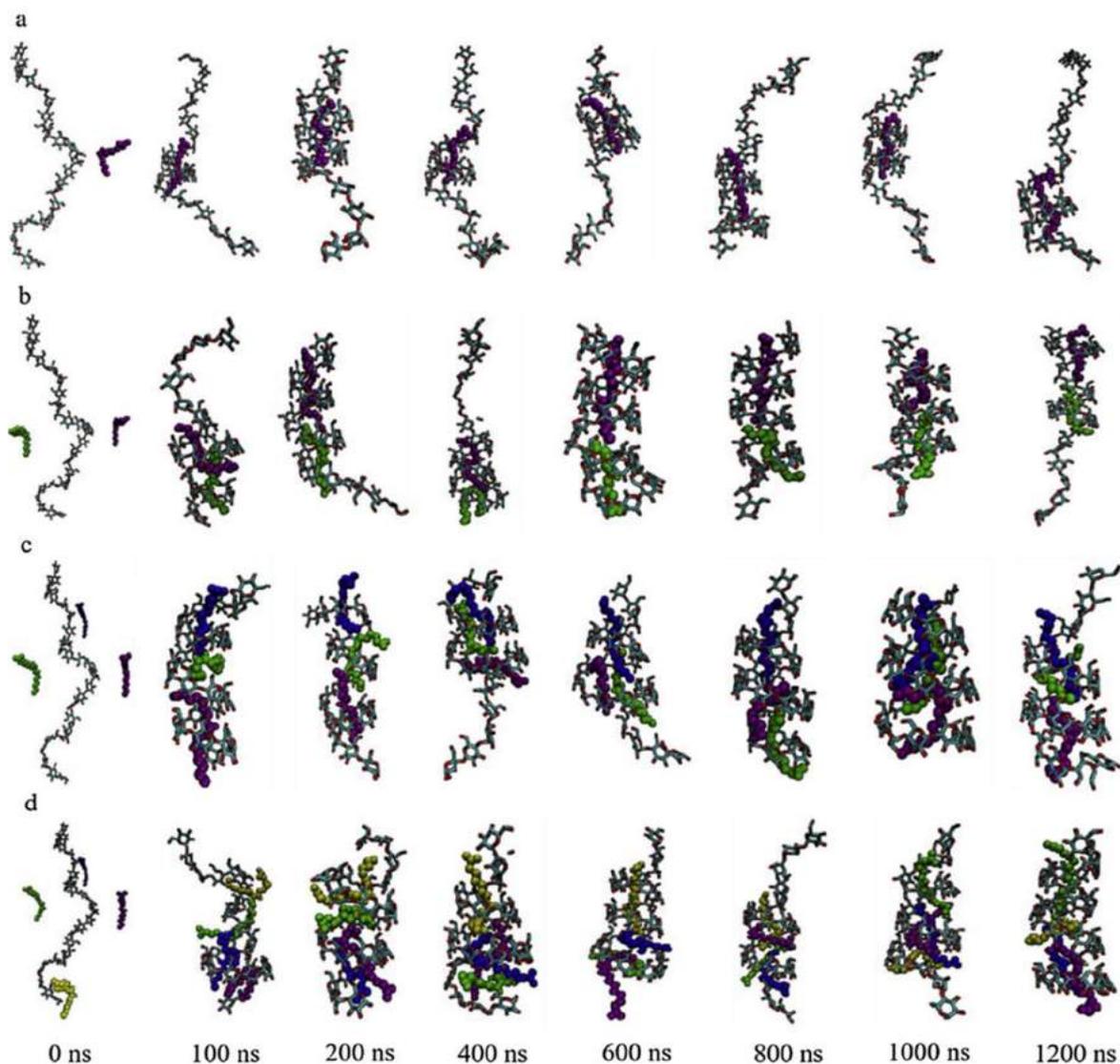


Figure 17: Structures obtenues lors de la complexation de l'amylose en présence d'une, deux trois ou quatre molécules d'acide linoléique. Extrait de Cheng 2019.

S'il existe relativement peu d'études entièrement théoriques, l'utilisation de la MD est devenue fréquente dans les articles étudiant les polymorphes  $V^{160,162-167}$ . Gao et co-auteurs ont étudié le mécanisme de relargage de certains composés bioactifs, encapsulés dans de l'amylose propionylée, à travers la barrière hématoencéphalique <sup>160</sup>. Ils ont notamment pu démontrer que les molécules de 1-palmitoyl-2-oleoyl phosphatidyléthanolamine (POPE) permettaient de déclencher le relargage de la molécule bioactive en formant des liaisons hydrogènes avec l'amylose, déstabilisant de ce fait les liaisons hydrogènes intramoléculaires au sein du polysaccharide. En parallèle, des expériences impliquant des barrières hématoencéphaliques artificielles ont démontré que les molécules bioactives étaient mieux délivrées à travers ces membranes lorsqu'elles étaient encapsulées dans l'amylose fonctionnalisée. Ils

finirent par confirmer cet effet sur des modèles animaux (souris et lapin), prouvant ainsi que l'encapsulation par l'amylose de composés bioactifs améliorerait leurs caractéristiques pharmacologiques.

Finalement, les divers travaux utilisant la mécanique quantique (QM) pour étudier les systèmes de types polysaccharides se sont résumés à l'étude de petits systèmes<sup>168</sup>, tels que des oligosaccharides (allant de 2 à une vingtaine de résidus)<sup>169-171</sup> ou des cyclodextrines<sup>172,173</sup>. Les calculs réalisés par Schnupf et co-auteurs sur la cyclodextrine de 26 résidus (principalement des optimisations de géométries à différents niveaux de calculs) sont, à ce jour, ceux qui se rapprochent le plus des calculs de mécanique quantique sur des complexes d'inclusion similaires au polymorphe V<sup>172</sup>. On peut aussi relever les calculs de Okuda et co-auteurs, qui réussirent à reproduire les spectres IR du complexe iode- $\alpha$ -cyclodextrine<sup>173</sup>.

L'utilisation de méthodes théoriques permet l'obtention de nouvelles informations sur la dynamique intrinsèque de ces complexes. Il s'agit cependant d'un domaine encore émergent, qui nécessite l'établissement d'une stratégie de calcul solide afin d'obtenir des résultats cohérents avec ceux obtenus expérimentalement. Nous reviendrons sur ce sujet dans le premier chapitre de ce manuscrit.

## 2. Polysaccharides et lipides bactériens

L'un des objectifs de ce travail de thèse aura été la caractérisation d'interactions entre polysaccharides et lipides de l'enveloppe de *Mycobacterium tuberculosis* (MTB), agent étiologique responsable de la tuberculose. Cette maladie est encore aujourd'hui extrêmement répandue et reste l'une des 10 premières causes de mortalité dans le monde. En 2018, plus de 10 millions de nouveaux cas ont été recensés par l'OMS, ainsi que plus de 1.2 millions de morts, dont la majorité est située en Afrique centrale et en Asie du Sud<sup>174</sup> (Figure 18). Environ un tiers de la population mondiale serait infectée, sans forcément déclencher la maladie.

### Estimated TB incidence rates, 2018

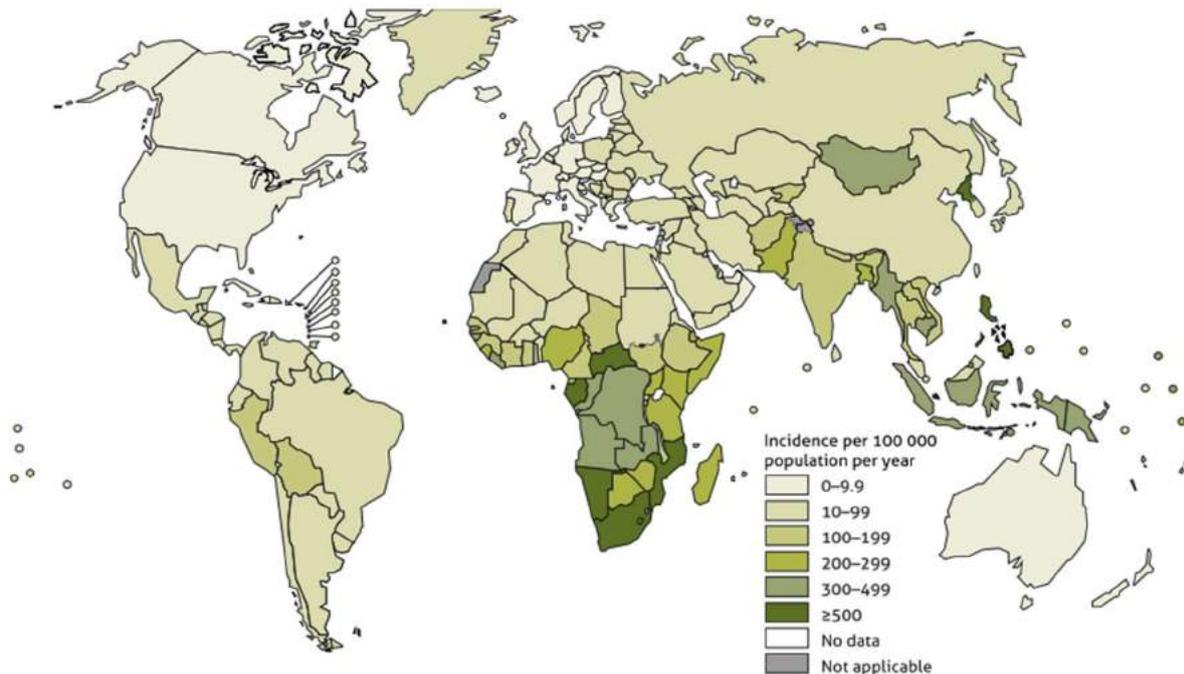


Figure 18: Taux d'incidence mondiale de la Tuberculose (WHO, Global Tuberculosis Report 2019)

Le traitement de la tuberculose consiste en une chimiothérapie durant entre 3 et 6 mois et comprenant 4 antibiotiques : l'isoniazide<sup>175</sup>, la rifampicine<sup>176</sup>, le pyrazinamide<sup>177</sup> et l'éthambutol<sup>178</sup>. La plus grande inquiétude concernant cette maladie réside dans l'apparition de souches résistantes aux antibiotiques utilisés. Deux sortes de tuberculoses résistantes sont définies : celle sur laquelle la rifampicine et l'isoniazide n'ont aucun effet, appelée « Multidrug resistant » (TB MDR) et celle sur laquelle au moins un antibiotique de seconde ligne est inefficace, appelée « Extensively drug resistant » (TB XDR).

MTB est un pathogène faisant partie de la famille des mycobactéries, caractérisées par leur enveloppe très complexe, particulièrement imperméable et hydrophobe du fait de leur haute teneur en lipides (60% de la masse sèche du bacille). Cet édifice commun à la majorité des différentes espèces de mycobactéries, est formé de quatre couches d'épaisseurs et de compositions différentes<sup>179,180</sup> (Figure 19): La membrane plasmique, le périplasme, la paroi ou « cell wall » et la capsule.

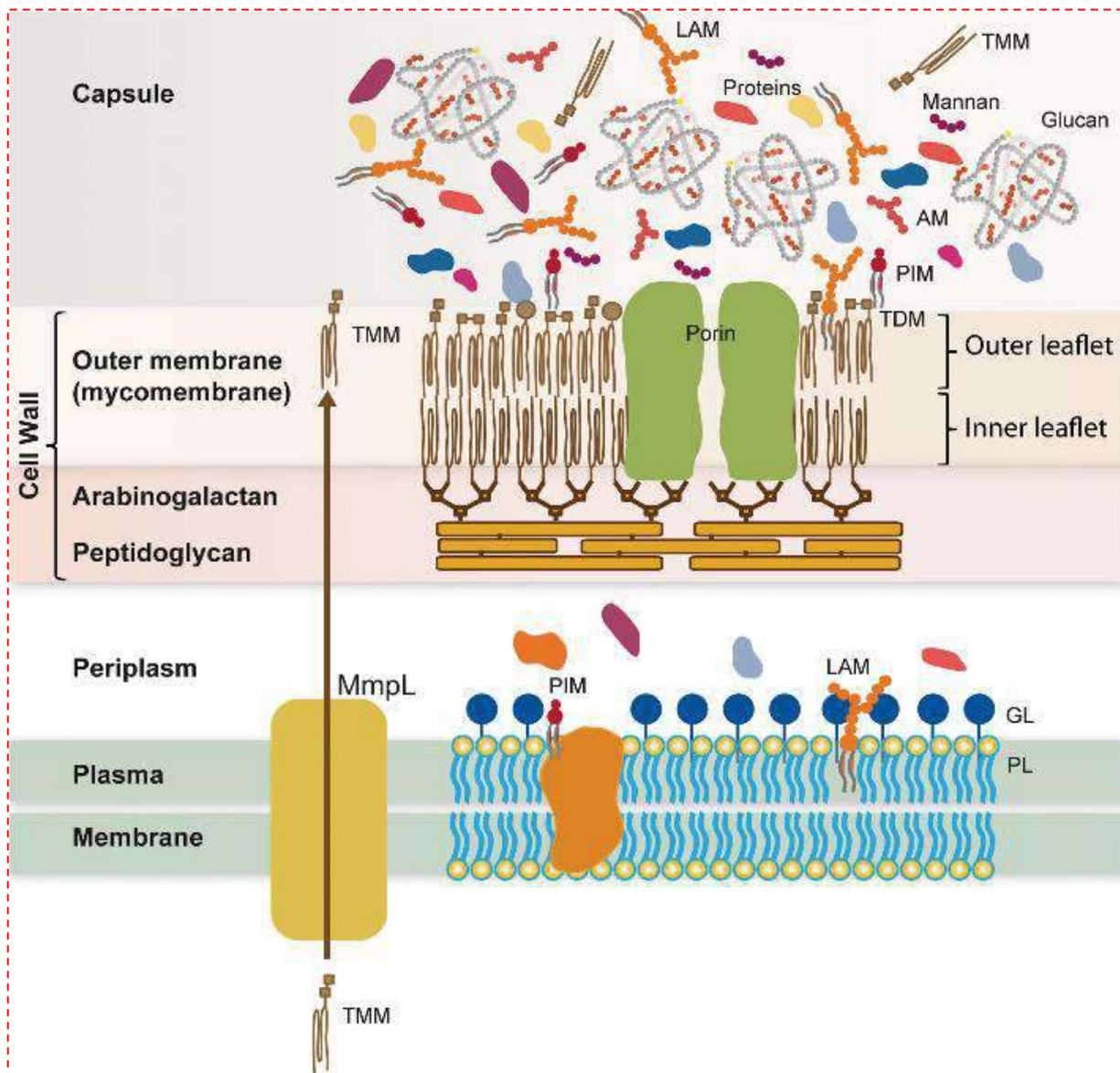


Figure 19: Représentation schématique de l'enveloppe de MTB. TMM: Monomycolate de tréhalose. PL: Phospholipides. GL: Glycolipides. PIM: Phosphatidylinositol mannoside. LAM: Lipoarabinomannane TDM: dimycolate de Tréhalose. AM: Arabinomannane.

Dans ce travail de thèse, nous nous sommes intéressés tout particulièrement à la capsule, partie la plus externe et interface directe entre la bactérie et l'hôte. C'est dans les années 90 qu'il a été montré<sup>8,9</sup> qu'une large partie de la capsule est relarguée dans le milieu environnant tandis qu'une autre semble organisée autour de la bactérie. Elle est constituée majoritairement d'un polysaccharide de haut poids moléculaire appelé  $\alpha$ -glucane mais aussi, en faible quantité (environ 2%) de certains lipides retrouvés dans d'autres parties de l'enveloppe, tels que les dimycolates de tréhalose<sup>181</sup>(DMT) et les dimycocérosate de phitiocérol<sup>182</sup> (DIM) (Figure 20). Il a

notamment été prouvé que ces trois molécules présentaient un impact sur le système immunitaire lors de l'infection causée par MTB<sup>183-188</sup>.

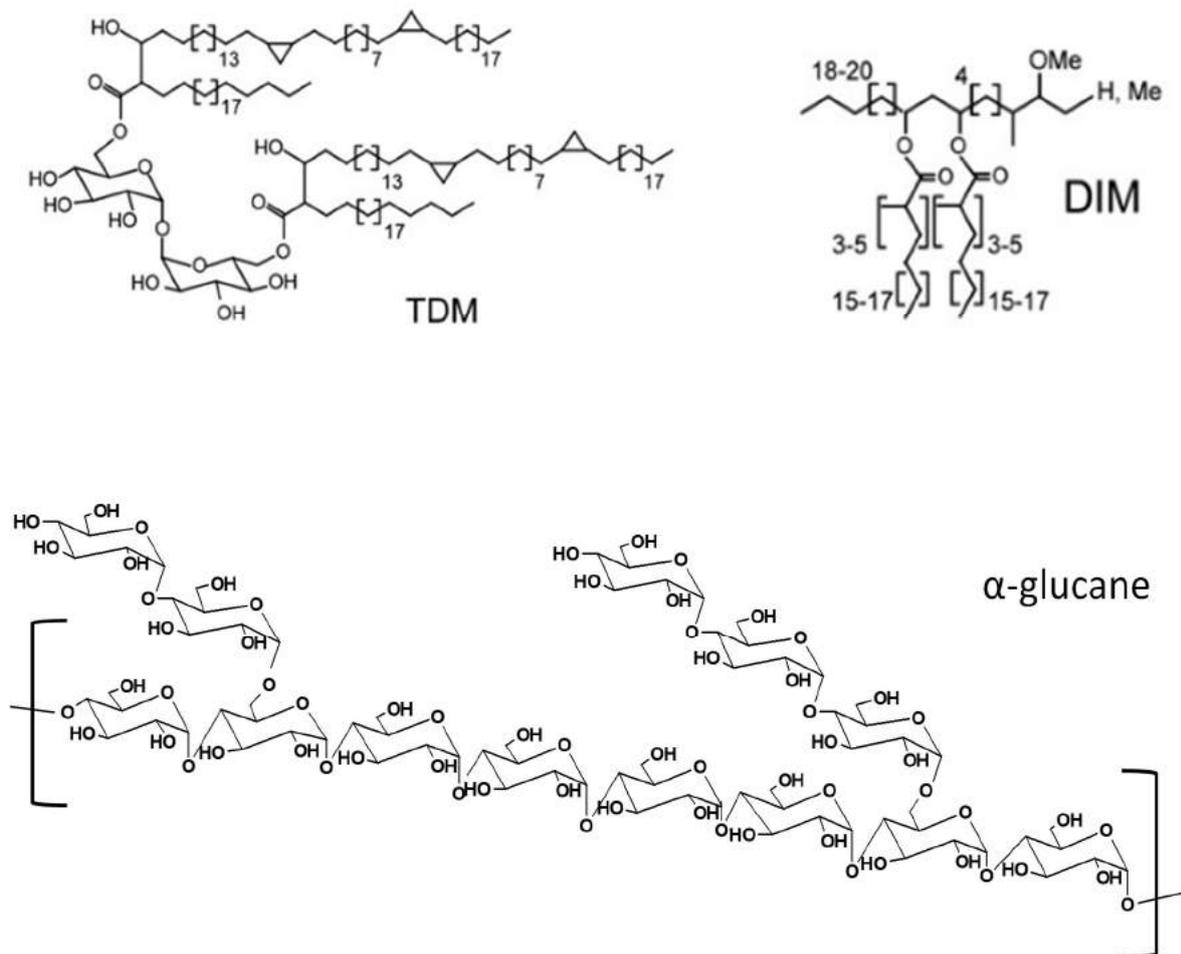


Figure 20: Structures des dimycolates de tréhalose (TDM), des dimycocérosates de phthiocérol (DIM). Modèle de structure possible de l' $\alpha$ -glucane.

Les structures primaires de ces lipides sont bien connues. En revanche, toutes les informations ultra-structurales concernant le polysaccharide ne sont pas encore établies. Il est constitué de résidus glucosyls liés en  $\alpha$ -(1- $\rightarrow$ 4) et certains résidus sont ramifiés en position 6<sup>8</sup>. Sa structure est donc proche de celles de l'amylopectine et du glycogène. Son poids moléculaire a été estimé de  $12 \cdot 10^6$  Da et le rayon moyen des particules observées en diffusion dynamique de la lumière (DLS) est de l'ordre de 28 nm<sup>189</sup>. Les analyses par RMN du proton sur l' $\alpha$ -glucane entier, et par spectrométrie de masse après dégradation des liaisons au niveau des branchements montrent qu'il possède un très haut taux de ramification et des chaînes contenant 2 à 9 résidus glucoses<sup>189,190</sup>. Cependant, ces approches donnent des valeurs moyennes, et n'ont pas encore permis d'étudier des domaines hétérogènes du polysaccharide, qui pourraient interagir avec les lipides minoritaires de la capsule.

Si la quantité de capsule relarguée varie d'une espèce à l'autre<sup>191</sup>, il persiste quand même une couche structurée qui semble similaire quelle que soit l'espèce étudiée<sup>7</sup>, qu'elle soit pathogène ou non. Nous savons que cette couche particulière n'est pas reliée de façon covalente à la mycomembrane et aux lipides, mais son organisation microscopique et les interactions moléculaires impliquées dans sa structuration restent une énigme. De plus, il s'agit finalement de la première couche de contact entre le bacille, pathogène, et les cellules de l'hôte, Or les lipides, cités ci-dessus jouent un rôle clé dans la virulence. La question de l'organisation entre la matrice constituée par l' $\alpha$ -glucane et les lipides capsulaires présentés à l'hôte doit donc être élucidée.

### 3. Objectifs de la thèse

L'objectif de ce travail de thèse est donc d'étudier les interactions entre différents types de polysaccharides et lipides afin d'obtenir plus d'informations sur leur structuration. La littérature nous offre de nombreuses connaissances sur les complexes d'inclusion réalisés avec de l'amylose en présence de lipides ne contenant qu'une chaîne lipidique, mais relativement peu sur les interactions avec des lipides en contenant plusieurs. A ce jour, très peu d'études ont permis la caractérisation des interactions entre polysaccharides branchés et lipides contenant plusieurs chaînes lipidiques.

Pour réaliser cette étude, il était nécessaire d'utiliser des méthodes transférables entre les systèmes de type amylose-lipide et ceux de type polysaccharides branchés-lipides mycobactériens. Notre choix s'est arrêté sur la RMN à l'état solide, car elle permet d'étudier des systèmes qui vont inévitablement s'avérer hétérogènes, non cristallins et dont le poids moléculaire ne permet pas l'étude à l'état liquide. Comme nous l'avons précisé, la RMN  $^{13}\text{C}$  à l'état solide a prouvé qu'elle était un outil de choix pour caractériser différentes structures tridimensionnelles de polysaccharides tels que l'amylose.

Il nous faut cependant rappeler que les noyaux  $^{13}\text{C}$  ne forment qu'1% des noyaux de carbones présents dans les échantillons. De par la difficulté et le prix de production d'un échantillon d' $\alpha$ -glucane marqué au carbone 13, il nous était impossible d'utiliser des spectres RMN multidimensionnels afin d'établir une structure tridimensionnelle de nos échantillons. Seul des spectres à une dimension sont accessibles en des temps

raisonnables sur des échantillons dont les noyaux  $^{13}\text{C}$  sont en abondance naturelle. De ce fait, nous avons utilisé des méthodes de modélisation afin de palier à ce problème, notamment la dynamique moléculaire (MD). La MD permet de modéliser le comportement de ces molécules au cours du temps et d'établir des structures tridimensionnelles des complexes étudiés. Enfin, afin de vérifier que les structures générées par MD peuvent être représentatives de celles générant les signaux RMN expérimentaux, des calculs de chimie quantique ont été mis en place afin d'accéder aux spectres RMN de ces dernières.

Le corps du travail réalisé durant ces trois ans de thèse s'agence de ce fait en trois parties :

I. **Etablissement de la stratégie d'étude théorique de complexes d'amylose :**

Au commencement de ce travail, les travaux théoriques disponibles sur les structures de l'amylose étaient relativement peu nombreux. Il était donc important d'établir une stratégie d'étude théorique solide. Le premier chapitre de ce manuscrit décrit la mise en place d'une étude théorique utilisant dynamique moléculaire (MD) et calculs de RMN par méthode de chimie quantique (DFT) pour caractériser des complexes d'inclusion amylose – acide palmitique. Cette étude a notamment permis de démontrer l'importance de la longueur du polysaccharide pour la stabilisation du polymorphe V, et l'importance de la prise en compte des effets dynamiques pour le calcul de paramètres RMN de ces complexes.

Dans la suite de ce premier chapitre seront aussi décrits les calculs de RMN réalisés sur des complexes d'amylose de type polymorphe B. Nous avons pu mettre en évidence la versatilité du réseau de liaisons hydrogène présent dans ces complexes et de son impact sur le spectre RMN obtenu. De plus nous avons pu démontrer que pour caractériser deux résidus différents au sein de ce polymorphe, il était nécessaire de prendre en compte le packing cristallin de l'échantillon.

II. **Etude expérimentale et théorique de complexes d'inclusion amylose-lipides contenant plusieurs chaînes hydrophobes :**

A ce jour, nous disposons d'un nombre d'informations important concernant les complexes

d'amylose en présence de diverses molécules telles que des acides gras ou de petites molécules hydrophobes. A notre connaissance il n'existe toujours pas de preuve formelle que des complexes d'inclusion puissent être formés avec des lipides classiques présents dans les membranes biologiques, tels que des phospholipides. Particulièrement, il est admis que les lipides contenant plusieurs chaînes hydrophobes reliées par une tête polaire ne puissent pas former de complexes d'inclusion. Dans le second chapitre de ce manuscrit la caractérisation formelle expérimentale et théorique de complexes d'inclusion entre des lipides présentant des chaînes multiples sera présentée. Il s'agit de la première caractérisation formelle de ce type de complexes, et en parallèle, un modèle expliquant les mécanismes en compétition lors du protocole de fabrication sera présenté concernant les échantillons contenant des acides gras et des phospholipides. Un lien sera notamment fait entre la structure primaire des lipides étudiés, leur concentration et leur capacité à induire de la structuration en polymorphe V dans l'échantillon. Nous apportons aussi la preuve qu'il est possible d'observer ce type de complexes en présence de lipides issus de mycobactéries et que ces derniers sont en accord avec le mécanisme de compétition décrit en présence de phospholipides et acides gras.

- III. **Etude expérimentale et théorique de d'interactions en polysaccharides ramifiés et lipides divers:** Le troisième chapitre présentera les travaux réalisés sur des échantillons contenant des polysaccharides ramifiés en présence de lipides classiques ou issus de corynébactéries et mycobactéries. Un modèle de polysaccharide branché a été développé théoriquement afin d'étudier la structure intrinsèque obtenue en présence de divers lipides. Ces dernières n'ont pour le moment pas encore été observées en présence de polysaccharides ramifiés, que ce soit théoriquement ou expérimentalement. Il a cependant été possible d'observer dans nos résultats théorique une structuration locale du polysaccharide, ce qui laisse entendre que les lipides pourraient être impliqués dans la structure de la capsule de MTB.

## Références

- (1) Obiro, W. C.; Sinha Ray, S.; Emmambux, M. N. V-Amylose Structural Characteristics, Methods of Preparation, Significance, and Potential Applications. *Food Reviews International*. 2012. <https://doi.org/10.1080/87559129.2012.660718>.
- (2) Wang, S.; Chao, C.; Cai, J.; Niu, B.; Copeland, L.; Wang, S. Starch–Lipid and Starch–Lipid–Protein Complexes: A Comprehensive Review. *Compr. Rev. Food Sci. Food Saf.* **2020**, *19* (3), 1056–1079. <https://doi.org/10.1111/1541-4337.12550>.
- (3) Le Bail, P.; Rondeau, C.; Buléon, A. Structural Investigation of Amylose Complexes with Small Ligands: Helical Conformation, Crystalline Structure and Thermostability. *Int. J. Biol. Macromol.* **2005**, *35* (1–2), 1–7. <https://doi.org/10.1016/j.ijbiomac.2004.09.001>.
- (4) Nuessli, J.; Putaux, J. L.; Le Bail, P.; Buléon, A. Crystal Structure of Amylose Complexes with Small Ligands. *Int. J. Biol. Macromol.* **2003**, *33* (4–5), 227–234. <https://doi.org/10.1016/j.ijbiomac.2003.08.009>.
- (5) Carbinatto, F. M.; Ribeiro, T. S.; Colnago, L. A.; Evangelista, R. C.; Cury, B. S. F. Preparation and Characterization of Amylose Inclusion Complexes for Drug Delivery Applications. *J. Pharm. Sci.* **2016**, *105* (1), 231–241. <https://doi.org/10.1002/jps.24702>.
- (6) Daffé, M. The Cell Envelope of Tubercle Bacilli. *Tuberculosis* **2015**. <https://doi.org/10.1016/j.tube.2015.02.024>.
- (7) Sani, M.; Houben, E. N. G.; Geurtsen, J.; Pierson, J.; De Punder, K.; Van Zon, M.; Wever, B.; Piersma, S. R.; Jiménez, C. R.; Daffé, M.; Appelmelk, B. J.; Bitter, W.; Van Wel, N. Der; Peters, P. J. Direct Visualization by Cryo-EM of the Mycobacterial Capsular Layer: A Labile Structure Containing ESX-1-Secreted Proteins. *PLoS Pathog.* **2010**, *6* (3). <https://doi.org/10.1371/journal.ppat.1000794>.
- (8) Lemassu, A.; Daffe, M. Structural Features of the Exocellular Polysaccharides

- of Mycobacterium Tuberculosis. *Biochem. J.* **1994**, 297 (2), 351–357. <https://doi.org/10.1042/bj2970351>.
- (9) Ortalo-Magné, A.; Lemassu, A.; Lanéelle, M. A.; Bardou, F.; Silve, G.; Gounon, P.; Marchal, G.; Daffé, M. Identification of the Surface-Exposed Lipids on the Cell Envelopes of Mycobacterium Tuberculosis and Other Mycobacterial Species. *J. Bacteriol.* **1996**, 178 (2), 456–461. <https://doi.org/10.1128/jb.178.2.456-461.1996>.
- (10) Nakamura, Y. *Starch: Metabolism and Structure*; Nakamura, Y., Ed.; Springer Japan: Tokyo, 2015; Vol. 369. <https://doi.org/10.1007/978-4-431-55495-0>.
- (11) Autio, K.; Baldwin, P. M.; Bassi, S. D.; Biliaderis, C. G.; Boyer, C. D.; Breuninger, W. F.; Chiu, C.; Eckhoff, S. R.; Eliasson, A. C.; Farris, P. L.; Gallant, D. J.; Garwood, D. L.; Grommers, H. E.; Hedges, A.; Hobbs, L.; Hoover, R.; Jane, J.; Lasater, G. D.; Maningat, C. C.; Mason, W. R.; Maurer, H. W.; Mitchell, C. R.; Perez, S.; Piyachomkwan, K.; Preiss, J.; Robyt, J. F.; Schwartz, D.; Seib, P. A.; Shannon, J. C.; Solarek, D.; Sriroth, K.; Van der Krogt, D. A.; Vasanthan, T.; Watson, S. A.; Whistler, R. L.; Willett, J. L.; Woo, K. *Starch Chemistry and Technology*; 2013; Vol. 369. <https://doi.org/10.1017/CBO9781107415324.004>.
- (12) Le, C. A. K.; Choisnard, L.; Wouessidjewe, D.; Putaux, J. L. Polymorphism of Crystalline Complexes of V-Amylose with Fatty Acids. *Int. J. Biol. Macromol.* **2018**, 119, 555–564. <https://doi.org/10.1016/j.ijbiomac.2018.07.163>.
- (13) Kim, H. I.; Kim, H. R.; Choi, S. J.; Park, C. S.; Moon, T. W. Preparation and Characterization of the Inclusion Complexes between Amylosucrase-Treated Waxy Starch and Palmitic Acid. *Food Sci. Biotechnol.* **2017**, 26 (2), 323–329. <https://doi.org/10.1007/s10068-017-0044-z>.
- (14) Le-Bail, P.; Lorentz, C.; Pencreac'h, G.; Sultani-Vigneron, S.; Pontoire, B.; Giraldo, L. J. L.; Villeneuve, P.; Hendrickx, J.; Tran, V. Trapping by Amylose of the Aliphatic Chain Grafted onto Chlorogenic Acid: Importance of the Graft Position. *Carbohydr. Polym.* **2015**, 117, 910–916. <https://doi.org/10.1016/j.carbpol.2014.10.029>.
- (15) Kong, L.; Perez-Santos, D. M.; Ziegler, G. R. Effect of Guest Structure on Amylose-Guest Inclusion Complexation. *Food Hydrocoll.* **2019**, 97 (June),

105188. <https://doi.org/10.1016/j.foodhyd.2019.105188>.
- (16) Kawai, K.; Takato, S.; Ueda, M.; Ohnishi, N.; Viriyarattanasak, C.; Kajiwara, K. Effects of Fatty Acid and Emulsifier on the Complex Formation and in Vitro Digestibility of Gelatinized Potato Starch. *Int. J. Food Prop.* **2017**, *20* (7), 1500–1510. <https://doi.org/10.1080/10942912.2016.1212877>.
- (17) Godet, M. C.; Buléon, A.; Tran, V.; Colonna, P. Structural Features of Fatty Acid-Amylose Complexes. *Carbohydr. Polym.* **1993**. [https://doi.org/10.1016/0144-8617\(93\)90003-M](https://doi.org/10.1016/0144-8617(93)90003-M).
- (18) Lebail, P.; Buleon, A.; Shifftan, D.; Marchessault, R. H. Mobility of Lipid in Complexes of Amylose-Fatty Acids by Deuterium And<sup>13</sup>C Solid State NMR. *Carbohydr. Polym.* **2000**. [https://doi.org/10.1016/S0144-8617\(00\)00180-6](https://doi.org/10.1016/S0144-8617(00)00180-6).
- (19) Hasjim, J.; Lee, S. O.; Hendrich, S.; Setiawan, S.; Ai, Y.; Jane, J. L. Characterization of a Novel Resistant-Starch and Its Effects on Postprandial Plasma-Glucose and Insulin Responses. *Cereal Chem.* **2010**, *87* (4), 257–262. <https://doi.org/10.1094/CCHEM-87-4-0257>.
- (20) Zhao, Y.; Ai, Y.; Li, L.; Jane, J. L.; Hendrich, S.; Birt, D. F. Inhibition of Azoxymethane-Induced Preneoplastic Lesions in the Rat Colon by a Stearic Acid Complexed High-Amylose Cornstarch Using Different Cooking Methods and Assessing Potential Gene Targets. *J. Funct. Foods* **2014**, *6* (1), 499–512. <https://doi.org/10.1016/j.jff.2013.11.016>.
- (21) Birt, D. F.; Boylston, T.; Hendrich, S.; Jane, J.-L.; Hollis, J.; Li, L.; McClelland, J.; Moore, S.; Phillips, G. J.; Rowling, M.; Schalinske, K.; Scott, M. P.; Whitley, E. M. Resistant Starch: Promise for Improving Human Health. *Adv. Nutr.* **2013**, *4* (6), 587–601. <https://doi.org/10.3945/an.113.004325>.
- (22) Nielsen, T. S.; Bendiks, Z.; Thomsen, B.; Wright, M. E.; Theil, P. K.; Scherer, B. L.; Marco, M. L. High-Amylose Maize, Potato, and Butyrylated Starch Modulate Large Intestinal Fermentation, Microbial Composition, and Oncogenic MiRNA Expression in Rats Fed a High-Protein Meat Diet. *Int. J. Mol. Sci.* **2019**, *20* (9). <https://doi.org/10.3390/ijms20092137>.
- (23) Moreno-Mendieta, S.; Barrios-Payán, J.; Mata-Espinosa, D.; Sánchez, S.; Hernández-Pando, R.; Rodríguez-Sanoja, R. Raw Starch Microparticles Have

- Immunostimulant Activity in Mice Vaccinated with BCG and Challenged with Mycobacterium Tuberculosis. *Vaccine* **2017**, *35* (38), 5123–5130. <https://doi.org/10.1016/j.vaccine.2017.08.012>.
- (24) Moreno-Mendieta, S.; Barrera-Rosales, A.; Mata-Espinosa, D.; Barrios-Payán, J.; Sánchez, S.; Hernández-Pando, R.; Rodríguez-Sanoja, R. Raw Starch Microparticles as BCG Adjuvant: Their Efficacy Depends on the Virulence of the Infection Strains. *Vaccine* **2019**, *37* (38), 5731–5737. <https://doi.org/10.1016/j.vaccine.2019.04.027>.
- (25) Pérez, S.; Bertoft, E. The Molecular Structures of Starch Components and Their Contribution to the Architecture of Starch Granules: A Comprehensive Review. *Starch/Staerke* **2010**, *62* (8), 389–420. <https://doi.org/10.1002/star.201000013>.
- (26) Hizukuri, S.; Takeda, Y.; Yasuda, M.; Suzuki, A. Multi-Branched Nature of Amylose and the Action of Debranching Enzymes. *Carbohydr. Res.* **1981**, *94* (2), 205–213. [https://doi.org/10.1016/S0008-6215\(00\)80718-1](https://doi.org/10.1016/S0008-6215(00)80718-1).
- (27) Gunning, A. P.; Giardina, T. P.; Faulds, C. B.; Juge, N.; Ring, S. G.; Williamson, G.; Morris, V. J. Surfactant-Mediated Solubilisation of Amylose and Visualisation by Atomic Force Microscopy. *Carbohydr. Polym.* **2003**, *51* (2), 177–182. [https://doi.org/10.1016/S0144-8617\(02\)00126-1](https://doi.org/10.1016/S0144-8617(02)00126-1).
- (28) Fredriksson, H.; Silverio, J.; Andersson, R.; Eliasson, A. C.; Åman, P. The Influence of Amylose and Amylopectin Characteristics on Gelatinization and Retrogradation Properties of Different Starches. *Carbohydr. Polym.* **1998**, *35* (3–4), 119–134. [https://doi.org/10.1016/S0144-8617\(97\)00247-6](https://doi.org/10.1016/S0144-8617(97)00247-6).
- (29) Carciofi, M.; Blennow, A.; Jensen, S. L.; Shaik, S. S.; Henriksen, A.; Buléon, A.; Holm, P. B.; Hebelstrup, K. H. Concerted Suppression of All Starch Branching Enzyme Genes in Barley Produces Amylose-Only Starch Granules. *BMC Plant Biol.* **2012**, *12*, 1–16. <https://doi.org/10.1186/1471-2229-12-223>.
- (30) Kubo, A.; Akdogan, G.; Nakaya, M.; Shojo, A.; Suzuki, S.; Satoh, H.; Kitamura, S. Structure, Physical, and Digestive Properties of Starch from Wx Ae Double-Mutant Rice. *J. Agric. Food Chem.* **2010**, *58* (7), 4463–4469. <https://doi.org/10.1021/jf904074k>.
- (31) Glaring, M. A.; Koch, C. B.; Blennow, A. Genotype-Specific Spatial Distribution

- of Starch Molecules in the Starch Granule: A Combined CLSM and SEM Approach. *Biomacromolecules* **2006**, 7 (8), 2310–2320. <https://doi.org/10.1021/bm060216e>.
- (32) Pilling, E.; Smith, A. M. Growth Ring Formation in the Starch Granules of Potato Tubers. *Plant Physiol.* **2003**, 132 (1), 365–371. <https://doi.org/10.1104/pp.102.018044>.
- (33) Hanson, E. A.; Katz, J. R. Abhandlungen Zur Physikalischen Chemie Der Stärke Und Der Brotbereitung. XVII. Über Versuche Die Gewachsene Struktur Des Stärkekorns Mikroskopisch Sichtbar Zu Machen Besonders an Lintnerisierter Stärke. *Zeitschrift für Phys. Chemie* **1934**, No. 1, 339–352. <https://doi.org/10.1515/zpch-1937-18034>.
- (34) Gallant, D. J.; Bouchet, B.; Baldwin, P. M. Microscopy of Starch: Evidence of a New Level of Granule Organization. *Carbohydr. Polym.* **1997**, 32 (3–4), 177–191. [https://doi.org/10.1016/S0144-8617\(97\)00008-8](https://doi.org/10.1016/S0144-8617(97)00008-8).
- (35) Chauhan, F.; Seetharaman, K. On the Organization of Chains in Amylopectin. *Starch/Staerke* **2013**, 65 (3–4), 191–199. <https://doi.org/10.1002/star.201200132>.
- (36) Tang, H.; Mitsunaga, T.; Kawamura, Y. Molecular Arrangement in Blocklets and Starch Granule Architecture. *Carbohydr. Polym.* **2006**, 63 (4), 555–560. <https://doi.org/10.1016/j.carbpol.2005.10.016>.
- (37) Sterling, C. A Low Angle Spacing in Starch. *J. Polym. Sci.* **1962**, 56 (163), S10–S12.
- (38) Bertoft, E. Understanding Starch Structure: Recent Progress. *Agronomy* **2017**, 7 (3). <https://doi.org/10.3390/agronomy7030056>.
- (39) Bertoft, E.; Piyachomkwan, K.; Chatakanonda, P.; Sriroth, K. Internal Unit Chain Composition in Amylopectins. *Carbohydr. Polym.* **2008**, 74 (3), 527–543. <https://doi.org/10.1016/j.carbpol.2008.04.011>.
- (40) Zhong, Y.; Bertoft, E.; Li, Z.; Blennow, A.; Liu, X. Amylopectin Starch Granule Lamellar Structure as Deduced from Unit Chain Length Data. *Food Hydrocoll.* **2020**, 108 (March), 106053. <https://doi.org/10.1016/j.foodhyd.2020.106053>.

- (41) Genkina, N. K.; Wikman, J.; Bertoft, E.; Yuryev, V. P. Effects of Structural Imperfection on Gelatinization Characteristics of Amylopectin Starches with A- and B-Type Crystallinity. *Biomacromolecules* **2007**, *8* (7), 2329–2335. <https://doi.org/10.1021/bm070349f>.
- (42) Daniels, D. R.; Donald, A. M. Soft Material Characterization of the Lamellar Properties of Starch: Smectic Side-Chain Liquid-Crystalline Polymeric Approach. *Macromolecules* **2004**, *37* (4), 1312–1318. <https://doi.org/10.1021/ma030360h>.
- (43) Koroteeva, D. A.; Kiseleva, V. I.; Krivandin, A. V.; Shatalova, O. V.; Błaszczak, W.; Bertoft, E.; Piyachomkwan, K.; Yuryev, V. P. Structural and Thermodynamic Properties of Rice Starches with Different Genetic Background. Part 2. Defectiveness of Different Supramolecular Structures in Starch Granules. *Int. J. Biol. Macromol.* **2007**, *41* (5), 534–547. <https://doi.org/10.1016/j.ijbiomac.2007.07.005>.
- (44) Kozlov, S. S.; Krivandin, A. V.; Shatalova, O. V.; Noda, T.; Bertoft, E.; Fornal, J.; Yuryev, V. P. Structure of Starches Extracted from Near-Isogenic Wheat Lines : PPPart 2. Molecular Organization of Amylopectin Clusters. *J. Therm. Anal. Calorim.* **2007**, *87* (2), 575–584. <https://doi.org/10.1007/s10973-006-7880-z>.
- (45) Donovan, J. W. Phase Transitions of the Starch–Water System. *Biopolymers* **1979**, *18* (2), 263–275. <https://doi.org/10.1002/bip.1979.360180204>.
- (46) Zobel, H. F.; Young, S. N.; Rocca, L. A. Starch Gelatinization: An X-Ray Diffraction Study'. *Cereal Chem.* **1988**, *65* (6), 443–446.
- (47) Cooke, D.; Gidley, M. J. Loss of Crystalline and Molecular Order during Starch Gelatinisation: Origin of the Enthalpic Transition. *Carbohydr. Res.* **1992**, *227* (C), 103–112. [https://doi.org/10.1016/0008-6215\(92\)85063-6](https://doi.org/10.1016/0008-6215(92)85063-6).
- (48) Waigh, T. A.; Gidley, M. J.; Komanshek, B. U.; Donald, A. M. The Phase Transformations in Starch during Gelatinisation: A Liquid Crystalline Approach. *Carbohydr. Res.* **2000**, *328* (2), 165–176. [https://doi.org/10.1016/S0008-6215\(00\)00098-7](https://doi.org/10.1016/S0008-6215(00)00098-7).
- (49) Jacobs, H.; Delcour, J. A. Hydrothermal Modifications of Granular Starch, with Retention of the Granular Structure: A Review. *J. Agric. Food Chem.* **1998**, *46*

- (8). <https://doi.org/10.1021/jf980169k>.
- (50) Maache-Rezzoug, Z.; Zarguili, I.; Loisel, C.; Queveau, D.; Buléon, A. Structural Modifications and Thermal Transitions of Standard Maize Starch after DIC Hydrothermal Treatment. *Carbohydr. Polym.* **2008**, *74* (4), 802–812. <https://doi.org/10.1016/j.carbpol.2008.04.047>.
- (51) Wang, S.; Wang, J.; Yu, J.; Wang, S. Effect of Fatty Acids on Functional Properties of Normal Wheat and Waxy Wheat Starches: A Structural Basis. *Food Chem.* **2016**, *190*, 285–292. <https://doi.org/10.1016/j.foodchem.2015.05.086>.
- (52) Wang, W.; Shi, Y. C. Gelatinization, Pasting and Retrogradation Properties of Hydroxypropylated Normal Wheat, Waxy Wheat, and Waxy Maize Starches. *Food Hydrocoll.* **2020**, *106* (March), 105910. <https://doi.org/10.1016/j.foodhyd.2020.105910>.
- (53) Zhi-guang, C.; Hong-hui, Z.; Keipper, W.; Hua-yin, P.; Qi, Y.; Chen-lu, F.; Guo-wei, S.; Jun-rong, H. The Analysis of the Effects of High Hydrostatic Pressure (HHP) on Amylose Molecular Conformation at Atomic Level Based on Molecular Dynamics Simulation. *Food Chem.* **2020**, *327* (May), 127047. <https://doi.org/10.1016/j.foodchem.2020.127047>.
- (54) Hermansson, A. M.; Svegmarm, K. Developments in the Understanding of Starch Functionality. *Trends Food Sci. Technol.* **1996**, *7* (11), 345–353. [https://doi.org/10.1016/S0924-2244\(96\)10036-4](https://doi.org/10.1016/S0924-2244(96)10036-4).
- (55) Wang, S.; Li, C.; Copeland, L.; Niu, Q.; Wang, S. Starch Retrogradation: A Comprehensive Review. *Compr. Rev. Food Sci. Food Saf.* **2015**, *14* (5), 568–585. <https://doi.org/10.1111/1541-4337.12143>.
- (56) Hoover, R.; Hughes, T.; Chung, H. J.; Liu, Q. Composition, Molecular Structure, Properties, and Modification of Pulse Starches: A Review. *Food Res. Int.* **2010**, *43* (2), 399–413. <https://doi.org/10.1016/j.foodres.2009.09.001>.
- (57) Tang, M. C.; Copeland, L. Investigation of Starch Retrogradation Using Atomic Force Microscopy. *Carbohydr. Polym.* **2007**, *70* (1), 1–7. <https://doi.org/10.1016/j.carbpol.2007.02.025>.
- (58) Ai, Y.; Hasjim, J.; Jane, J. L. Effects of Lipids on Enzymatic Hydrolysis and

- Physical Properties of Starch. *Carbohydr. Polym.* **2013**, *92* (1), 120–127. <https://doi.org/10.1016/j.carbpol.2012.08.092>.
- (59) Wang, S.; Zheng, M.; Yu, J.; Wang, S.; Copeland, L. Insights into the Formation and Structures of Starch-Protein-Lipid Complexes. *J. Agric. Food Chem.* **2017**, *65* (9), 1960–1966. <https://doi.org/10.1021/acs.jafc.6b05772>.
- (60) Eliasson, A. C.; Krog, N. Physical Properties of Amylose-Monoglyceride Complexes. *J. Cereal Sci.* **1985**, *3* (3), 239–248. [https://doi.org/10.1016/S0733-5210\(85\)80017-5](https://doi.org/10.1016/S0733-5210(85)80017-5).
- (61) Meyer, K. H.; Brentano, W.; Bernfeld, P. Recherches Sur l'amidon II. Sur La Nonhomogénéité de l'amidon. *Helv. Chim. Acta* **1940**, *1103* (1929), 845–853.
- (62) McCready, R. M.; Hassid, W. Z. The Separation and Quantitative Estimation of Amylose and Amylopectin in Potato Starch. *J. Am. Chem. Soc.* **1943**, *65* (6), 1154–1157. <https://doi.org/10.1021/ja01246a038>.
- (63) Juliano, B. O.; Tuaño, A. P. P.; Monteroso, D. N.; Aoki, N.; Mestres, C.; Duldulao, J. B. A.; Bergonio, K. B. Replacement of Acetate with Ammonium Buffer to Determine Apparent Amylose Content of Milled Rice. *Cereal Foods World* **2012**, *57* (1), 14–19. <https://doi.org/10.1094/CFW-57-1-0014>.
- (64) Zhou, X.; Wang, R.; Yoo, S. H.; Lim, S. T. Water Effect on the Interaction between Amylose and Amylopectin during Retrogradation. *Carbohydr. Polym.* **2011**, *86* (4), 1671–1674. <https://doi.org/10.1016/j.carbpol.2011.06.082>.
- (65) Ohdan, K.; Fujii, K.; Yanase, M.; Takaha, T.; Kuriki, T. Enzymatic Synthesis of Amylose. *Biocatal. Biotransformation* **2006**, *24* (1–2), 77–81. <https://doi.org/10.1080/10242420600598152>.
- (66) Yanase, M.; Takata, H.; Fujii, K.; Takaha, T.; Kuriki, T. Cumulative Effect of Amino Acid Replacements Results in Enhanced Thermostability of Potato Type L  $\alpha$ -Glucan Phosphorylase. *Appl. Environ. Microbiol.* **2005**, *71* (9), 5433–5439. <https://doi.org/10.1128/AEM.71.9.5433-5439.2005>.
- (67) Roblin, P.; Potocki-Véronèse, G.; Guieysse, D.; Guerin, F.; Axelos, M. A. V.; Perez, J.; Buleon, A. SAXS Conformational Tracking of Amylose Synthesized by Amylosucrases. *Biomacromolecules* **2013**. <https://doi.org/10.1021/bm301651y>.

- (68) Emond, S.; Mondeil, S.; Jaziri, K.; André, I.; Monsan, P.; Remaud-Siméon, M.; Potocki-Véronèse, G. Cloning, Purification and Characterization of a Thermostable Amylosucrase from *Deinococcus Geothermalis*. *FEMS Microbiol. Lett.* **2008**, *285* (1), 25–32. <https://doi.org/10.1111/j.1574-6968.2008.01204.x>.
- (69) Potocki-Veronese, G.; Putaux, J. L.; Dupeyre, D.; Albenne, C.; Remaud-Siméon, M.; Monsan, P.; Buleon, A. Amylose Synthesized in Vitro by Amylosucrase: Morphology, Structure, and Properties. *Biomacromolecules* **2005**. <https://doi.org/10.1021/bm049326g>.
- (70) Potocki De Montalk, G.; Remaud-Simeon, M.; Willemot, R. M.; Sarçabal, P.; Planchot, V.; Monsan, P. Amylosucrase from *Neisseria Polysaccharea*: Novel Catalytic Properties. *FEBS Lett.* **2000**. [https://doi.org/10.1016/S0014-5793\(00\)01406-X](https://doi.org/10.1016/S0014-5793(00)01406-X).
- (71) Guérin, F.; Barbe, S.; Pizzut-Serin, S.; Potocki-Véronèse, G.; Guieysse, D.; Guillet, V.; Monsan, P.; Mourey, L.; Remaud-Siméon, M.; André, I.; Tranier, S. Structural Investigation of the Thermostability and Product Specificity of Amylosucrase from the Bacterium *Deinococcus Geothermalis*. *J. Biol. Chem.* **2012**. <https://doi.org/10.1074/jbc.M111.322917>.
- (72) Rondeau-Mouro, C.; Veronese, G.; Buleon, A. High-Resolution Solid-State NMR of B-Type Amylose. *Biomacromolecules* **2006**. <https://doi.org/10.1021/bm060330x>.
- (73) Popov, D.; Buléon, A.; Burghammer, M.; Chanzy, H.; Montesanti, N.; Putaux, J.-L.; Potocki-Véronèse, G.; Riekkel, C. Crystal Structure of A-Amylose: A Revisit from Synchrotron Microdiffraction Analysis of Single Crystals. *Macromolecules* **2009**, *42* (4), 1167–1174. <https://doi.org/10.1021/ma801789j>.
- (74) Banks, W.; Greenwood, C. T. The Conformation of Amylose in Dilute Solution. *Starch - Stärke* **1971**, *23* (9), 300–314. <https://doi.org/10.1002/star.19710230903>.
- (75) Sarko, A.; Zeitlin, B. R.; Germino, F. J. Cold Water Soluble Amylose, May 1963.
- (76) Jane, J.; Robyt, J. F.; Huang, D.-H. <sup>13</sup>C-N.M.R. Study of the Conformation of Helical Complexes of Amylodextrin and of Amylose in Solution. *Carbohydr. Res.* **1985**, *140* (5), 21–35.

- (77) Hayashi, A.; Kinoshita, K.; Miyake, Y. The Conformation of Amylose in Solution. *Polym. J.* **1981**, *13* (6), 537–541. <https://doi.org/10.1002/star.19710230903>.
- (78) Tusch, M.; Krüger, J.; Fels, G. Structural Stability of V-Amylose Helices in Water-DMSO Mixtures Analyzed by Molecular Dynamics. *J. Chem. Theory Comput.* **2011**, *7* (9), 2919–2928. <https://doi.org/10.1021/ct2005159>.
- (79) Koneru, J. K.; Zhu, X.; Mondal, J. A Quantitative Assessment of the Conformational Heterogeneity in Amylose across Force Fields. *J. Chem. Theory Comput.* **2019**. <https://doi.org/10.1021/acs.jctc.9b00630>.
- (80) Foster, J. F.; Zucker, D. Length of the Amylose-Iodine Complex as Determined by Streaming Dichroism. *J. Phys. Chem.* **1952**, *56*, 170–173.
- (81) Bluhm, T. L.; Zugenmaier, P. Detailed Structure of the Vh-Amylose-Iodine Complex: A Linear Polyiodine Chain. *Carbohydr. Res.* **1981**, *89* (1), 1–10. [https://doi.org/10.1016/S0008-6215\(00\)85224-6](https://doi.org/10.1016/S0008-6215(00)85224-6).
- (82) Gessler, K.; Uson, I.; Takaha, T.; Krauss, N.; Smith, S. M.; Okada, S.; Sheldrick, G. M.; Saenger, W. V-Amylose at Atomic Resolution: X-Ray Structure of a Cycloamylose with 26 Glucose Residues (Cyclomaltohexacosose). *Proc. Natl. Acad. Sci.* **1999**, *96* (8), 4246–4251. <https://doi.org/10.1073/pnas.96.8.4246>.
- (83) Hollo, J.; Szejtli, J. Die Struktur Der Amylose-Moleküle in Wässrigen Lösungen. *Starch - Stärke* **1958**, *10* (3), 49–52.
- (84) Rao, V. S. R.; Foster, J. F. Studies of the Conformation of Carboxymethyl Amylose in Aqueous Solutions. *Biopolymers* **1963**, *3* (2), 185–193. <https://doi.org/10.1002/bip.360030206>.
- (85) Banks, W.; Greenwood, C. T. Amylose: A Non-Helical Biopolymer in Aqueous Solution. *Polymer (Guildf)*. **1971**, *12* (2), 141–145. [https://doi.org/10.1016/0032-3861\(71\)90005-X](https://doi.org/10.1016/0032-3861(71)90005-X).
- (86) Pfannemüller, B.; Mayerhöfer, H.; Schulz, R. C. Conformation of Amylose in Aqueous Solution: Optical Rotatory Dispersion and Circular Dichroism of Amylose–Iodine Complexes and Dependence on Chain Length of Retrogradation of Amylose. *Biopolymers* **1971**, *10* (2), 243–261. <https://doi.org/10.1002/bip.360100203>.

- (87) Buléon, A.; Véronèse, G.; Putaux, J. L. Self-Association and Crystallization of Amylose. *Aust. J. Chem.* **2007**, *60* (10), 706–718. <https://doi.org/10.1071/CH07168>.
- (88) Buleon, A.; Duprat, F.; Booy, F. P.; Chanzy, H. Single Crystals of Amylose with a Low Degree of Polymerization. *Carbohydr. Polym.* **1984**, *4* (3), 161–173. [https://doi.org/10.1016/0144-8617\(84\)90009-2](https://doi.org/10.1016/0144-8617(84)90009-2).
- (89) Imberty, A.; Chanzy, H.; Perez, S.; Buleon, A.; Tran, V. The Double-Helical Nature of the Crystalline Part of A-Starch. *J. Mol. Biol.* **1988**, *201*, 365–378.
- (90) Imberty, A.; Perez, S. A Revisit to the Three-Dimensional Structure of B-Type Starch. *Biopolymers* **1988**, *27*, 1205–1221.
- (91) Takahashi, Y.; Kumano, T.; Nishikawa, S. Crystal Structure of B-Amylose. *Macromolecules* **2004**, *37* (18), 6827–6832. <https://doi.org/10.1021/ma0490956>.
- (92) Wu, H.-C. H.; Sarko, A. The Double-Helical Molecular Structure of Crystalline A-Amylose. *Carbohydr. Res.* **1978**, *60* (2), 27–40. [https://doi.org/10.1016/s0008-6215\(00\)83568-5](https://doi.org/10.1016/s0008-6215(00)83568-5).
- (93) Poulhazan, A.; Arnold, A. A.; Warschawski, D. E.; Marcotte, I. Unambiguous Ex Situ and in Cell 2D 13 C Solid-State Nmr Characterization of Starch and Its Constituents. *Int. J. Mol. Sci.* **2018**, *19* (12). <https://doi.org/10.3390/ijms19123817>.
- (94) Katz, J. R. Abhandlungen Zur Physikalischen Chemie Der Stärke and Der Brotbereitung. I. Über Die Änderungen Im Röntgenspektrum Der Stärke Beim Backen Und Beim Altbackenwerden Des Brotes. *Zeitschrift für Phys. Chemie* **1930**, *150A* (1), 37–59.
- (95) Katz, J. R.; Itallie, T. B. Abhandlungen Zur Physikalischen Chemie Der Stärke and Der Brotbereitung. V. Alle Stärkearten Haben Das Gleiche Retrogradationsspektrum. *Zeitschrift für Phys. Chemie* **1930**, *150A* (1), 90–99.
- (96) Putaux, J. L.; Nishiyama, Y.; Mazeau, K.; Morin, M.; Cardoso, M. B.; Chanzy, H. Helical Conformation in Crystalline Inclusion Complexes of V-Amylose: A Historical Perspective. *Macromol. Symp.* **2011**, *303* (1), 1–9. <https://doi.org/10.1002/masy.201150501>.

- (97) Putseys, J. A.; Lamberts, L.; Delcour, J. A. Amylose-Inclusion Complexes: Formation, Identity and Physico-Chemical Properties. *J. Cereal Sci.* **2010**, *51* (3), 238–247. <https://doi.org/10.1016/j.jcs.2010.01.011>.
- (98) Rundle, R. E.; Foster, J. F.; Baldwin, R. R. On the Nature of the Starch-Iodine Complex. *J. Am. Chem. Soc.* **1943**, *66* (5), 2116–2120.
- (99) Rundle, R. E.; Edwards, C. The Configuration of Starch in the Starch-Iodine Complex. IV. An X-Ray Dif- Fraction Investigation of Butanol-Precipitated Amylose. *J. Am. Chem. Soc.* **1943**, *65* (1942), 2200–2203.
- (100) Nimz, O.; Geßler, K.; Usón, I.; Laettig, S.; Welfle, H.; Sheldrick, G. M.; Saenger, W. X-Ray Structure of the Cyclomaltohexaicosaoose Triiodide Inclusion Complex Provides a Model for Amylose-Iodine at Atomic Resolution. *Carbohydr. Res.* **2003**, *338* (9), 977–986. [https://doi.org/10.1016/S0008-6215\(03\)00016-8](https://doi.org/10.1016/S0008-6215(03)00016-8).
- (101) Bear, S. Complex Formation between Starch and Organic Molecules. *J. Am. Chem. Soc.* **1944**, *66*, 2122–2123.
- (102) Yamashita, Y. Single Crystals of Amylose V Complexes. *J. Polym. Sci. A.* **1965**, *3* (9), 3251–3260. <https://doi.org/10.1002/pol.1965.100030919>.
- (103) Oguchi, T.; Yamasato, H.; Limmatvapirat, S.; Yonemochi, E.; Yamamoto, K. Structural Change and Complexation of Strictly Linear Amloose Induced by Sealed-Heaing with Salicyclic Acid. *J. Chem. Soc. - Faraday Trans.* **1998**, *94* (7), 923–927. <https://doi.org/10.1039/a707848j>.
- (104) Cardoso, M. B.; Putaux, J. L.; Nishiyama, Y.; Helbert, W.; Hütch, M.; Silveira, N. P.; Chanzy, H. Single Crystals of V-Amylose Complexed with  $\alpha$ -Naphthol. *Biomacromolecules* **2007**, *8* (4), 1319–1326. <https://doi.org/10.1021/bm0611174>.
- (105) Le, C. A. K. Inclusion Complexes of Amylose: Morphogenesis, Crystal Structure and Release of Bioactive Molecules, Ph.D Thesis from Grenoble Alpes University, 2018.
- (106) Karkalas, J.; Ma, S.; Morrison, W. R.; Pethrick, R. A. Some Factors Determining the Thermal Properties of Amylose Inclusion Complexes with Fatty Acids. *Carbohydr. Res.* **1995**. [https://doi.org/10.1016/0008-6215\(94\)00336-E](https://doi.org/10.1016/0008-6215(94)00336-E).

- (107) Biliaderis, C. G.; Page, C. M.; Slade, L.; Sirett, R. R. Thermal Behavior of Amylose-Lipid Complexes. *Carbohydr. Polym.* **1985**, *5* (5), 367–389. [https://doi.org/10.1016/0144-8617\(85\)90044-X](https://doi.org/10.1016/0144-8617(85)90044-X).
- (108) Tufvesson, F.; Wahlgren, M.; Eliasson, A.-C. Formation of Amylose-Lipid Complexes and Effects of Temperature Treatment. Part 2. Fatty Acids. *Starch - Stärke* **2003**, *55* (34), 138–149. <https://doi.org/10.1002/star.200390028>.
- (109) Takeo, K.; Kuge, T. Complexes of Starchy Materials with Organic Compounds. *Agric. Biol. Chem.* **1969**, *33* (8), 1174–1180. <https://doi.org/10.1080/00021369.1969.10859434>.
- (110) Tufvesson, F.; Wahlgren, M.; Eliasson, A. C. Formation of Amylose-Lipid Complexes and Effects of Temperature Treatment - Part 1. Monoglycerides. *Starch/Staerke* **2003**, *55* (2), 61–71. <https://doi.org/10.1002/star.200390018>.
- (111) MERCIER, C.; CHARBONNIERE, R.; GALLANT, D.; GUILBOT, A. Structural Modification of Various Starches By Extrusion Cooking With a Twin-Screw French Extruder. *Polysaccharides Food* **1979**, 153–170. <https://doi.org/10.1016/b978-0-408-10618-4.50015-9>.
- (112) Mercier, C.; Charbonniere, R.; Grebaut, J.; Gueriviere, J. F. Formation of Amylose-Lipid Complexes by Twin-Screw Extrusion Cooking of Manioc Starch. *Cereal Chem.* **1980**, *57* (1), 4–9.
- (113) Bhatnagar, S.; Hanna, M. A. Extrusion Processing Conditions for Amylose-Lipid Complexing. *Cereal Chem.* **1994**, *71* (6), 587–593.
- (114) Bhatnagar, S.; Hanna, M. a. Amylose-Lipid Complex Formation During Single-Screw Extrusion of Various Corn Starches. *Cereal Chem.* **1994**, *71*, 582–587.
- (115) Bhatnagar, S.; Hanna, M. A. Starch-Stearic Acid Complex Development within Single and Twin Screw Extruders. *J. Food Sci.* **1996**, *61* (4), 778–782. <https://doi.org/10.1111/j.1365-2621.1996.tb12202.x>.
- (116) De Pilli, T.; Jouppila, K.; Ikonen, J.; Kansikas, J.; Derossi, A.; Severini, C. Study on Formation of Starch-Lipid Complexes during Extrusion-Cooking of Almond Flour. *J. Food Eng.* **2008**, *87* (4), 495–504. <https://doi.org/10.1016/j.jfoodeng.2007.12.028>.

- (117) De Pilli, T.; Derossi, A.; Talja, R. A.; Jouppila, K.; Severini, C. Study of Starch-Lipid Complexes in Model System and Real Food Produced Using Extrusion-Cooking Technology. *Innov. Food Sci. Emerg. Technol.* **2011**, *12* (4), 610–616. <https://doi.org/10.1016/j.ifset.2011.07.011>.
- (118) Lesmes, U.; Barchechath, J.; Shimoni, E. Continuous Dual Feed Homogenization for the Production of Starch Inclusion Complexes for Controlled Release of Nutrients. *Innov. Food Sci. Emerg. Technol.* **2008**, *9* (4), 507–515. <https://doi.org/10.1016/j.ifset.2007.12.008>.
- (119) Shimoni, E.; Lesmes, U.; Ungar, Y. Non-Covalent Complexes of Bioactive Agents with Starch for Oral Delivery, Patent WO 2007/122624 A3, 2007.
- (120) Meng, S.; Ma, Y.; Cui, J.; Sun, D. W. Preparation of Corn Starch-Fatty Acid Complexes by High-Pressure Homogenization. *Stand. News* **2014**, *66* (9–10), 809–817. <https://doi.org/10.1002/star.201400022>.
- (121) Fanta, G. F.; Shogren, R. L.; Salch, J. H. Steam Jet Cooking of High-Amylose Starch-Fatty Acid Mixtures. An Investigation of Complex Formation. *Carbohydr. Polym.* **1999**, *38* (1), 1–6. [https://doi.org/10.1016/S0144-8617\(98\)00104-0](https://doi.org/10.1016/S0144-8617(98)00104-0).
- (122) Fanta, G. F.; Felker, F. C.; Shogren, R. L. Formation of Crystalline Aggregates in Slowly-Cooled Starch Solutions Prepared by Steam Jet Cooking. *Carbohydr. Polym.* **2002**, *48* (2), 161–170. [https://doi.org/10.1016/S0144-8617\(01\)00230-2](https://doi.org/10.1016/S0144-8617(01)00230-2).
- (123) Fanta, G. F.; Felker, F. C.; Shogren, R. L.; Byars, J. A.; Salch, J. H. Crystalline Particles Formed in Slowly-Cooled Cornstarch Dispersions Prepared by Steam Jet Cooking. the Effect of Starch Concentration, Added Oil and Rate of Cooling. *Carbohydr. Polym.* **2005**, *61* (2), 222–230. <https://doi.org/10.1016/j.carbpol.2005.05.013>.
- (124) Fanta, G. F.; Felker, F. C.; Shogren, R. L.; Salch, J. H. Effect of Fatty Acid Structure on the Morphology of Spherulites Formed from Jet Cooked Mixtures of Fatty Acids and Defatted Cornstarch. *Carbohydr. Polym.* **2006**, *66* (1), 60–70. <https://doi.org/10.1016/j.carbpol.2006.02.017>.
- (125) Fanta, G. F.; Felker, F. C.; Shogren, R. L.; Salch, J. H. Preparation of Spherulites from Jet Cooked Mixtures of High Amylose Starch and Fatty Acids. Effect of Preparative Conditions on Spherulite Morphology and Yield. *Carbohydr. Polym.*

- 2008**, 71 (2), 253–262. <https://doi.org/10.1016/j.carbpol.2007.05.034>.
- (126) Tang, M. C.; Copeland, L. Analysis of Complexes between Lipids and Wheat Starch. *Carbohydr. Polym.* **2007**, 67 (1), 80–85. <https://doi.org/10.1016/j.carbpol.2006.04.016>.
- (127) Chao, C.; Yu, J.; Wang, S.; Copeland, L.; Wang, S. Mechanisms Underlying the Formation of Complexes between Maize Starch and Lipids. *J. Agric. Food Chem.* **2018**, 66 (1), 272–278. <https://doi.org/10.1021/acs.jafc.7b05025>.
- (128) Bhosale, R. G.; Ziegler, G. R. Preparation of Spherulites from Amylose-Palmitic Acid Complexes. *Carbohydr. Polym.* **2010**, 80 (1), 53–64. <https://doi.org/10.1016/j.carbpol.2009.10.069>.
- (129) Genkina, N. K.; Kiseleva, V. I.; Martirosyan, V. V. Different Types of v Amylose-Lipid Inclusion Complexes in Maize Extrudates Revealed by DSC Analysis. *Starch/Staerke* **2015**, 67 (9–10), 752–755. <https://doi.org/10.1002/star.201500012>.
- (130) Kadokawa, J. I.; Kaneko, Y.; Nagase, S. I.; Takahashi, T.; Tagaya, H. Vine-Twining Polymerization: Amylose Twines around Polyethers to Form Amylose-Polyether Inclusion Complexes. *Chem. - A Eur. J.* **2002**, 8 (15), 3321–3326. [https://doi.org/10.1002/1521-3765\(20020802\)8:15<3321::AID-CHEM3321>3.0.CO;2-E](https://doi.org/10.1002/1521-3765(20020802)8:15<3321::AID-CHEM3321>3.0.CO;2-E).
- (131) Kaneko, Y.; Kadokawa, J. I. Vine-Twining Polymerization: A New Preparation Method for Well-Defined Supramolecules Composed of Amylose and Synthetic Polymers. *Chem. Rec.* **2005**, 5 (1), 36–46. <https://doi.org/10.1002/tcr.20031>.
- (132) Kaneko, Y.; Beppu, K.; Kadokawa, J. I. Preparation of Amylose/Polycarbonate Inclusion Complexes by Means of Vine-Twining Polymerization. *Macromol. Chem. Phys.* **2008**, 209 (10), 1037–1042. <https://doi.org/10.1002/macp.200800006>.
- (133) Gelders, G. G.; Goesaert, H.; Delcour, J. A. Potato Phosphorylase Catalyzed Synthesis of Amylose-Lipid Complexes. *Biomacromolecules* **2005**, 6 (5), 2622–2629. <https://doi.org/10.1021/bm0502011>.
- (134) Putseys, J. A.; Derde, L. J.; Lamberts, L.; Goesaert, H.; Delcour, J. A. Production

- of Tailor Made Short Chain Amylose-Lipid Complexes Using Varying Reaction Conditions. *Carbohydr. Polym.* **2009**, *78* (4), 854–861. <https://doi.org/10.1016/j.carbpol.2009.07.002>.
- (135) Gelders, G. G.; Vanderstukken, T. C.; Goesaert, H.; Delcour, J. A. Amylose-Lipid Complexation: A New Fractionation Method. *Carbohydr. Polym.* **2004**. <https://doi.org/10.1016/j.carbpol.2004.03.012>.
- (136) Kim, J. Y.; Lim, S. T. Preparation of Nano-Sized Starch Particles by Complex Formation with n-Butanol. *Carbohydr. Polym.* **2009**, *76* (1), 110–116. <https://doi.org/10.1016/j.carbpol.2008.09.030>.
- (137) Kim, J. Y.; Yoon, J. W.; Lim, S. T. Formation and Isolation of Nanocrystal Complexes between Dextrins and N-Butanol. *Carbohydr. Polym.* **2009**, *78* (3), 626–632. <https://doi.org/10.1016/j.carbpol.2009.05.026>.
- (138) Reddy, C. K.; Choi, S. M.; Lee, D. J.; Lim, S. T. Complex Formation between Starch and Stearic Acid: Effect of Enzymatic Debranching for Starch. *Food Chem.* **2018**, *244* (September 2017), 136–142. <https://doi.org/10.1016/j.foodchem.2017.10.040>.
- (139) Reddy, C. K.; Lee, D. J.; Lim, S. T.; Park, E. Y. Enzymatic Debranching of Starches from Different Botanical Sources for Complex Formation with Stearic Acid. *Food Hydrocoll.* **2019**, *89*, 856–863. <https://doi.org/10.1016/j.foodhyd.2018.11.059>.
- (140) Leyva-porras, C.; Cruz-alcantar, P.; Espinosa-sol, V.; Saavedra-leos, M. Z. Application of Differential Scanning Calorimetry ( DSC ) and Modulated Differential Scanning Calorimetry (MDSC) in Food and Drug Industries. *Polymers (Basel)*. **2019**, *12* (5), 1–21.
- (141) Goderis, B.; Putseys, J. A.; Gommers, C. J.; Bosmans, G. M.; Delcour, J. A. The Structure and Thermal Stability of Amylose-Lipid Complexes: A Case Study on Amylose-Glycerol Monostearate. *Cryst. Growth Des.* **2014**. <https://doi.org/10.1021/cg4016355>.
- (142) Blazek, J.; Gilbert, E. P. Application of Small-Angle X-Ray and Neutron Scattering Techniques to the Characterisation of Starch Structure: A Review. *Carbohydr. Polym.* **2011**, *85* (2), 281–293.

<https://doi.org/10.1016/j.carbpol.2011.02.041>.

- (143) Godet, M. C.; Bizot, H.; Buléon, A. Crystallization of Amylose-Fatty Acid Complexes Prepared with Different Amylose Chain Lengths. *Carbohydr. Polym.* **1995**. [https://doi.org/10.1016/0144-8617\(95\)00034-5](https://doi.org/10.1016/0144-8617(95)00034-5).
- (144) Biliaderis, C. G.; Seneviratne, H. D. On the Supramolecular Structure and Metastability of Glycerol Monostearate-Amylose Complex. *Carbohydr. Polym.* **1990**, *13* (2), 185–206. [https://doi.org/10.1016/0144-8617\(90\)90083-5](https://doi.org/10.1016/0144-8617(90)90083-5).
- (145) Rappenecker, G.; Zugenmaier, P. Detailed Refinement of the Crystal Structure of Vh-Amylose. *Carbohydr. Res.* **1981**, *89* (1), 11–19. [https://doi.org/10.1016/S0008-6215\(00\)85225-8](https://doi.org/10.1016/S0008-6215(00)85225-8).
- (146) Zabar, S.; Lesmes, U.; Katz, I.; Shimoni, E.; Bianco-Peled, H. Studying Different Dimensions of Amylose-Long Chain Fatty Acid Complexes: Molecular, Nano and Micro Level Characteristics. *Food Hydrocoll.* **2009**. <https://doi.org/10.1016/j.foodhyd.2009.02.004>.
- (147) Schoch, T. J. Fractionation of Starch by Selective Precipitation with Butanol. *J. Am. Chem. Soc.* **1942**, *64* (12), 2957–2961. <https://doi.org/10.1021/ja01264a065>.
- (148) Kerr, R. W.; Severson, G. M. On the Multiple Amylose Concept on Starch. III. The Isolation of an Amylose in Crystalline Form. *J. Am. Chem. Soc.* **1943**, *65* (2), 193–198. <https://doi.org/10.1021/ja01242a015>.
- (149) Hirai, N.; Yasui, T.; Fujita, S.; Yamashita, Y. Physicochemical Study on Crystalline Texture of Starch. I. Single Crystal of  $\alpha$ -Amylose. *Kobunshi Kagaku* **1963**, *20* (219), 413–416. <https://doi.org/10.1017/CBO9781107415324.004>.
- (150) Yamashita, Y. Physicochemical Study on Crystalline Texture of Starch II. The Fold Structure of Amylose Helices in Single Crystals. *Kobunshi Kagaku* **1964**, *21* (226), 103–108.
- (151) Nishiyama, Y.; Mazeau, K.; Morin, M.; Cardoso, M. B.; Chanzy, H.; Putaux, J. L. Molecular and Crystal Structure of 7-Fold V-Amylose Complexed with 2-Propanol. *Macromolecules* **2010**, *43* (20), 8628–8636. <https://doi.org/10.1021/ma101794w>.

- (152) Marinopoulou, A.; Papastergiadis, E.; Raphaelides, S. N.; Kontominas, M. G. Morphological Characteristics, Oxidative Stability and Enzymic Hydrolysis of Amylose-Fatty Acid Complexes. *Carbohydr. Polym.* **2016**, *141*, 106–115. <https://doi.org/10.1016/j.carbpol.2015.12.062>.
- (153) Zhu, F. NMR Spectroscopy of Starch Systems. *Food Hydrocoll.* **2017**, *63*, 611–624. <https://doi.org/10.1016/j.foodhyd.2016.10.015>.
- (154) Gidley, M. J.; Bociek, S. M. <sup>13</sup>C CP/MAS NMR Studies of Amylose Inclusion Complexes, Cyclodextrins, and the Amorphous Phase of Starch Granules: Relationships Between Glycosidic Linkage Conformation and Solid-State <sup>13</sup>C Chemical Shifts. *J. Am. Chem. Soc.* **1988**, *110* (12), 3820–3829. <https://doi.org/10.1021/ja00220a016>.
- (155) Zhang, L.; Zhao, Y.; Hu, W.; Qian, J. Y.; Ding, X. L.; Guan, C. R.; Lu, Y. Q.; Cao, Y. Multi-Scale Structures of Cassava and Potato Starch Fractions Varying in Granule Size. *Carbohydr. Polym.* **2018**, *200* (March), 400–407. <https://doi.org/10.1016/j.carbpol.2018.08.022>.
- (156) Ren, N.; Ma, Z.; Xu, J.; Hu, X. Insights into the Supramolecular Structure and Techno-Functional Properties of Starch Isolated from Oat Rice Kernels Subjected to Different Processing Treatments. *Food Chem.* **2020**, *317* (October 2019), 126464. <https://doi.org/10.1016/j.foodchem.2020.126464>.
- (157) Feng, T.; Li, M.; Zhou, J.; Zhuang, H.; Chen, F.; Ye, R.; Campanella, O.; Fang, Z. Application of Molecular Dynamics Simulation in Food Carbohydrate Research—a Review. **2015**. <https://doi.org/10.1016/j.ifset.2015.06.015>.
- (158) Cheng, L.; Feng, T.; Zhang, B.; Zhu, X.; Hamaker, B.; Zhang, H.; Campanella, O. A Molecular Dynamics Simulation Study on the Conformational Stability of Amylose-Linoleic Acid Complex in Water. *Carbohydr. Polym.* **2018**, *196*, 56–65. <https://doi.org/10.1016/j.carbpol.2018.04.102>.
- (159) Cheng, L.; Zhu, X.; Hamaker, B. R.; Zhang, H.; Campanella, O. H. Complexation Process of Amylose under Different Concentrations of Linoleic Acid Using Molecular Dynamics Simulation. *Carbohydr. Polym.* **2019**. <https://doi.org/10.1016/j.carbpol.2019.04.013>.
- (160) Gao, W.; Liu, Y.; Jing, G.; Li, K.; Zhao, Y.; Sha, B.; Wang, Q.; Wu, D. Rapid and

- Efficient Crossing Blood-Brain Barrier: Hydrophobic Drug Delivery System Based on Propionylated Amylose Helix Nanoclusters. *Biomaterials* **2017**, *113*, 133–144. <https://doi.org/10.1016/j.biomaterials.2016.10.045>.
- (161) López, C. A.; De Vries, A. H.; Marrink, S. J. Amylose Folding under the Influence of Lipids. *Carbohydr. Res.* **2012**. <https://doi.org/10.1016/j.carres.2012.10.007>.
- (162) Bhopatkar, D.; Feng, T.; Chen, F.; Zhang, G.; Carignano, M.; Park, S. H.; Zhuang, H.; Campanella, O. H.; Hamaker, B. R. Self-Assembled Nanoparticle of Common Food Constituents That Carries a Sparingly Soluble Small Molecule. *J. Agric. Food Chem.* **2015**, *63* (17), 4312–4319. <https://doi.org/10.1021/acs.jafc.5b00037>.
- (163) Kaneko, Y.; Ueno, K.; Yui, T.; Nakahara, K.; Kadokawa, J. I. Amylose's Recognition of Chirality in Polylactides on Formation of Inclusion Complexes in Vine-Twining Polymerization. *Macromol. Biosci.* **2011**, *11* (10), 1407–1415. <https://doi.org/10.1002/mabi.201100133>.
- (164) Zhang, X.; Wang, Y. S.; Chen, H. H. Effect of Annealing Temperature on Morphology and Physicochemical Properties of Cornstarch Complexed with Oleic Acid and Molecular Dynamics Simulation. *Cereal Chem.* **2019**, *96* (4), 668–677. <https://doi.org/10.1002/cche.10163>.
- (165) Sang, S.; Chen, Y.; Zhu, X.; Narsimhan, G.; Hu, Q.; Jin, Z.; Xu, X. Effect of Egg Yolk Lipids on Structure and Properties of Wheat Starch in Steamed Bread. *J. Cereal Sci.* **2019**, *86* (October 2018), 77–85. <https://doi.org/10.1016/j.jcs.2019.01.009>.
- (166) Chen, L.; Xiong, Z.; Din, Z. ud; Nawaz, A.; Xiong, H.; Cai, J. Interfacial Modification of Starch at High Concentration by Sodium Dodecylsulfate as Revealed by Experiments and Molecular Simulation. *J. Mol. Liq.* **2020**, *310*, 113190. <https://doi.org/10.1016/j.molliq.2020.113190>.
- (167) Wang, Y. S.; Liu, W. H.; Zhang, X.; Chen, H. H. Preparation of VII-Type Normal Cornstarch-Lauric Acid Complexes with High Yield and Stability Using a Combination Treatment of Debranching and Different Complexation Temperatures. *Int. J. Biol. Macromol.* **2020**, *154*, 456–465. <https://doi.org/10.1016/j.ijbiomac.2020.03.142>.

- (168) Toukach, F. V.; Ananikov, V. P. Recent Advances in Computational Predictions of NMR Parameters for the Structure Elucidation of Carbohydrates: Methods and Limitations. *Chem. Soc. Rev.* **2013**, *42* (21), 8376–8415. <https://doi.org/10.1039/c3cs60073d>.
- (169) Kasat, R. B.; Wang, N. H. L.; Franses, E. I. Effects of Backbone and Side Chain on the Molecular Environments of Chiral Cavities in Polysaccharide-Based Biopolymers. *Biomacromolecules* **2007**, *8* (5), 1676–1685. <https://doi.org/10.1021/bm070006h>.
- (170) Tang, S.; Zhao, Z.; Chen, G.; Su, Y.; Lu, L.; Li, B.; Liang, D.; Jin, R. Fabrication of Ampicillin/Starch/Polymer Composite Nanofibers with Controlled Drug Release Properties by Electrospinning. *J. Sol-Gel Sci. Technol.* **2016**, *77* (3), 594–603. <https://doi.org/10.1007/s10971-015-3887-x>.
- (171) Liang, J.; Su, Q.; Zhao, Y.; Wang, Y. Theoretical Insights into Three Types of Oxidized Starch-Based Adhesives: Chemical Stability, Water Resistance, and Shearing Viscosity from a Molecular Viewpoint. *J. Chem.* **2016**, *2016*. <https://doi.org/10.1155/2016/2369739>.
- (172) Schnupf, U.; Momany, F. A. DFT Energy Optimization of a Large Carbohydrate: Cyclomaltohexaicosaoase (CA-26). *J. Phys. Chem. B* **2012**, *116* (23), 6618–6627. <https://doi.org/10.1021/jp208927v>.
- (173) Okuda, M.; Hiramatsu, T.; Yasuda, M.; Ishigaki, M.; Ishigaki, M.; Ozaki, Y.; Ozaki, Y.; Hayashi, M.; Tominaga, K.; Tominaga, K.; Chatani, E. Theoretical Modeling of Electronic Structures of Polyiodide Species Included in  $\alpha$ -Cyclodextrin. *J. Phys. Chem. B* **2020**, *124* (20), 4089–4096. <https://doi.org/10.1021/acs.jpcc.0c01749>.
- (174) *Global Tuberculosis Report from the World Health Organization*; 2019. <https://doi.org/10.16309/j.cnki.issn.1007-1776.2003.03.004>.
- (175) Bernstein, J.; Lott, W. A.; Steinberg, B. A.; Yale, H. L. Chemotherapy of Experimental Tuberculosis. V. Isonicotinic Acid Hydrazide (Nydrasid) and Related Compounds. *Am. Rev. Tuberc.* **1952**, *65* (4), 357–364.
- (176) Hobby, G. L.; Lenert, T. F. The Antimycobacterial Activity of Rifampin. *Am. Rev. Respir. Dis.* **1968**, *97* (4), 713–714.

- (177) Malone, L.; Schurr, A.; Lindh, H.; McKenzie, D.; Kiser, J. S.; Williams, J. H. The Effect of Pyrazinamide (Aldinamide) on Experimental Tuberculosis in Mice. *Am. Rev. Tuberc.* **1952**, *65* (5), 511–518.
- (178) Thomas, J. P.; Baughn, C. O.; Wilkinson, R. G.; Shepherd, R. G. A New Synthetic Compound with Antituberculous Activity in Mice: Ethambutol (Dextro-2,2'-(Ethylenediimino)-Di-1-Butanol). *Am. Rev. Respir. Dis.* **1961**, *83* (6), 891–893.
- (179) Kalscheuer, R.; Palacios, A.; Anso, I.; Cifuentes, J.; Anguita, J.; Jacobs, W. R.; Guerin, M. E.; Prados-Rosales, R. The Mycobacterium Tuberculosis Capsule: A Cell Structure with Key Implications in Pathogenesis. *Biochem. J.* **2019**, *476* (14), 1995–2016. <https://doi.org/10.1042/BCJ20190324>.
- (180) Daffé, M.; Marrakchi, H. Unraveling the Structure of the Mycobacterial Envelope. *Microbiol. Spectr.* **2019**, *7* (4), 1087–1095. <https://doi.org/10.1128/microbiolspec.gpp3-0027-2018>.
- (181) Asselineau, C.; Asselineau, J. Trehalose-Containing Glycolipids. *Prog. Chem. Fats Other Lipids* **1978**, *16*, 59–99.
- (182) Daffe, M.; Laneelle, M. A. Distribution of Phthiocerol Diester, Phenolic Mycosides and Related Compounds in Mycobacteria. *J. Gen. Microbiol.* **1988**, *134* (7), 2049–2055. <https://doi.org/10.1099/00221287-134-7-2049>.
- (183) Gagliardi, M. C.; Lemassu, A.; Teloni, R.; Mariotti, S.; Sargentini, V.; Pardini, M.; Daffé, M.; Nisini, R. Cell Wall-Associated Alpha-Glucan Is Instrumental for Mycobacterium Tuberculosis to Block CD1 Molecule Expression and Disable the Function of Dendritic Cell Derived from Infected Monocyte. *Cell. Microbiol.* **2007**. <https://doi.org/10.1111/j.1462-5822.2007.00940.x>.
- (184) Arbues, A.; Lugo-Villarino, G.; Neyrolles, O.; Guilhot, C.; Astarie-Dequeker, C. Playing Hide-and-Seek with Host Macrophages through the Use of Mycobacterial Cell Envelope Phthiocerol Dimycocerosates and Phenolic Glycolipids. *Front. Cell. Infect. Microbiol.* **2014**, *4* (December), 1–7. <https://doi.org/10.3389/fcimb.2014.00173>.
- (185) Cambier, C. J.; Takaki, K. K.; Larson, R. P.; Hernandez, R. E.; Tobin, D. M.; Urdahl, K. B.; Cosma, C. L.; Ramakrishnan, L. Mycobacteria Manipulate

- Macrophage Recruitment through Coordinated Use of Membrane Lipids. *Nature* **2014**, *505* (7482), 218–222. <https://doi.org/10.1038/nature12799>.
- (186) Cambier, C. J.; Falkow, S.; Ramakrishnan, L. Host Evasion and Exploitation Schemes of Mycobacterium Tuberculosis. *Cell* **2014**, *159* (7), 1497–1509. <https://doi.org/10.1016/j.cell.2014.11.024>.
- (187) Ishikawa, E.; Ishikawa, T.; Morita, Y. S.; Toyonaga, K.; Yamada, H.; Takeuchi, O.; Kinoshita, T.; Akira, S.; Yoshikai, Y.; Yamasaki, S. Direct Recognition of the Mycobacterial Glycolipid, Trehalose Dimycolate, by C-Type Lectin Mincle. *J. Exp. Med.* **2009**, *206* (13), 2879–2888. <https://doi.org/10.1084/jem.20091750>.
- (188) Jessica Indrigo<sup>1</sup>, R. L. H. J. and J. K. A. Cord Factor ( $\alpha,\alpha$ -Trehalose 6,6'-Dimycolate) Inhibits Fusion between Phospholipid Vesicles. *Proc. Natl. Acad.* **1991**, *88* (February), 737–740.
- (189) Dinadayala, P.; Sambou, T.; Daffé, M.; Lemassu, A. Comparative Structural Analyses of the  $\alpha$ -Glucan and Glycogen from Mycobacterium Bovis. *Glycobiology* **2008**, *18* (7), 502–508. <https://doi.org/10.1093/glycob/cwn031>.
- (190) Sambou, T.; Dinadayala, P.; Stadthagen, G.; Barilone, N.; Bordat, Y.; Constant, P.; Levillain, F.; Neyrolles, O.; Gicquel, B.; Lemassu, A.; Daffé, M.; Jackson, M. Capsular Glucan and Intracellular Glycogen of Mycobacterium Tuberculosis: Biosynthesis and Impact on the Persistence in Mice. *Mol. Microbiol.* **2008**. <https://doi.org/10.1111/j.1365-2958.2008.06445.x>.
- (191) Lemassu, A.; Ortalo-magne, A.; Bardou, F.; Silve, G.; Lanéelle, M. A.; Daffe, M. Extracellular and Surface-Exposed Polysaccharides of Non-Tuberculous Mycobacteria. *Microbiology* **1996**, *142*, 1513–1520.
- (192) Goesaert, H.; Brijs, K.; Veraverbeke, W. S.; Courtin, C. M.; Gebruers, K.; Delcour, J. A. Wheat Flour Constituents: How They Impact Bread Quality, and How to Impact Their Functionality. *Trends Food Sci. Technol.* **2005**, *16* (1–3), 12–30. <https://doi.org/10.1016/j.tifs.2004.02.011>.

## **Méthodes**



# Chapitre 1. La Résonance Magnétique Nucléaire

La plus grande partie de ce travail de thèse repose sur des expériences de Résonance Magnétique Nucléaire (RMN), qu'elles soient expérimentales ou théoriques. Il apparaît donc indispensable de décrire la méthodologie sous-jacente. Par soucis d'homogénéité, la partie théorique quant à elle sera traitée dans le second chapitre de cette partie.

Nous commencerons ici par un bref historique dans lequel nous reverrons en quelques dates comment nous en sommes arrivés à l'état de l'art actuel et en suivant nous nous attarderons sur les différents principes sur lesquels reposent cette méthode avant de finir avec son formalisme mathématique.

## 1. Historique

La RMN est une méthode expérimentale dont l'histoire est parsemée de découvertes scientifiques majeures. S'il est difficile de trouver un réel point de départ à l'historique de cette méthode, nous considérerons que cela commence avec la description du moment cinétique nucléaire. En effet, cette grandeur, décrite en 1924 par W. Pauli<sup>1</sup>, est au centre de toute la théorie utilisée dans la RMN.

Ce n'est que 14 ans plus tard qu'une première expérience concernant l'étude des propriétés magnétiques des noyaux est publiée. Il s'agit de l'expérience de I. Rabi<sup>2,3</sup>, publiée en 1938, mesurant le moment magnétique nucléaire de différents noyaux.

Jusqu'ici nous ne parlions pas de RMN, il s'agissait simplement d'études sur les propriétés intrinsèques des noyaux atomiques, réalisées sur des échantillons gazeux à haute pression. C'est en 1946 que F. Bloch<sup>4</sup> et E. Purcell<sup>5</sup>, chacun de leur côté, établirent les fondements de cette méthode. Purcell fut le premier à décrire ce phénomène dans un matériau solide, la paraffine, prouvant ainsi qu'il existait dans un autre état de la matière. Bloch quant à lui, est connu pour son travail sur ce qu'il appelait « l'induction nucléaire ». Dans l'article portant ce titre, il y définit le formalisme permettant de décrire le comportement temporel de l'aimantation nucléaire. C'est dans

cet article que sont définies pour la première fois les notions de relaxation longitudinale et relaxation transversale, notions de première importance dans cette méthode.

A partir de cette période, de nombreuses découvertes, concernant la RMN en tant que telle, sont apparues. Ici nous nous contenterons de citer les travaux de W. G. Proctor et F. C. Yu<sup>6</sup>, W. C. Dickinson<sup>7</sup>, E. R. Andrew<sup>8</sup> et R. R. Ernst<sup>9</sup> qui préfigurent les méthodes utilisées au cours de cette thèse.

Commençons par Proctor et Yu, qui simultanément avec Dickinson, découvrirent le concept de déplacement chimique en 1950. Ils notèrent que la fréquence de résonance d'un noyau donné dépendait intrinsèquement de son environnement chimique. C'est cette découverte fondamentale qui fit entrer la RMN dans les méthodes applicables afin d'étudier la structure des molécules. M. Goldman rapportera par ailleurs la boutade de Félix Bloch à propos de ces découvertes « Quand les chimistes pénètrent dans un domaine, c'est qu'il est temps d'en sortir ! »<sup>10</sup>.

Suite à cela, Andrew établit en 1959 le concept de rotation à l'angle magique ou « magic angle spinning », permettant d'annuler les composantes anisotropes des couplages dipolaire, quadripolaire et du déplacement chimique (notions qui seront redéfinies plus bas). Cette méthode nous permet notamment aujourd'hui de réaliser des spectres avec une meilleure résolution sur des échantillons à l'état solide.

Enfin, nous devons à Ernst la RMN par transformée de Fourier, décrite en 1966. Elle est aujourd'hui utilisée partout et elle permit notamment l'arrivée à la RMN multidimensionnelle.

Comme précisé au début de cet historique, ces découvertes scientifiques font partie des avancées majeures du siècle dernier. Si on peut compter parmi les scientifiques cités pas moins de cinq prix Nobel (quatre en physique et un en chimie) décernés pour leurs travaux réalisés dans ce domaine, il nous faut aussi rappeler que ces méthodes ont aussi amenés des avancées importantes en biologie et en médecine. En effet, K. Wüthrich<sup>11,12</sup> démontrera dans les années 1970-1980 qu'il est possible d'utiliser la RMN pour étudier la structure tridimensionnelle des protéines. Au même moment, P. Lauterbur et P. Mansfield développeront indépendamment les prémices de l'imagerie par résonance magnétique (IRM). Ces trois scientifiques seront récompensés d'un prix Nobel en 2002, et 2003 respectivement.

Nous allons passer maintenant au formalisme permettant de décrire la RMN d'un point de vue mathématique. Cela nous permettra aussi de définir les expériences utilisées en laboratoire afin d'étudier les échantillons de cette thèse.

## 2. Principe et fonctionnement

Depuis Pauli, nous savons que la plupart des noyaux possèdent un spin nucléaire noté  $I$ . Voici le traditionnel tableau présentant les spins nucléaires de certains noyaux d'intérêt, principalement ceux rencontrés au cours de cette thèse :

Tableau 4: Abondance naturelle, Spin nucléaire et rapport gyromagnétique de noyaux communs au monde biologique.

Noyau	Abondance naturelle (%)	$I$	$\gamma$ ( $/10^6 \text{ s}^{-1} \text{ T}^{-1}$ )
$^1\text{H}$	99,98	1/2	267,52
$^2\text{H}$	0,02	1	41,06
$^{12}\text{C}$	98,93	0	-
$^{13}\text{C}$	1,07	1/2	67,28
$^{14}\text{N}$	99,63	1	19,34
$^{15}\text{N}$	0,37	1/2	-27,13
$^{16}\text{O}$	99,96	0	-
$^{17}\text{O}$	0,04	5/2	-36,28
$^{31}\text{P}$	~100,00	1/2	108,39

Le phénomène de RMN ne peut être observé que sur des noyaux dont le spin nucléaire est non-nul. La présence d'un spin nucléaire engendre un moment cinétique de spin noté  $\vec{I}$ . Pour chaque noyau  $j$ , on définit le moment magnétique de la manière suivante :

$$\vec{\mu}_j = \gamma_j \hbar \vec{I}_j \quad (2.1)$$

où  $\gamma_j$  est le rapport gyromagnétique. Comme on peut le voir dans le Tableau 4, cette grandeur peut prendre des valeurs positives comme négatives. De ce fait,  $\vec{\mu}$  est un vecteur parallèle ou antiparallèle à  $\vec{I}$ .

Dans un milieu quelconque, les moments magnétiques de chaque atome sont distribués de manière aléatoire. De plus, étant soumis à l'agitation thermique due à la température, ils sont en constante réorientation au sein du milieu.

Pour comprendre ce qu'il se passe dans un échantillon soumis à un champs magnétique intense, il nous faut tout d'abord redéfinir une notion fondamentale : l'effet Zeeman. A partir de maintenant, nous considérerons que le champ magnétique est appliqué selon l'axe z, sauf mention contraire.

Nous parlons ici de particules soumises aux lois de la mécanique quantique. Un état nucléaire de spin  $I$  est dit dégénéré  $(2I+1)$  fois, il comprend  $2I+1$  sous états de même énergie. Sous l'action d'un champ magnétique externe, la dégénérescence de ces  $I$  états est levée, proportionnellement à l'intensité de ce dernier (Figure 21).

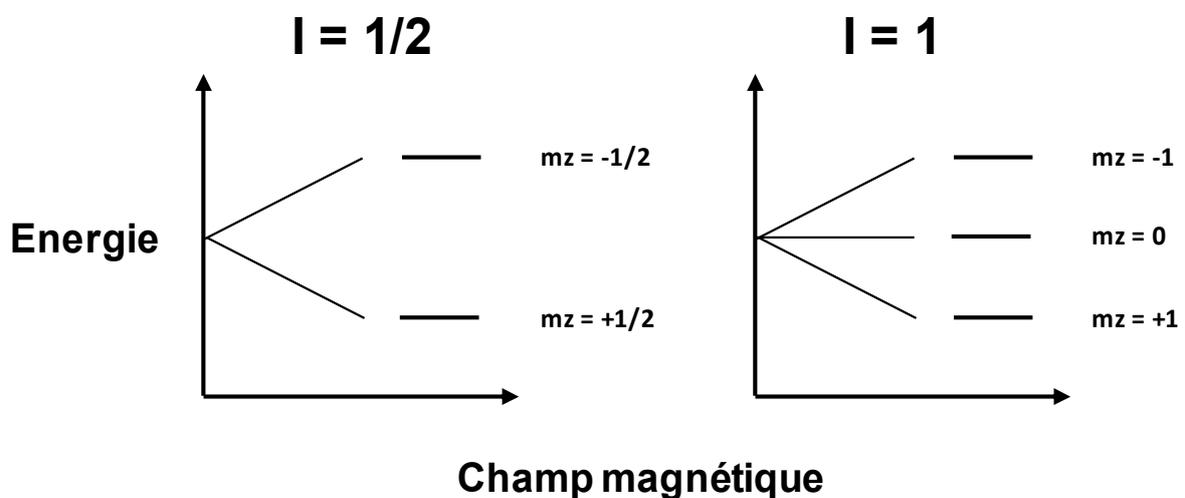


Figure 21: Effet Zeeman appliqué aux sous-états de spin d'un noyau de spin 1/2 et d'un noyau de spin 1

Cet effet porte le nom d'effet Zeeman. Imaginons maintenant un ensemble d'atomes/molécules soumis à un champ magnétique extérieur  $\vec{B}_0$ . Chaque moment magnétique individuel va entrer en précession autour de l'axe de  $\vec{B}_0$ , et rester dans cet état, là où ils se réorientaient constamment en absence de ce dernier. La vitesse de précession autour de cet axe, pour un noyau  $j$ , est caractérisée par la fréquence de Larmor :

$$\nu_0 = \frac{\gamma_j}{2\pi} \|\vec{B}_0\| \quad (2.2)$$

Nous savons que toutes nos particules présentent  $2I+1$  sous états non dégénérés, mais concrètement que représentent-ils ? En réalité, nous pouvons considérer qu'il s'agit de la projection du moment magnétique de chaque particule sur l'axe  $z$ . Cette projection dépend uniquement de l'angle entre  $\vec{\mu}$  et  $\vec{B}_0$  au moment précis de son application. On peut calculer l'énergie magnétique de la façon suivante :

$$E_{mag} = h\nu_0 = -\vec{\mu} \cdot \vec{B}_0 \quad (2.3)$$

Cette équation nous donne une information primordiale : si on considère un rapport gyromagnétique positif, l'énergie magnétique due à un moment magnétique en présence d'un champ magnétique extérieur se minimise si les deux vecteurs « pointent » dans le même sens, et son minimum absolu apparaît quand ils sont complètement parallèles. Attention, rappelons que l'énergie est ici discrétisée, et non pas continue.

On peut donc imaginer que les  $2I+1$  états décrits correspondent à des états, dont l'énergie représente l'alignement du vecteur de magnétisation avec le champ  $\vec{B}_0$ . Les états de plus haute énergie représenteront ceux qui « pointent » dans une direction opposée à ce dernier. L'agitation thermique du milieu permet de peupler ces états en suivant une distribution de Boltzman :

$$\frac{N_1}{N_0} = e^{-\frac{\Delta E}{kT}} \quad (2.4)$$

Et la différence d'énergie entre les deux états considérés est la suivante :

$$\Delta E = E_1 - E_0 = \gamma_i \hbar \|\vec{B}_0\| \quad (2.5)$$

Si les contributions de chaque moment magnétique  $\vec{\mu}$  vont statistiquement se compenser dans le plan perpendiculaire à  $\vec{B}_0$ , la différence de population entre les  $2I + 1$  niveaux considérés implique un plus grand nombre de contributions qui « pointent » dans le même sens que  $\vec{B}_0$ . De ce fait, en sommant maintenant tous les moments magnétiques présents en solution, on se retrouve avec un moment magnétique total  $\vec{M}$  aligné avec le champ magnétique extérieur  $\vec{B}_0$ .

Le moment magnétique total ne peut pas être mesuré quand il est aligné sur l'axe  $z$ , choisi par convention comme l'axe du champ magnétique extérieur  $\vec{B}_0$ . Pour pouvoir être mesuré, il doit être basculé dans ce qu'on appelle le plan transverse, orthogonal à  $z$ . Ce basculement n'est possible qu'en présence d'un autre champ  $\vec{B}_1$ , perpendiculaire à  $\vec{B}_0$ , qui tourne autour de ce dernier à la fréquence de Larmor du noyau étudié (Figure 22).

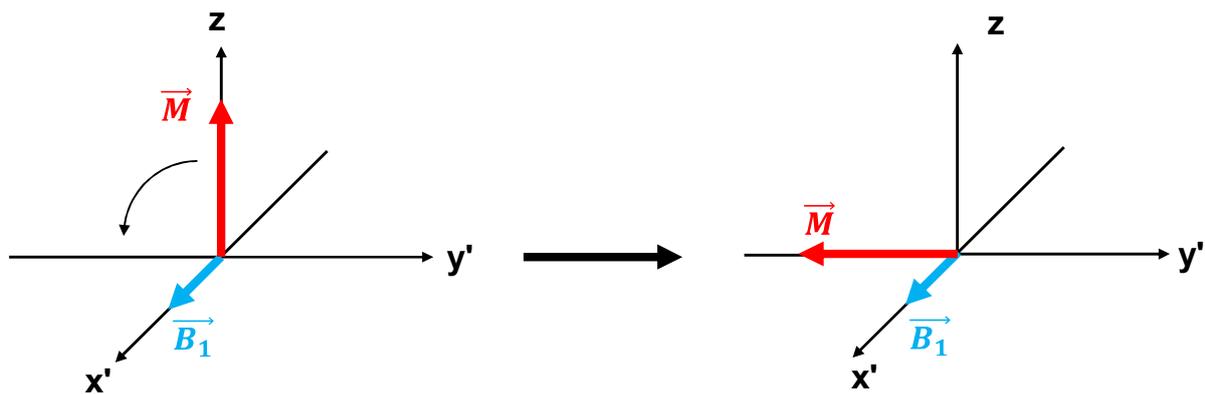


Figure 22: Représentation vectorielle du basculement dans le plan transverse. Adapté de Levitt, 2008<sup>28</sup>.

Ici, le référentiel utilisé  $\{x', y', z\}$  est appelé référentiel tournant et tourne autour de  $z$  à la même vitesse que  $\vec{B}_1$ . On utilise ce référentiel pour simplifier la représentation de ce phénomène. Le moment magnétique total va donc maintenant entrer en précession autour de  $\vec{B}_0$ , et sa composante  $\{z\}$  va progressivement diminuer au profit de ses composantes  $\{x, y\}$ . Lorsque le champ magnétique  $\vec{B}_1$  est retiré, le moment magnétique total va se réorienter selon l'axe  $z$ . Ce retour à l'équilibre est appelé relaxation. On différencie deux types de relaxation : la relaxation longitudinale, dont le temps caractéristique est appelé  $T_1$ , et la relaxation transverse, dont le temps caractéristique est appelé  $T_2$ . La relaxation longitudinale concerne la composante  $\{z\}$

du moment magnétique et la relaxation transverse concerne les composantes  $\{x, y\}$ . Elles sont régies par les équations de Bloch :

$$\text{Relaxation Transverse : } \begin{cases} M_x = M_{ini} \sin(\omega_0 t) e^{-\frac{t}{T_2}} \\ M_y = -M_{ini} \cos(\omega_0 t) e^{-\frac{t}{T_2}} \end{cases} \quad (2.6)$$

$$\text{Relaxation longitudinale : } M_z = M_0(1 - e^{-\frac{t}{T_1}}) \quad (2.7)$$

Ici, malgré le caractère contre-intuitif de cette information, il faut bien comprendre que le retour à l'équilibre ne correspond pas aux mêmes phénomènes physiques selon l'axe considéré. La relaxation longitudinale correspond à un retour à l'équilibre, dans lequel les atomes vont revenir aux  $2I+1$  états décrits plus haut. La relaxation transversale est due à la perte de cohérence entre les précessions de tous les moments magnétiques  $\vec{\mu}$  dans le plan  $xy$ . On peut tracer leur comportement au cours du temps :

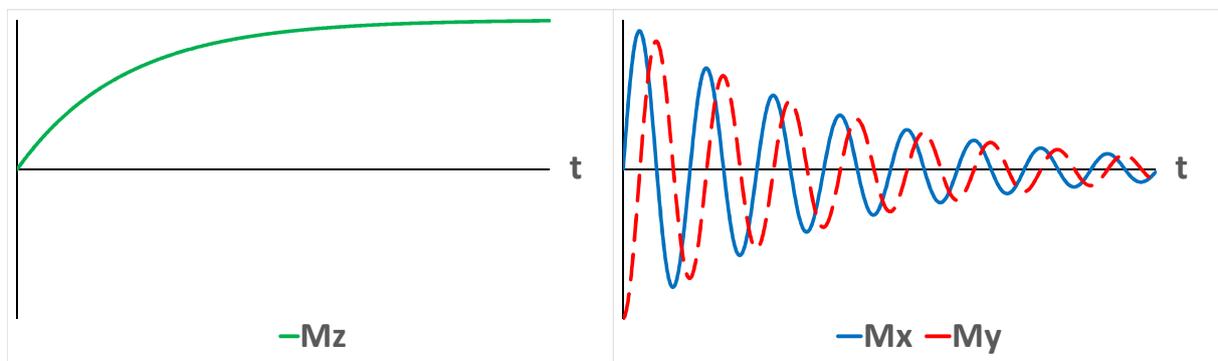


Figure 23: Comportement temporelle des aimantations transverses et longitudinales. Adapté de Levitt, 2008

De tels comportements indiquent que la composante dans le plan transverse va tourner autour de  $\vec{B}_0$  en disparaissant peu à peu. C'est cette dernière qui permet le signal RMN. En effet, afin de générer le champ tournant  $\vec{B}_1$ , en pratique on utilise un courant alternatif, appelé pulsation radiofréquence, que l'on fait passer dans un solénoïde perpendiculaire à  $\vec{B}_0$ . Une fois que le moment magnétique a atteint le plan transverse, la précession de ce dernier induit un nouveau champ magnétique tournant

très faible. Ce dernier génère néanmoins un courant alternatif dans le solénoïde de départ, qui est observable, et qui s'appelle Free Induction Decay (FID).

Les FID générés expérimentalement sont des combinaisons de plusieurs signaux tels que ceux présentés Figure 23. Ces signaux présentent différentes fréquences d'oscillation et vitesse de décohérence, rendant l'analyse du FID en tant que tel particulièrement fastidieuse. Il existe cependant un outil mathématique, appelé transformée de Fourier, qui permet de séparer les signaux superposés au sein du FID en passant de l'espace temporel à l'espace des fréquences. A chaque fréquence sera associé un pic, et le résultat de cette opération aboutit aux spectres RMN couramment utilisés de nos jours.

Voici dans les grandes lignes le fonctionnement de la RMN telle qu'on la connaît aujourd'hui. Nous allons passer maintenant à la description du formalisme mathématique sous-jacent, qui nous permettra d'aborder les quelques notions fondamentales utilisées dans cette thèse.

### 3. Formalisme mathématique

#### I. Hamiltonien de spin :

La dynamique temporelle d'un spin nucléaire est décrite à travers l'équation de Schrödinger dépendante du temps :

$$i \frac{d}{dt} |\Psi_{spin}(t)\rangle = \hat{\mathcal{H}}_{spin} |\Psi_{spin}(t)\rangle \quad (3.8)$$

On peut définir l'opérateur Hamiltonien de spin comme étant la somme de deux composantes qui décrivent les effets internes ou externes subis par le noyau considéré:

$$\hat{\mathcal{H}}_{spin} = \hat{\mathcal{H}}_{Int} + \hat{\mathcal{H}}_{Ext} \quad (3.9)$$

Les interactions externes sont dues au champ magnétique statique  $\vec{B}_0$  et au champ radiofréquence  $\vec{B}_1$  :

$$\hat{\mathcal{H}}_{Ext} = \hat{\mathcal{H}}_{stat} + \hat{\mathcal{H}}_{RF} \quad (3.10)$$

Les interactions internes sont dues à de multiples effets : le déplacement chimique, le couplage dipolaire, le couplage scalaire et le couplage quadripolaire :

$$\hat{\mathcal{H}}_{Int} = \hat{\mathcal{H}}_{DC} + \hat{\mathcal{H}}_D + \hat{\mathcal{H}}_J + \hat{\mathcal{H}}_Q \quad (3.11)$$

A ce stade, nous devons préciser une chose pour la bonne compréhension du lecteur. Les formes mathématiques des composantes de cet Hamiltonien sont compliquées. Il existe cependant deux approximations permettant de les simplifier : l'approximation des champs forts aussi appelée approximation séculaire et l'approximation tenant compte des réorientations rapides de l'échantillon. L'objectif de ce chapitre étant de décrire de manière pédagogique le formalisme sous-jacent, ces deux approximations ne seront pas décrites en détail. Leur utilisation sera cependant précisée quand cela est nécessaire.

## II. Interactions externes

### a. Le champ statique

Le champ statique  $\vec{B}_0$  induit l'effet Zeeman décrit précédemment. Comme précisé plus haut, le champ statique étant par convention aligné selon l'axe z, l'Hamiltonien résultant pour un noyau j sera le suivant :

$$\hat{\mathcal{H}}_{stat} = -\gamma_j \hbar B_0 \hat{I}_{jz} \quad (3.12)$$

Il en résultera les  $2I+1$  niveaux d'énergie associés.

### b. Le champ radiofréquence

Le champ radiofréquence a une forme un peu particulière au cours du temps. En effet, rappelons-nous qu'il s'agit d'un champ qui est en rotation lui-même autour du champ statique. De ce fait sa projection sur l'axe x ou y est la suivante :

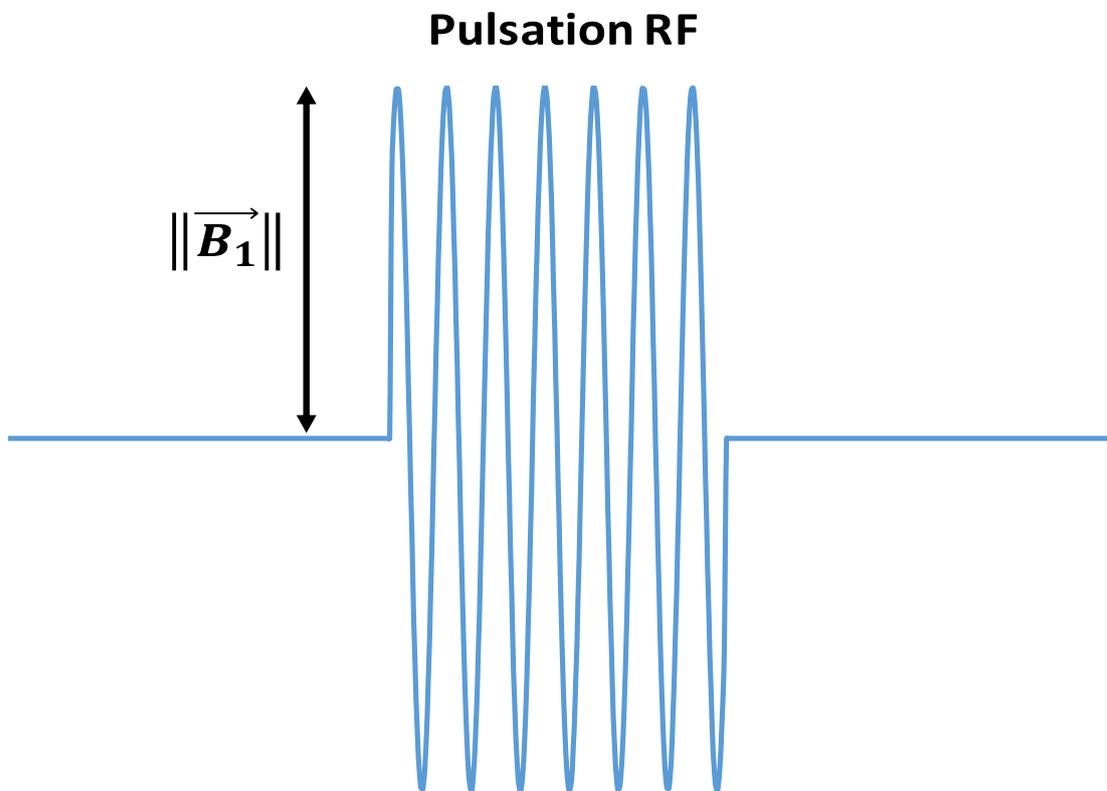


Figure 24: Projection sur l'axe x ou y dans le repère du laboratoire du champ radiofréquence. Adapté de Levitt, 2008

La description mathématique d'un tel champ nous amène au résultat suivant :

$$\hat{\mathcal{H}}_{RF} = \gamma_j \hbar B_1 (\hat{I}_{jx} \cos(\omega_{RF}t) + \hat{I}_{jy} \sin(\omega_{RF}t)) \quad (3.13)$$

Nous avons donc finalement une contribution selon l'axe z due au champ statique et des composantes sur les axes x et y qui dépendent exclusivement du champ radiofréquence.

### III. Interactions internes

#### a. Le déplacement chimique

Quand une molécule est plongée dans un champ magnétique, les électrons de cette dernière vont se mettre à circuler au sein du nuage électronique. Cette circulation va induire un champ magnétique local autour de chaque noyau. Deux noyaux ayant un environnement chimique local différent (ex : CH<sub>3</sub> et CH<sub>2</sub>) présentent une densité électronique différente. Le champ magnétique induit par la circulation des électrons sera différent pour chaque groupement : c'est ce qu'on appelle le déplacement chimique, et son interaction sera notée  $\hat{H}_{DC}$ . Le champ magnétique local sera donc la somme du champ magnétique externe, et du champ magnétique induit :

$$\vec{B}_{local} = \vec{B}_0 + \vec{B}_{induit} \quad (3.14)$$

On définit le champ magnétique induit comme suivant :

$$\vec{B}_{induit} = -\sigma \cdot \vec{B}_0 \quad (3.15)$$

Le signe négatif ici est dû au fait que  $\vec{B}_{induit}$  est en général opposé à  $\vec{B}_0$ .  $\sigma$  est ce qu'on appelle le tenseur d'écrantage chimique représenté par une matrice 3x3 :

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \quad (3.16)$$

Ce tenseur est au centre de cette thèse. C'est la grandeur qui permettra d'effectuer les comparaisons entre résultats théoriques et expérimentaux. On peut remarquer que comme  $\vec{B}_0$  est orienté selon z,  $\vec{B}_{induit}$  présentera des composantes en x et en y dépendantes de  $\sigma_{xz}$  et  $\sigma_{yz}$ . Cela implique que le champ magnétique induit n'est pas forcément parallèle au champ extérieur appliqué. Il existe cependant un repère orthogonal, dont les axes sont appelés axes propres. Si  $\vec{B}_0$  est appliqué de façon parallèle à l'un de ces axes, le champ magnétique induit sera parallèle à  $\vec{B}_0$ . Ces trois axes propres sont obtenus en diagonalisant  $\sigma$  :

$$\sigma_{diag} = \begin{pmatrix} \sigma_{XX} & 0 & 0 \\ 0 & \sigma_{YY} & 0 \\ 0 & 0 & \sigma_{ZZ} \end{pmatrix} \quad (3.17)$$

Aux valeurs propres  $\{\sigma_{XX}, \sigma_{YY}, \sigma_{ZZ}\}$  correspondent respectivement les vecteurs propres  $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$  qui définissent le repère propre (Figure 25).

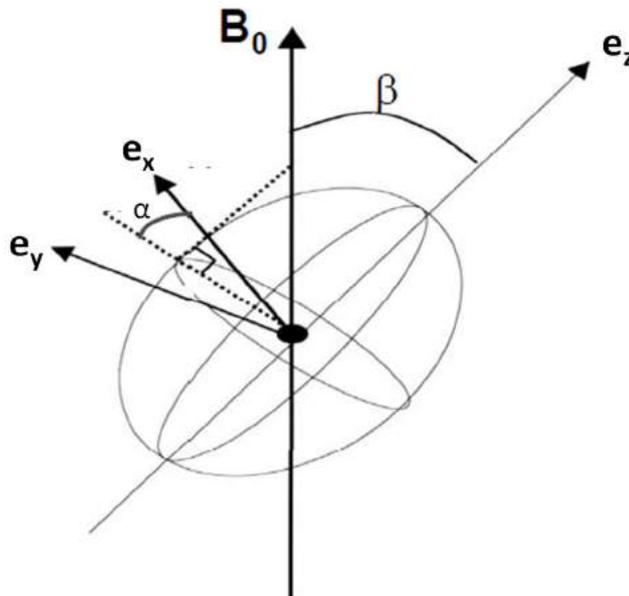


Figure 25: Définition des angles  $\alpha$  et  $\beta$ , angle polaires définissant l'orientation de  $\vec{B}_0$  dans le repère propre définis par  $\{\vec{e}_x, \vec{e}_y, \vec{e}_z\}$

Ce tenseur permet notamment de calculer trois paramètres expliquant les spectres expérimentaux, qui sont l'écrantage chimique isotrope, l'anisotropie réduite et le paramètre d'asymétrie :

$$\sigma_{iso} = \frac{1}{3}(\sigma_{XX} + \sigma_{YY} + \sigma_{ZZ}) \quad (3.18)$$

$$\Delta\sigma = \frac{3}{2}(\sigma_{ZZ} - \sigma_{iso}) \quad (3.19)$$

$$\eta_\sigma = \frac{\sigma_{YY} - \sigma_{XX}}{\Delta\sigma} \quad (3.20)$$

On notera que, selon la convention que nous avons choisie,  $\sigma_{XX}$ ,  $\sigma_{YY}$  et  $\sigma_{ZZ}$  sont classés de la façon suivante<sup>13-15</sup> :

$$|\sigma_{ZZ} - \sigma_{iso}| > |\sigma_{XX} - \sigma_{iso}| > |\sigma_{YY} - \sigma_{iso}| \quad (3.21)$$

Ces grandeurs nous seront utiles ici afin de définir l'Hamiltonien de déplacement chimique. Sa forme générale est la suivante :

$$\hat{\mathcal{H}}_{DC} = -\vec{\mu} \cdot \vec{B}_{local} = -\gamma\hbar\vec{I} \cdot \vec{B}_0 + \gamma\hbar\vec{\sigma} \cdot \vec{B}_0 \quad (3.22)$$

On peut montrer, grâce à l'approximation des champs forts, que seule la composante selon l'axe z est importante dans notre cas, ce qui amène à l'équation suivante :

$$\hat{\mathcal{H}}_{DC} = -\gamma\hbar B_0(1 - \sigma_{ZZ})\hat{I}_z \quad (3.23)$$

Attention, ici nous voyons pour la première fois apparaître le terme  $\hat{I}_z$ . Il est important ici de ne pas confondre ce terme avec un vecteur. Il s'agit d'un opérateur qui génère les  $2I+1$  niveaux d'énergie décrivant l'effet Zeeman. Notre objectif est maintenant d'exprimer  $\sigma_{ZZ}$  dans le repère des axes principaux, en utilisant les angles  $\alpha$  et  $\beta$  définis plus haut. Il a été montré qu'on peut arriver au résultat suivant :

$$\sigma_{ZZ} = \sigma_{iso} + \frac{1}{2} \Delta\sigma [(3 \cos^2(\beta) - 1) - \eta_\sigma \sin^2(\beta) \cos(2\alpha)] \quad (3.24)$$

De ce fait, en réutilisant les formules (2.2) et (2.3) on peut en déduire la fréquence de résonance à laquelle on observera un noyau j :

$$\nu_j = \nu_0 \left[ 1 - \left( \sigma_{iso} + \frac{1}{2} \Delta\sigma [(3 \cos^2(\beta) - 1) - \eta_\sigma \sin^2(\beta) \cos(2\alpha)] \right) \right] \quad (3.25)$$

Que nous dit cette équation ? La fréquence de résonance d'un noyau I correspond à la fréquence de Larmor, à laquelle on enlève un facteur proportionnel aux effets de déplacement chimique isotrope et anisotrope. Attardons-nous un peu sur le terme anisotrope, qui dépend des angles entre le repère d'axes propres et le repère du laboratoire (dont l'axe z est défini suivant  $\vec{B}_0$ ). Dans un liquide, nous étudions des noyaux se réorientant rapidement et de ce fait, pour chaque couple d'angles  $\{\alpha, \beta\}$ , nous allons aussi observer au cours du temps les contributions  $\{\pi + \alpha, \pi + \beta\}$ . Lors d'une expérience RMN, ces contributions vont donc s'annuler, ne laissant que la valeur isotrope de l'interaction de déplacement chimique. Dans cette thèse nous travaillons exclusivement avec des échantillons solides au sein desquels la réorientation rapide est limitée. De ce fait, les contributions anisotropes ne seront pas annulées. Nous verrons dans une des parties suivantes qu'il est possible de moyennner artificiellement les composantes de déplacement chimique anisotrope dans ce type d'échantillon.

En pratique, si on mesure en effet directement  $\nu_j$ , ce n'est pas la valeur que l'on utilise pour caractériser un échantillon. On détermine la différence entre la transformée de Fourier de cette fréquence,  $\nu_j$  et celle d'un échantillon de référence,  $\nu_{ref}$  et on divise le tout par la fréquence de Larmor du noyau observé au champ magnétique du spectromètre utilisé,  $\nu_{spec}$  :

$$\delta = \frac{\nu_j - \nu_{ref}}{\nu_{spec} \cdot 10^6} \quad (3.26)$$

Cette opération permet de rendre le signal observé indépendant de l'appareil d'étude et donc de le normaliser. Comme cette fréquence est exprimée en MHz,  $\delta$  est exprimé en parties par millions (ppm).

*b. Le couplage dipolaire :*

Passons maintenant au couplage dipolaire  $\hat{\mathcal{H}}_D$ . Dans notre échantillon, nous avons un ensemble de noyaux possédant chacun leur moment magnétique  $\vec{\mu}_j$ . Ces moments magnétiques interagissent forcément les uns avec les autres, que ce soit au sein de la molécule ou entre deux molécules différentes. Pour deux noyaux j et k, le couplage dipolaire est défini de la manière suivante :

$$\hat{\mathcal{H}}_D = -d_{jk} [3(\vec{I}_j \cdot \vec{r}_{jk})(\vec{I}_k \cdot \vec{r}_{jk}) - \vec{I}_j \cdot \vec{I}_k] \quad (3.27)$$

Avec  $d_{jk}$ , la constante dipolaire entre les deux noyaux définis par :

$$d_{jk} = \frac{\mu_0 \gamma_j \gamma_k}{4\pi r_{jk}^3} \quad (3.28)$$

On représente usuellement le couplage dipolaire grâce au schéma présenté dans la Figure 26. Une des particularités du couplage dipolaire réside dans le fait qu'il se définit différemment selon la nature des noyaux j et k. On différencie le couplage hétéronucléaire du couplage homonucléaire. Le couplage homonucléaire entre deux noyaux du même élément, dans l'approximation des champs forts, prend la forme suivante :

$$\hat{\mathcal{H}}_{jk}^{Homo} = -d_{jk} \frac{1}{2} (3 \cos^2(\beta) - 1) [3(\widehat{I}_{zj} \cdot \widehat{I}_{zk}) - (\widehat{I}_{zj} \cdot \widehat{I}_{zk} + \widehat{I}_{xj} \cdot \widehat{I}_{xk} + \widehat{I}_{yj} \cdot \widehat{I}_{yk})] \quad (3.29)$$

$$= -d_{jk} \frac{1}{2} (3 \cos^2(\beta) - 1) [2(\widehat{I}_{zj} \cdot \widehat{I}_{zk}) - (\widehat{I}_{xj} \cdot \widehat{I}_{xk} + \widehat{I}_{yj} \cdot \widehat{I}_{yk})]$$

Afin d'arriver à cette équation, nous sommes partis de l'équation (3.24). Nous avons d'abord extrait la contribution du vecteur  $\vec{r}_{jk}$ , ce qui aboutit au terme « $(3 \cos^2(\beta) - 1)$ ». Dans l'approximation des champs forts le premier terme entre crochet peut être assimilé aux contributions des opérateurs selon l'axe z, de manière similaire à ce qui est décrit dans l'équation (3.23). Enfin, le produit scalaire entre les deux moments cinétiques de spin mène au terme  $(\widehat{I}_{xj} \cdot \widehat{I}_{xk} + \widehat{I}_{yj} \cdot \widehat{I}_{yk})$  qui dépend des coordonnées x et y. Ce dernier terme est non négligeable pour deux noyaux du même élément.

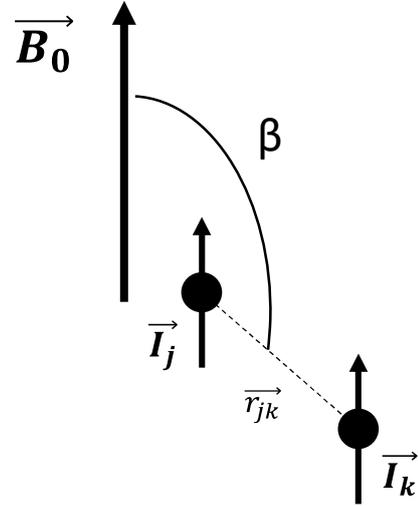


Figure 26: Représentation schématique du couplage dipolaire dans le repère du laboratoire.  $\beta$  définit l'angle entre le vecteur

Dans le cas de deux noyaux d'éléments différents, on obtient l'équation suivante :

$$\widehat{\mathcal{H}}_{jk}^{\text{Hétéro}} = -d_{jk} \frac{1}{2} (3 \cos^2(\beta) - 1) [2(\widehat{I}_{zj} \cdot \widehat{S}_{zk})] \quad (3.30)$$

Encore une fois nous voyons un terme prenant la forme suivante :  $(3 \cos^2(\beta) - 1)$ . De manière analogue à l'interaction de déplacement chimique, ce terme se moyenne à zéro en solution, et est observable en milieu solide.

### c. Les couplages scalaires et quadripolaires :

Les couplages scalaires et quadripolaires ne concernent pas vraiment ce travail. De ce fait, ils seront mentionnés et très légèrement décrits sans être approfondis.

Le couplage scalaire  $\widehat{\mathcal{H}}_j$  est similaire au couplage dipolaire, en ceci qu'il concerne le couplage entre deux spins nucléaires. Cependant, il est dû aux électrons formant la liaison entre les deux noyaux concernés. C'est pourquoi il est aussi appelé parfois le

couplage dipolaire indirect. En pratique, dans les échantillons solides il est difficile à différencier du couplage dipolaire et de ce fait ne sera pas plus décrit.

Enfin, l'interaction quadripolaire  $\hat{H}_Q$  est la seule interaction d'origine électrique dans les interactions internes. Cette interaction ne concerne que les noyaux de spin nucléaire supérieur à 1/2. Au sein de ce type de noyau, il existe une distribution des charges qui interagit avec le gradient de champ électrique généré par les électrons environnants. Etant donné que seuls des noyaux de spin 1/2 sont étudiés ici, nous n'approfondirons pas plus sa description.

Au final tous ces termes peuvent être classés selon leur ordre de grandeur, et leur existence dans les échantillons solides ou liquides (Tableau 5).

Tableau 5: Ordre de grandeur des interactions présentes dans les échantillons selon leur état.

Interaction	Ordre de grandeur (Hz)	Intensité typique en phase solide (Hz)	Intensité typique en phase liquide (Hz)
Effet Zeeman	$10^6$ - $10^9$	$10^8$	$10^8$
Déplacement chimique (DC)	$0$ - $10^3$	$10^2$	$10^2$
Couplage dipolaire	$0$ - $10^5$	$10^4$	$10^4$
Couplage scalaire	$0$ - $10^5$	$10^2$	0
Couplage quadripolaire	$0$ - $10^9$	$10^6$	0

Ces interactions ont des ordres de grandeurs différents, le déplacement chimique étant celui qui présente le moins élevé. Comme il s'agit de celle permettant la caractérisation des différentes structures présentes au sein des échantillons étudiés ici, il est nécessaire de s'abroger des autres effets.

## 4. Techniques expérimentales utilisées

Les échantillons étudiés au cours de cette thèse présentent deux contraintes majeures : il s'agit d'échantillons à l'état solide, et les noyaux présents sont en abondance naturelle. Pour répondre à ces contraintes nous allons maintenant discuter de deux techniques essentielles employées au cours de cette thèse : la rotation à l'angle magique et la polarisation croisée.

### I. La rotation à l'angle magique

Comme nous en avons discuté précédemment nous étudions des échantillons à l'état solide. De ce fait, la contribution anisotrope du déplacement chimique ainsi que le couplage dipolaire seront observables. Concernant le couplage dipolaire, nous reviendrons sur ce point dans la partie suivante de ce chapitre. Le problème engendré par la contribution anisotrope du déplacement chimique réside dans le fait qu'elle élargit énormément le signal obtenu, rendant toute interprétation impossible. Il existe un moyen technique permettant de moyenniser cette interaction à zéro. Rappelons que ces deux contributions dépendent d'un terme ayant la forme  $(3 \cos^2(\beta) - 1)$ . Il existe donc un angle  $\beta$  pour lequel cette contribution s'annule. Cet angle s'appelle l'angle magique et prend la valeur suivante :

$$\beta_{mag} = \cos^{-1}\left(\sqrt{1/3}\right) = \tan^{-1}(\sqrt{2}) \cong 54.74^\circ \quad (4.31)$$

Evidemment, tous les moments magnétiques dans l'échantillon considéré ne formeront pas cet angle avec le champ  $\vec{B}_0$ . En pratique, l'échantillon est inséré sous

forme de poudre dans un rotor qui est lui-même orienté de manière à former cet angle avec  $\vec{B}_0$ . Si le rotor est mis en rotation autour de son axe, chaque moment magnétique devrait former l'équivalent d'un cône autour de cet angle (Figure 27). Le moment magnétique considéré s'en retrouve moyenné à un vecteur aligné avec l'axe de rotation. Les effets sur le spectre dépendront de la vitesse de rotation (Figure 28).

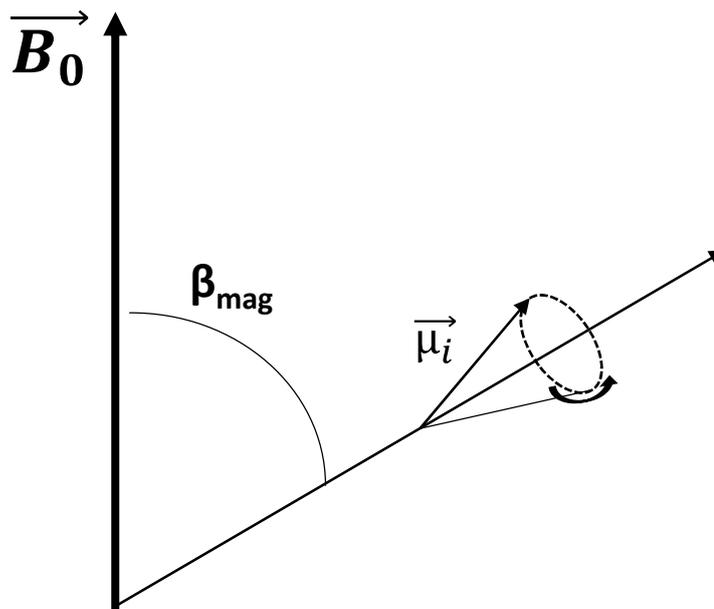


Figure 27: Représentation schématique de l'effet de la rotation à l'angle magique sur  $\vec{\mu}_i$

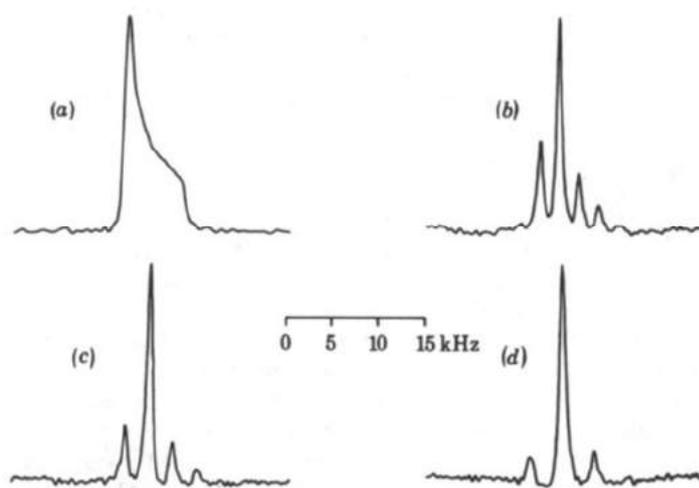


Figure 28: Effet de la rotation à l'angle magique pour des fréquences de 2.1, 2.6 et 3.6 kHz sur le spectre d'un échantillon de  $^{111}\text{Cd}$ . Extrait de Andrew et al., 1974.<sup>29</sup>

Le pic dont la fréquence est indépendante de la vitesse de rotation correspond à la contribution isotrope du déplacement chimique et les autres pics correspondent à ce qu'on appelle des bandes de rotation. Ces bandes apparaissent à des fréquences multiples de la vitesse de rotation du rotor de part et d'autre du pic isotrope. Leur intensité dépend l'amplitude de l'interaction anisotrope considérée et de la vitesse de

rotation du rotor. Quand la vitesse de rotation du rotor est supérieure à l'anisotrope de déplacement chimique (exprimée en Hz), l'intensité les bandes de rotation est nulle et seul le pic isotrope persiste avec une intensité plus élevée.

## II. La polarisation croisée

Comme nous l'avons abordé plus haut, l'anisotropie de déplacement chimique n'est pas la seule limitation rencontrée dans la caractérisation de ces échantillons. Les différentes structures de nos échantillons sont caractérisées par des différences de déplacement chimique des atomes de carbone. Les polysaccharides étudiés (issus de la pomme de terre ou des mycobactéries) ne pouvant pas être marqués au carbone 13 (isotope du carbone de spin  $\frac{1}{2}$ ) à un coût raisonnable, les échantillons doivent être caractérisés en abondance naturelle. De ce fait, seulement 1% (ratio naturel entre les isotopes  $^{13}\text{C}$  et  $^{12}\text{C}$ ) de chaque carbone de l'échantillon entrera en résonance lors de l'expérience. A ceci s'ajoute le couplage dipolaire hétéronucléaire proton-carbone, qui aura comme effet de diviser le signal en plusieurs signaux d'intensité réduite. Il nous faut donc un moyen d'augmenter l'intensité du signal d'une manière ou d'une autre afin de pouvoir étudier ces échantillons.

Concernant le couplage dipolaire hétéronucléaire, il peut être éliminé de manière relativement simple en appliquant un champ radiofréquence à la fréquence de Larmor des atomes d'hydrogène au cours de l'acquisition du FID à la fréquence de Larmor du carbone.

Il est possible dans certaines conditions de transférer l'aimantation d'un noyau vers un autre. L'objectif ici est d'utiliser un noyau de spin I beaucoup plus abondant et ayant un rapport gyromagnétique plus élevé (typiquement  $^1\text{H}$ ) afin d'augmenter le signal du noyau de spin S plus difficile à observer ( $^{13}\text{C}$ ). Pour ceci il faut que l'énergie de résonance des deux noyaux soit égale (Figure 29).

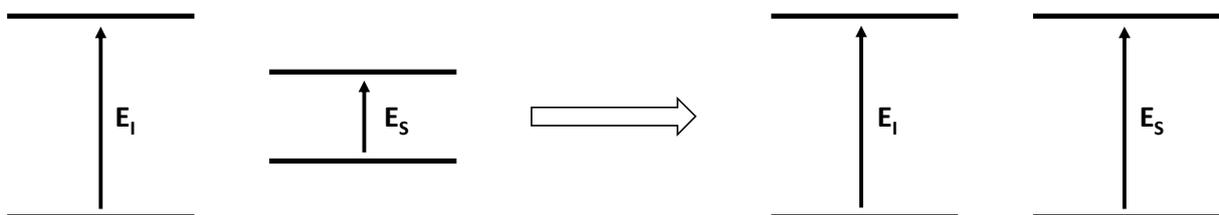


Figure 29: Schéma représentant la condition de Hartmann-Hahn.

Pour obtenir de telles conditions, il est nécessaire d'appliquer deux champs radiofréquence de manière à obtenir l'égalité suivante :

$$\gamma_I B_I = \gamma_S B_S + n\omega_{mas} \quad (4.32)$$

Où  $\gamma_I$  et  $\gamma_S$  sont respectivement les rapports gyromagnétiques du spin I et du spin S,  $B_I$  et  $B_S$  respectivement les valeurs des champs rf appliqués à la fréquence du spin I et du spin S,  $\omega_{mas}$  la fréquence de rotation et n un entier. Cette égalité est appelée la condition de Hartmann-Hahn<sup>16</sup>. Elle explique aussi la présence des bandes de résonance décrites dans la partie précédente. Le rapport gyromagnétique du proton étant environ quatre fois plus élevé que celui du carbone, le rapport dit « signal/bruit » sera multiplié par 4 grâce au transfert de polarisation.

En mettant en place les deux solutions « techniques » proposées, nous avons donc finalement la possibilité d'étudier des échantillons carbonés à l'état solide, et en abondance naturelle. L'expérience de RMN qui sera la plus utilisée dans cette thèse est donc l'expérience de polarisation croisée en rotation à l'angle magique ou CP-MAS, et nous permettra d'étudier les signaux RMN de noyaux  $^{13}\text{C}$  des échantillons étudiés (Figure 30).

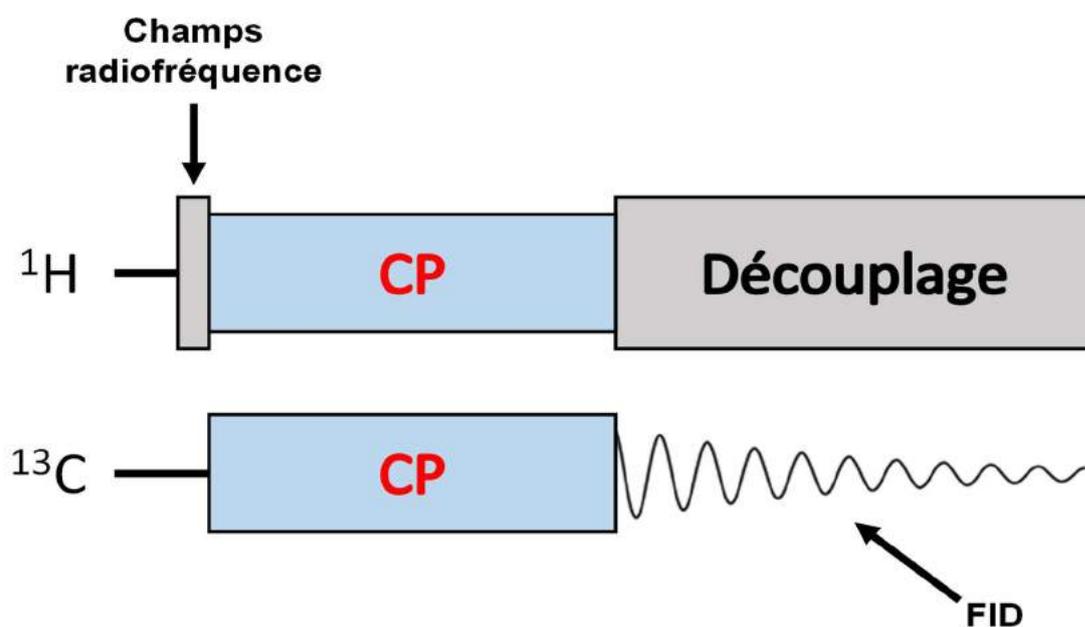


Figure 30: Séquence d'impulsion pour l'expérience de CP-MAS.

## Chapitre 2. Outils théoriques utilisés

Afin de pouvoir établir une relation entre la structure des molécules étudiées et les spectres RMN obtenus, nous avons réalisés des calculs de tenseurs d'écrantage chimique en utilisant différentes méthodes. Les quelques paragraphes qui suivent visent à la description de ces dernières et feront notamment référence à des notions décrites dans l'annexe de mécanique quantique.

### 1. Calculs d'écrantage chimique

Nous allons tout d'abord discuter du formalisme mathématique utilisé pour les calculs d'écrantage chimique. Cela nous permettra d'aborder un problème fondamental des méthodes théoriques : l'origine de jauge. Puis nous décrirons les méthodes utilisées pour pallier ce problème, d'abord sur molécule unique, puis en conditions périodiques.

#### I. Formalisme mathématique

Nous avons vu précédemment que l'application d'un champ magnétique extérieur  $\vec{B}_0$  induit un courant et donc une densité de courant  $\vec{j}(\vec{r})$  au sein du nuage électronique du système considéré. En utilisant la loi de Biot et Savart, on sait que le champ magnétique induit à la position  $\vec{r}_j$  par cette densité de courant est de la forme :

$$\vec{B}_{induit} = -\frac{\mu_0}{4\pi} \int \frac{\vec{j}(\vec{r}) \times (\vec{r} - \vec{r}_j)}{|\vec{r} - \vec{r}_j|^3} d\vec{r} \quad (1.33)$$

Comme nous l'avons vu précédemment, la densité de courant générée va donc modifier l'hamiltonien électronique du système et la relation entre la densité de courant et le tenseur d'écrantage chimique peut être déduit des équations (1.33) et (3.15). Dans l'annexe de mécanique quantique est défini l'hamiltonien électronique du

système sans interaction :  $\hat{\mathcal{H}}_{elec}$ . En présence d'un champ magnétique externe, le nouvel hamiltonien est le suivant :

$$\hat{\mathcal{H}}_{magnétique} = \hat{\mathcal{H}}_{elec} + \vec{A} \cdot \vec{p} + \frac{1}{2} \vec{A}^2 \quad (1.34)$$

Nous introduisons ici la notion de potentiel vecteur  $\vec{A}$ . Il s'agit d'un vecteur défini de la manière suivante :

$$\vec{B}_0 = \vec{\nabla} \times \vec{A} \quad (1.35)$$

La forme de ce dernier peut être choisie arbitrairement, aussi on le trouve sous la forme suivante :

$$\vec{A} = \frac{1}{2} \vec{B}_0 (\vec{r} - \vec{R}_0) \quad (1.36)$$

Ici on voit que le potentiel vecteur est défini en fonction de  $\vec{R}_0$  qui représente ce qu'on appelle *l'origine de jauge*. Finalement, en insérant l'équation (1.36) dans l'équation (1.34), l'hamiltonien final du système dans un champ magnétique s'exprime donc :

$$\hat{\mathcal{H}}_{mag} = \hat{\mathcal{H}}_{elec} + \frac{1}{2} [(\vec{r} - \vec{R}_0) \times \vec{p}] \vec{B}_0 + \frac{1}{8} [\vec{B}_0 \times (\vec{r} - \vec{R}_0)]^2 \quad (1.37)$$

Il est expliqué dans l'annexe de mécanique quantique comment obtenir les spin-orbitales  $\varphi_i^{(0)}$  du système en absence de champs magnétique, à partir de  $\hat{\mathcal{H}}_{elec}$  en utilisant les méthodes issues de la Théorie de la Fonctionnelle de la Densité (DFT). En présence d'un champ magnétique, les orbitales vont être impactées par ce dernier, et on peut les redéfinir, ainsi que la densité de courant en utilisant la théorie de la perturbation :

$$\varphi_i = \varphi_i^{(0)} + i\varphi_i^{(1)} + O(B_0^2) \quad (1.38)$$

$$\vec{j} = \vec{j}^{(1)} + O(B_0^2) \quad (1.39)$$

Ici,  $\varphi_i^{(0)}$  représente l'orbitale  $i$  en absence de champ magnétique et  $\varphi_i^{(1)}$  sa réponse en présence d'un champ magnétique. Concernant la densité de courant, elle n'existe que si l'on est en présence d'un champ magnétique. A partir de ces orbitales  $\varphi_i$ , on peut calculer la densité de courant de la manière suivante :

$$\vec{j} = i \sum_{i=1}^n (\varphi_i^* \vec{\nabla} \varphi_i - (\vec{\nabla} \varphi_i^*) \varphi_i) - 2\vec{A} \sum_{i=1}^n \varphi_i^* \varphi_i \quad (1.40)$$

Si les équations de Kohn et Sham sont résolues exactement,  $\vec{j}$  ne devrait pas dépendre de l'origine de jauge. Or dans les faits, on ne résout pas exactement ces équations, car nous n'utilisons pas de base complète (ou infinie) pour décrire nos orbitales. Ce problème est décrit plus en profondeur dans l'annexe de mécanique quantique. En incluant la nouvelle définition des orbitales perturbées dans l'équation (1.40), on aboutit à l'équation suivante concernant

$$\vec{j} = 2 \sum_{i=1}^n (\varphi_i^{(1)} \vec{\nabla} \varphi_i^{(0)} - \varphi_i^{(0)} \vec{\nabla} \varphi_i^{(1)}) - 2\vec{A} \sum_{i=1}^n \varphi_i^{(0)} \varphi_i^{(0)} \quad (1.41)$$

La première somme concerne ce qu'on appelle la partie paramagnétique de la densité de courant, et la seconde la partie diamagnétique. Comme nous pouvons le voir, seule la partie paramagnétique contient les orbitales perturbées. Il se trouve que la réponse des orbitales au champ magnétique peut-être exprimée en fonction des orbitales natives du noyau considéré :

$$\varphi_i^{(1)} = \frac{1}{2}(\vec{B}_0 \times \vec{R}_0)\vec{r}\varphi_i^{(0)} \quad (1.42)$$

A partir d'ici, l'annexe de mécanique quantique permet d'appréhender de manière plus claire les problématiques sous-jacentes au problème d'origine de jauge. Dans ce dernier, nous avons vu que les orbitales  $\varphi_i^{(0)}$  sont définies en sommes de fonctions de bases, et que ces sommes devraient être infinies pour en permettre la définition exacte. En pratique, il est impossible d'utiliser un nombre de fonctions de bases infinies, nous utilisons donc des fonctions gaussiennes contractées (CGF) afin de mimer le comportement de fonctions de type « Slater orbitals ». C'est de là que vient le principal problème dû à l'origine de jauge : la réponse au champ magnétique des orbitales du système nécessite un nombre de CGF beaucoup trop important pour que l'on puisse réaliser ce genre de calculs sur des systèmes tels que ceux étudiés ici.

Pour illustrer cette affirmation, prenons un exemple au travers du Tableau 6. Dans celui-ci sont représentées les constantes isotropes d'écrantage chimique du proton dans une molécule de méthane en fonction de l'origine de jauge utilisée. Par soucis de comparaison, il faut savoir que l'on rencontre usuellement l'utilisation de base dites « double-zêta », dans lesquelles on utilise deux fonctions pour modéliser le comportement des orbitales. La base minimale utilisée dans le Tableau 6 est une base « triple-zêta », c'est-à-dire utilisant trois fonctions pour modéliser les orbitales, étant de ce fait plus précise que les fonctions double-zêta. Même avec l'utilisation de ce type de bases, les résultats sont extrêmement divers concernant le calcul de l'écrantage chimique du proton.

L'ajout de fonctions de polarisations (3P), ainsi que de fonctions non-contactées sur l'atome de carbone (dec) permet de s'approcher de la convergence. Cependant ce n'est qu'en utilisant une base étendue (extended) contenant un très grand nombre de fonctions (carbone : 61 fonctions, hydrogène : 24 fonctions) que le résultat devient similaire à celui obtenu par l'utilisation de méthode indépendantes de jauge (IGLO, GIAO).

Tableau 6 : Constante isotrope d'écrantage chimique du proton pour la molécule de méthane. Extrait de Kaup et al. 2004<sup>17</sup>

Basis Set	Common Gauge Origin at			IGLO	GIAO
	C atom	same H	other H		
TZ	28.58	62.01	17.43	32.48	32.57
TZ +3P	30.97	41.22	27.56	31.29	31.35
TZ-dec +3P	30.98	32.07	30.62	31.20	31.34
Extended	31.23	31.26	31.22	31.24	31.25

Le problème d'origine de jauge est donc un problème numérique. Le méthane est une petite molécule, mais si l'on désire réaliser des calculs sur des molécules contenant un nombre d'atomes plus important, il faut trouver un moyen de calculer ces écrantages chimiques sans avoir à utiliser une quantité de fonctions de base aussi grande.

## II. Invariance de Jauge

Pour ce faire, on peut profiter de la décomposition de la densité présentée dans l'équation (1.41). En effet, dans le cas où l'origine de jauge est située au même endroit que l'atome considéré, la partie paramagnétique de la densité de courant s'annule pour des raisons de symétrie :

$$\vec{j}(\vec{r}) = -2\vec{A} \sum_{i=1}^n \varphi_i^{(0)} \varphi_i^{(0)} = -(\vec{B}_0 \times \vec{r}) \sum_{i=1}^n \varphi_i^{(0)} \varphi_i^{(0)} \quad (1.43)$$

Ce qui nous amène à la définition suivante concernant le champ induit :

$$\vec{B}_{induit} = -\frac{\mu_0}{4\pi} \sum_{i=1}^n \langle \varphi_i^{(0)} | \frac{(\vec{B}_0 \times \vec{r}) \times \vec{r}}{r^3} | \varphi_i^{(0)} \rangle \quad (1.44)$$

$$= -\vec{B}_0 \frac{2}{3} \frac{\mu_0}{4\pi} \sum_{i=1}^n \langle \varphi_i^{(0)} \left| \frac{1}{r} \right| \varphi_i^{(0)} \rangle$$

En se rappelant que  $\vec{B}_{induit} = -\boldsymbol{\sigma} \cdot \vec{B}_0$  nous voyons apparaitre une nouvelle façon de définir le tenseur d'écrantage chimique :  $\boldsymbol{\sigma} = \frac{2}{3} \frac{\mu_0}{4\pi} \sum_{i=1}^n \langle \varphi_i^{(0)} \left| \frac{1}{r} \right| \varphi_i^{(0)} \rangle$ . Une façon d'éviter le recours à une base de fonctions trop grande serait donc d'utiliser non pas une, mais plusieurs jauges, distribuées de façon à approximer au mieux le calcul de la partie paramagnétique du courant.

### a. La méthode GIAO

Soit un atome situé à une position A et l'origine de jauge étant 0. On considère que les densités électroniques de tous les atomes du système considéré ne se superposent pas. En définissant les orbitales de cet atome comme étant localisées sur A, nous pouvons définir la densité de courant comme ceci :

$$\begin{aligned} \vec{j}^{(1)} &\approx -(\vec{B}_0 \times (\vec{r} - \vec{R}_A)) \sum_{i=1}^n \varphi_i^{(0)} \varphi_i^{(0)} \\ &= -\sum_1^n (\vec{B}_0 \times (\vec{r} - \vec{R}_i)) \varphi_i^{(0)} \varphi_i^{(0)} \quad (1.45) \\ &= \sum_{i=1}^n (\vec{B}_0 \times \vec{R}_i) \varphi_i^{(0)} \varphi_i^{(0)} - \sum_{i=1}^n (\vec{B}_0 \times \vec{r}) \varphi_i^{(0)} \varphi_i^{(0)} \end{aligned}$$

Nous voyons ici que la deuxième somme est la définition exacte de la partie diamagnétique du courant induit. De ce fait, la première somme est une approximation de la partie paramagnétique du courant, exprimée seulement en fonction des orbitales non-perturbées. Formulé autrement : une partie de la contribution paramagnétique de la densité de courant peut être exprimée comme une contribution diamagnétique à laquelle nous avons attribué une origine de jauge propre :  $\vec{R}_A = \vec{R}_i$ . Les orbitales

perturbées par le champ magnétique se définissent donc, par analogie avec l'équation (1.42) :

$$\varphi_i^{(1)}(\vec{r}) = \frac{1}{2}(\vec{B}_0 \times \vec{R}_i)\vec{r} \varphi_i^{(0)} + \widetilde{\varphi}_i^{(1)}(\vec{r}) \quad (1.46)$$

Ceci reste une approximation, car notre jeu de fonctions de bases ne nous permet pas de définir de manière exacte les orbitales du système. De ce fait, il reste une partie inconnue  $\widetilde{\varphi}_i^{(1)}(\vec{r})$  dans la définition des orbitales, due à l'approximation d'origine sur le recouvrement des densités électroniques qui est omis. Heureusement, en pratique cette partie inconnue est une contribution assez faible pour pouvoir être négligée et amener des résultats corrects. En effet, le terme  $\frac{1}{2}(\vec{B}_0 \times \vec{R}_i)\vec{r}$ , qui est appelé le « terme de jauge », relocalise l'orbitale  $\varphi_i^{(1)}(\vec{r})$  sur l'atome d'où elle provient, et de ce fait contient la plus grande part des fonctions qui la définisse.

Ce que nous avons décrit jusqu'ici définit les fondements de la méthode dite Individual Gauge for Localized Orbital (IGLO)<sup>18</sup>.

La méthode GIAO<sup>19,20</sup> se base sur ce principe aussi, cependant ce sont les fonctions de base elles-mêmes qui subissent une transformation de jauge. Comme nous l'avons déjà précisé, notamment dans l'annexe de mécanique quantique, les orbitales sont définies par rapport à un jeu de fonctions de bases :

$$\varphi_i^{(0)} = \sum_{\mu=1}^{N_{base}} c_{\mu i} \eta_{\mu} \quad (1.47)$$

On définit de ce fait les orbitales de London, par analogie avec l'équation (1.46) :

$$\varphi_i^{(1)}(\vec{r}) = - \sum_{\mu=1}^{N_{base}} c_{\mu i} \frac{1}{2}(\vec{B}_0 \times \vec{R}_{\mu})\vec{r} \eta_{\mu} + \widetilde{\varphi}_i^{(1)}(\vec{r}) \quad (1.48)$$

Et la densité de courant générée peut donc être calculée de la façon suivante :

$$\begin{aligned}
\vec{j}^{(1)} = & 2 \sum_{i=1}^N \widetilde{\varphi}_i^{(1)} \vec{\nabla} \varphi_i^{(0)} - \varphi_i^{(0)} \vec{\nabla} \widetilde{\varphi}_i^{(1)} \\
& + \sum_{i=1}^N \sum_{\mu, \nu=1}^{N_{base}} c_{\mu i} c_{\nu i} \left( \vec{B}_0 \times \left[ \vec{r} - \frac{1}{2} (\vec{R}_\mu + \vec{R}_\nu) \right] \right) \eta_\mu \eta_\nu \\
& + \sum_{i=1}^N \sum_{\mu, \nu=1}^{N_{base}} c_{\mu k} c_{\nu k} \frac{1}{2} \vec{B}_0 \times (\vec{R}_\nu - \vec{R}_\mu) \vec{r} (\eta_\mu \vec{\nabla} \eta_\nu - \eta_\nu \vec{\nabla} \eta_\mu)
\end{aligned} \tag{1.49}$$

Par analogie avec l'équation 1.41, nous retrouvons donc la partie paramagnétique restante (première somme) et la partie diamagnétique totale (deuxième somme) définies en fonction des nouvelles fonctions de bases corrigées par le changement de jauge. La troisième somme présente dans cette équation est nécessaire ici, car les fonctions de bases modifiées par le changement de jauge ne sont pas orthogonales. Il en résulte des termes de recouvrement qui doivent être compensés.

Ce dernier terme a longtemps été difficile à calculer, rendant de ce fait la méthode IGLO plus populaire. A partir du moment où il a été possible de calculer ce genre d'intégrales plus facilement, la méthode GIAO est devenue plus utilisée car elle permettait de s'affranchir des orbitales localisées. Comme montré dans le Tableau 2, les deux méthodes permettent d'obtenir de très bons résultats et une convergence rapide malgré un set de base limité.

### b. La méthode GIPAW

La méthode dite Gauge Including Projector Augmented Wave (GIPAW)<sup>21</sup> est originaire de la physique du solide et a été utilisée dans peu de calculs concernant cette thèse. Elle sera donc décrite dans les grandes lignes, sans entrer en profondeur dans les équations. Plusieurs précisions sont tout d'abord nécessaires : i) comme indiqué dans

son nom, il s'agit d'une méthode utilisant comme base des ondes planes ; ii) cette méthode est réellement utile si nous étudions un système comprenant des conditions de symétrie périodiques, tel qu'un cristal.

Pour comprendre la théorie sous-jacente derrière la méthode GIPAW, il va tout d'abord falloir discuter des méthodes Pseudopotential Planewave (PP) et Linearized Augmented Plane-Wave (LAPW).

Comme nous l'avons précisé plus haut, nous utilisons des ondes planes comme bases pour décrire la densité électronique du système. Ce type de fonction présente l'avantage de ne pas dépendre des positions atomiques et d'être relativement simple. En contrepartie, ces bases présentent des oscillations extrêmement fortes au voisinage du noyau considéré et vont nécessiter un très grand nombre de fonctions pour reproduire de manière précise le comportement oscillant de la fonction d'onde proche du noyau (Figure 31). L'idée derrière la méthode PP, est de remplacer le potentiel proche du noyau par un pseudo-potential qui permettra de reproduire le comportement des électrons de valence. Cela permet de réduire grandement la durée du calcul<sup>22</sup>.

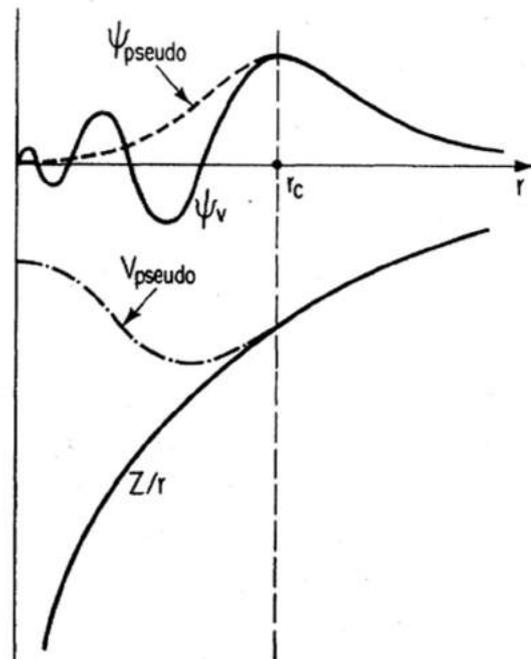


Figure 31: Représentation schématique du comportement de la fonction d'onde et du potentiel en considérant tous les électrons (trait plein) ou en utilisant un pseudopotential (pointillés). Extrait de Payne (1992).

La méthode LAPW est basée sur une approche légèrement différente. Au lieu d'utiliser un pseudo potentiel afin d'approximer le comportement de la fonction d'onde aux alentours du noyau, on sépare le système en deux parties<sup>23</sup> : d'un côté on définit une sphère autour de chaque noyaux, et de l'autre on considère le reste du système comme une région unique (Figure 32). La région interstitielle est définie en utilisant des ondes planes classiques comme fonctions de base, et les régions sphériques autour de chaque atome sont définies grâce à des fonctions dites « atomic-like » similaires à celles définies dans l'annexe de mécanique quantique.

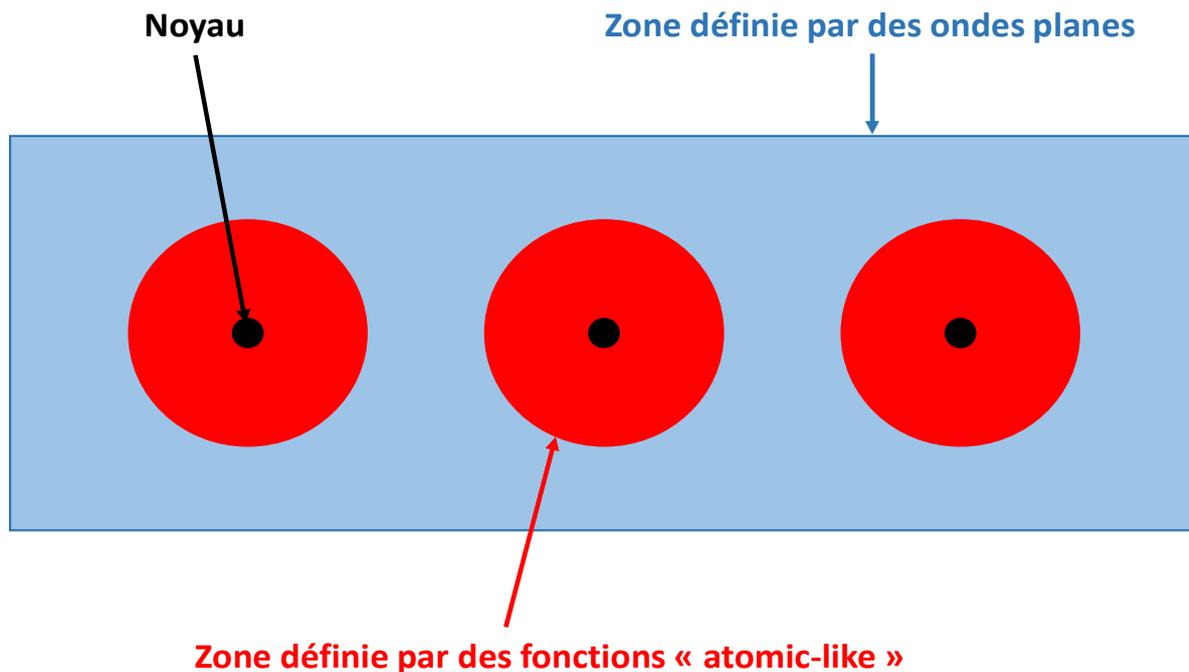


Figure 32: Description schématique de la méthode LAPW. Adapté de John Pask(14).

Cette méthode permet de retrouver la définition précise de la fonction d'onde proche du noyau, mais en contrepartie est plus coûteuse en ressources calculatoires.

La méthode que nous utilisons est appelée Projector Augmented Wave (PAW) et permet d'unifier les deux approches décrites ci-dessus en utilisant un pseudo potentiel au sein des parties sphériques, et en utilisant des ondes planes dans les parties interstitielles. En utilisant cette approche, il est possible de calculer, à partir de l'utilisation de pseudo-potentiels, des propriétés qui normalement nécessitent la description des tous les électrons du système. L'approche PAW reste cependant dépendante de l'origine de jauge utilisée dans le cas de calculs de paramètres RMN. L'approche GIPAW, qui est finalement analogue à l'approche GIAO, inclue un terme de correction de jauge.

L'intérêt principal de l'utilisation de cette méthode, réside dans le fait qu'il est possible d'inclure le caractère périodique du système étudié dans les ondes planes utilisées. En effet, les ondes planes qui sont utilisées dans la méthode GIPAW sont des ondes de Bloch, définies de la manière suivante :

$$\psi(r) = e^{ikr} u(r) \tag{1.50}$$

Ici,  $k$  représente le vecteur d'onde,  $r$  est la position dans l'espace et  $u(r)$  est une fonction périodique ayant la même périodicité que celle de l'échantillon étudié. Dans le cadre de ce travail de thèse, la méthode GIPAW sera donc utilisée pour l'étude d'échantillon d'amylose présentant des conditions périodiques, telles qu'un cristal de polymorphe B.

En conclusion, nous disposons donc de deux méthodes pour calculer des écrantages chimiques : l'approche GIAO, qui sera la méthode principalement utilisée dans cette thèse, et l'approche GIPAW, qui nous sera utile afin d'étudier certains échantillons connus pour leur haute cristallinité.

## 2. Génération des structures moléculaires utilisées

Maintenant que nous savons comment obtenir les paramètres RMN de nos systèmes, il nous faut des structures sur lesquelles effectuer ces calculs. Les parties à venir visent à décrire les méthodes quantiques et classiques qui seront utilisées dans les chapitres suivants.

### I. Optimisations de Géométrie

Il est difficile de définir une méthode d'optimisation de géométrie sans parler en premier lieu de surface d'énergie potentielle. Une molécule composée d'un nombre  $N$  d'atomes pourra prendre un grand nombre de conformations, dirigées par ses  $3N$  degrés de liberté. On peut associer une énergie à chaque conformation, et de ce fait construire ce qu'on appelle une hypersurface d'énergie potentielle de  $3N$  dimensions. La Figure 33<sup>24</sup> montre un exemple de projection de cette surface selon deux dimensions.

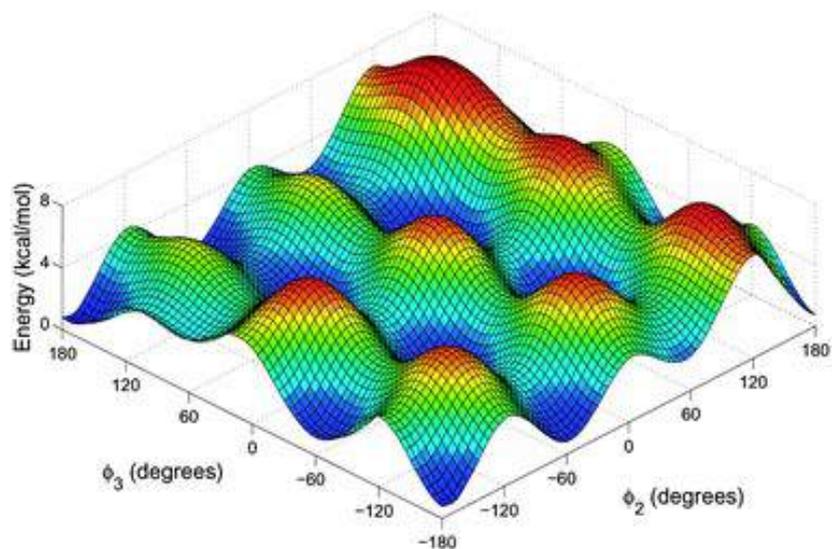


Figure 33: SEP de la tryptamine en fonction des angles  $\phi_2$  et  $\phi_3$ . Extraite de Y. Sturdy et al, 2007

L'objectif d'une optimisation de géométrie est de trouver la conformation qui amènera l'énergie la plus basse appelée minimum global de la surface d'énergie potentielle (SEP) qui sera représentative de la structure la plus stable du système (Figure 34). En pratique, sur les systèmes étudiés dans ce manuscrit, il est impossible de savoir si cette conformation est atteinte, car le nombre de degrés de liberté est trop important pour pouvoir être échantillonné entièrement. La plupart des algorithmes d'optimisation de géométrie, mènent au minimum d'énergie le plus proche de la structure de départ utilisée. Si ce dernier n'est pas le minimum global de la SEP, on parle donc de minimum local. Il est possible que les structures obtenues dans les minimums locaux ne correspondent pas à celles que l'on trouve expérimentalement. Les calculs réalisés de façon subséquente sur ces structures seront difficilement corrélables avec les observables expérimentales. En revanche, si l'on dispose d'une structure que l'on sait proche de la structure expérimentale, l'optimisation de géométrie est toute désignée afin de trouver la structure qui permettra d'étudier les observables désirées.

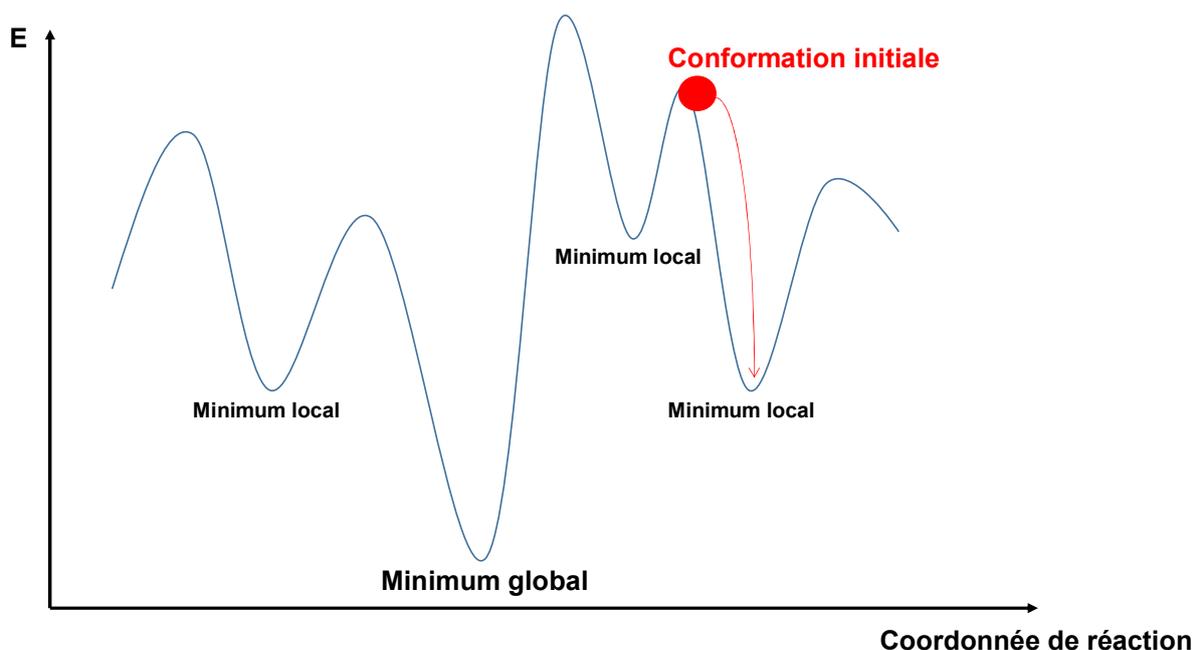


Figure 34: Représentation schématique d'une optimisation de géométrie.

## II. Dynamique moléculaire classique

Outre les problèmes de minima locaux, il existe cependant des cas dans lesquels l'étude d'une structure unique n'est pas suffisante, ou n'est tout simplement pas possible. On peut notamment mentionner les cas où des structures cristallographiques ne sont pas disponibles, ou encore les cas nécessitant la prise en compte de la dynamique du système mais aussi la prise en compte explicite de molécules de solvant afin de calculer les paramètres désirés.

La dynamique moléculaire est « une méthode de calcul qui génère une série de points corrélés dans l'espace des phases en propageant un jeu initial de coordonnées et vitesses en suivant les lois de Newton par l'utilisation de pas de temps finis »<sup>25</sup>. De manière schématique, cette méthode de calcul permet de déplacer le système sur la surface d'énergie potentielle au cours du temps. Cela nous permet d'étudier les propriétés dynamiques du système et de les inclure dans le calcul de nos observables expérimentales.

### a. Propagateurs et ensembles statistiques

La description complète du formalisme utilisé en dynamique moléculaire mériterait un chapitre complet. Dans ce travail, il s'agit simplement d'un outil permettant de générer des structures qui seront utilisées pour des calculs de RMN. De ce fait, nous nous intéresserons seulement à la définition de l'hamiltonien et la façon dont l'évolution temporelle des positions atomiques est calculée.

Pour commencer nous devons préciser certaines approximations réalisées. Contrairement aux méthodes de mécanique quantique, lors d'un calcul de dynamique moléculaire classique, les électrons ne sont pas pris en compte explicitement. De plus, les  $M$  atomes de masse  $m_I$  sont considérés comme des charges ponctuelles et leur masse est considérée suffisamment grande pour obéir aux équations de Newton. L'hamiltonien du système se définit de la façon suivante :

$$\hat{\mathcal{H}}_{MD}(\{\vec{R}_I\}, \{\vec{P}_I\}) = \sum_{I=1}^M \frac{\vec{P}_I^2}{2m_I} + V(\{\vec{R}_I\}) \quad (2.51)$$

Ce dernier comprend les termes d'énergie cinétique et potentielle du système qui, dans ce cas, dépendent respectivement de l'impulsion de chaque noyau  $\{\vec{P}_I\}$  et de leur position  $\{\vec{R}_I\}$ . Le comportement de chaque noyau peut donc être simulé en suivant la seconde loi de Newton, ce qui amène l'équation suivante :

$$m \frac{d^2 \vec{R}}{dt^2} = - \frac{dV}{d\vec{R}} \quad (2.52)$$

Le vecteur  $\vec{R}$  contient toutes les coordonnées de toutes les particules du système et  $V$  correspond à l'énergie potentielle à la position  $\vec{R}$ . On définit ainsi la vitesse et l'accélération au temps  $\Delta t$  d'un noyau localisé à la position  $\vec{R}_I$  de la façon suivante :

$$\vec{v}_I(\Delta t) = \frac{\partial \vec{R}}{\partial t}(\Delta t) = \frac{\vec{P}_I}{M_I} \quad (2.53)$$

$$\vec{a}_I(\Delta t) = \frac{\partial^2 \vec{R}}{\partial t^2}(\Delta t) = -\frac{1}{m_I} \frac{\partial V}{\partial \vec{R}_I} = \frac{\vec{F}_I(\{\vec{R}_I\})}{m_I} \quad (2.54)$$

L'utilisation d'une série de Taylor nous permet de définir la nouvelle position  $\vec{R}_{I+1}$  atteinte au temps  $\Delta t$  :

$$\begin{aligned} \vec{R}_{I+1} &= \vec{R}_I + \frac{\partial \vec{R}}{\partial t}(\Delta t) + \frac{1}{2} \frac{\partial^2 \vec{R}}{\partial t^2}(\Delta t)^2 + \frac{1}{6} \frac{\partial^3 \vec{R}}{\partial t^3}(\Delta t)^3 + \dots \\ \vec{R}_{I+1} &= \vec{R}_I + \vec{v}_I(\Delta t) + \frac{1}{2} \vec{a}_I(\Delta t)^2 + \frac{1}{6} \frac{\partial^3 \vec{R}}{\partial t^3}(\Delta t)^3 + \dots \end{aligned} \quad (2.55)$$

De la même manière, on peut définir la position  $\vec{R}_{I-1}$  atteinte à  $-\Delta t$  :

$$\vec{R}_{I-1} = \vec{R}_I - \vec{v}_I(\Delta t) + \frac{1}{2} \vec{a}_I(\Delta t)^2 - \frac{1}{6} \frac{\partial^3 \vec{R}}{\partial t^3}(\Delta t)^3 + \dots \quad (2.56)$$

Il est possible de s'affranchir du calcul de certaines des dérivées en additionnant les équations 2.55 et 2.56 :

$$\vec{R}_{I+1} = (2\vec{R}_I - \vec{R}_{I-1}) + \vec{a}_I(\Delta t)^2 + \dots \quad (2.57)$$

Cela revient donc à calculer la position  $\vec{R}_{I+1}$  au temps  $\Delta t$  en fonction des positions aux temps actuel et précédent et en fonction de l'accélération au pas de temps actuel.

Cette accélération est accessible depuis le potentiel au grâce à l'utilisation de l'équation 2.54.

L'équation 2.57 est appelée l'algorithme de Verlet et comme on tronque la série de Taylor après le terme d'ordre deux, elle présente l'avantage d'être correcte jusqu'au troisième ordre en  $\Delta t$ , les dérivées à l'ordre 3 de la position étant annulées. Cependant, la vitesse n'apparaît pas de façon explicite dans cette équation. Or c'est une donnée nécessaire pour pouvoir définir les ensembles statistiques qui présentent une température constante, tels que l'ensemble canonique (NVT) et l'ensemble Isotherme-isobare (NPT).

L'algorithme appelé « leap-frog » utilise un raisonnement proche de l'algorithme de Verlet, mais en utilisant un pas de temps de  $\frac{1}{2}\Delta t$ , ce qui donne les équations suivantes :

$$\vec{R}_{I+1} = \vec{R}_I + \vec{v}_{I+\frac{1}{2}}\Delta t \quad (2.58)$$

$$\vec{v}_{I+\frac{1}{2}} = \vec{v}_{I-\frac{1}{2}} + \vec{a}_I\Delta t \quad (2.59)$$

Cet algorithme ne permet cependant pas d'obtenir la vitesse et la position au même temps. En effet, si on connaît la position au temps  $t$ , on ne connaîtra la vitesse qu'aux temps  $t + \frac{1}{2}\Delta t$  et  $t - \frac{1}{2}\Delta t$ . Afin de résoudre ce problème, on utilise en pratique l'algorithme appelé « Velocity-Verlet » qui aboutit aux équations suivantes :

$$\vec{R}_{I+1} = \vec{R}_I + \vec{v}_I\Delta t + \frac{1}{2} \vec{a}_I\Delta t^2 \quad (2.60)$$

$$\vec{v}_{I+1} = \vec{v}_I + \frac{1}{2} (\vec{a}_I + \vec{a}_{I+1})\Delta t \quad (2.61)$$

Cet algorithme de propagation présente tous les avantages : il est précis jusqu'à l'ordre 3 de dérivation de la position, et il présente explicitement la vitesse et la position au même temps.

Comme cités ci-dessus, il existe différents ensembles statistiques dans lesquels différentes grandeurs sont définies comme constantes. L'ensemble naturel généré par les simulations de dynamique moléculaires est l'ensemble micro canonique (NVE), dans lequel le nombre de particule (N), le volume (V) et l'énergie totale (E) sont constants. Il arrive cependant que l'utilisateur désire travailler dans un ensemble dont la température est constante (NVT) ou dans un ensemble dont la température et la pression sont constantes (NPT). Ces ensembles sont notamment utilisés afin de reproduire les conditions rencontrées en laboratoire et donc de s'approcher de l'expérience.

Rappelons que la température T est proportionnelle à la moyenne de l'énergie cinétique du système :

$$\langle E_{cin} \rangle = \sum_{i=1}^N \frac{1}{2} m_i v_i^2 = \frac{3}{2} NkT \quad (2.62)$$

Avec N la somme des « i » particules du système et k la constante de Boltzmann. Au cours d'une simulation dans l'ensemble NVE, l'énergie cinétique étant la différence entre l'énergie totale et l'énergie potentielle du système, celle-ci va inévitablement fluctuer, entraînant de ce fait la fluctuation de la température du système. Il existe plusieurs moyens de contraindre la température et de la rendre constante, le plus simple étant de multiplier la vitesse de chaque particule par un facteur permettant de retrouver la température désirée, tous les « n » pas de dynamique. Une telle correction peut amener des artéfacts importants sur le comportement du système, notamment un comportement périodique qui n'est pas représentatif de la réalité.

Dans ce travail, les systèmes étudiés en dynamique moléculaire seront couplés à un thermostat de Langevin afin de contraindre la température à une valeur constante. A cette fin, les équations de Newton sont corrigées de la manière suivante :

$$m \frac{d^2 R}{dt^2} = -\zeta \frac{dR}{dt} + F_{\text{système}} + F_{\text{aléatoire}} \quad (2.63)$$

L'accélération de toutes les particules est redéfinie comme une somme d'un terme de friction proportionnel à la vitesse de toutes les particules, de toutes les forces normalement définies du système et d'une force aléatoire. Le terme de friction aura pour effet d'enlever de l'énergie au système, et la force aléatoire sera présente pour compenser la perte d'énergie du système. Le facteur  $\zeta$  est appelé « damping coefficient » et peut être considéré comme un facteur de viscosité du système.

Concernant la pression, elle peut être maintenue constante par le biais d'un barostat, similaire au thermostat définit ci-dessus<sup>26</sup>, que l'on appelle piston de Langevin.

#### b. Définition de l'énergie potentielle : utilisation de champs de force

Nous avons défini les équations de propagation des positions, vitesses et accélérations du système, discuté des méthodes de régulation de la température et de la pression, il nous reste maintenant à décrire la méthode de calcul de l'énergie potentielle du système. L'obtention du potentiel  $V$  est un peu différente de celle qui a été décrite dans la partie concernant les méthodes de mécanique quantique et est basée sur l'utilisation d'un champ de force. On peut séparer ce potentiel en deux parties, les termes liés et non-liés :

$$V(\{\vec{R}_I\}) = V_{\text{liés}} + V_{\text{non-liés}} \quad (2.64)$$

La définition de ces deux potentiels dépend de la catégorie de champ de force utilisé. Dans le cas d'un champ de force dit de première génération, le potentiel lié est défini à travers les liaisons covalentes successives allant de 2 jusqu'à 4 noyaux :

$$V_{\text{liés}} = V_{\text{liaison}} + V_{\text{angle}} + V_{\text{dihedre}} + V_{\text{impropre}} \quad (2.65)$$

Ces termes sont développés de la manière suivante :

$$V_{liaison} = \frac{1}{2} \sum_{i,j \in liaisons} k_{ij} (d_{ij} - d_{ij}^0)^2 \quad (2.66)$$

$$V_{angle} = \sum_{i,j,k \in angles} k_{ijk} (\theta_{ijk} - \theta_{ijk}^0)^2 \quad (2.67)$$

$$V_{dihedre} = \sum_{i,j,k,l \in angles} k_{ijkl} [1 + \cos(n\phi_{i,j,k,l} + \omega)] \quad (2.68)$$

$$V_{impropre} = \sum_{i,j,k,l \in plans} k'_{i,j,k,l} (\phi_{ijkl} - \phi_{ijkl}^0)^2 \quad (2.69)$$

Ici,  $V_{bond}$  est représentatif de la déformation d'une liaison  $d_{ij}$  entre deux noyaux  $i$  et  $j$ ,  $V_{angle}$  de la déformation de l'angle  $\theta_{ijk}$  entre 3 noyaux  $i$ ,  $j$  et  $k$ ,  $V_{dihedre}$  de la rotation d'angle  $\phi_{i,j,k,l}$  autour d'une liaison entre deux noyaux  $j$  et  $k$  et enfin  $V_{impropre}$  de la déformation entre les plans formés par les noyaux  $i$ ,  $j$  et  $k$  et  $j$ ,  $k$  et  $l$  respectivement, dont l'angle entre les plans est  $\phi_{ijkl}$ . Schématiquement, on peut représenter ces termes comme ceci :

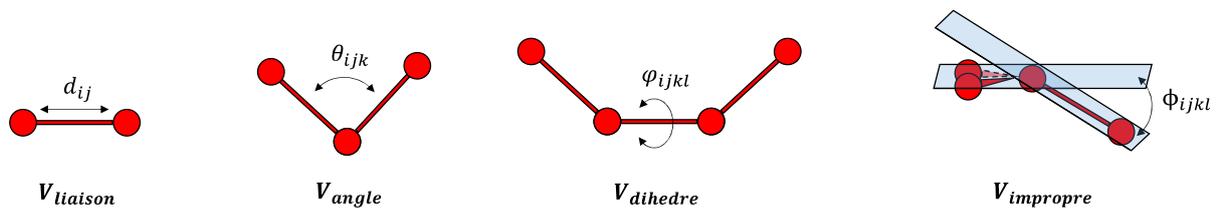


Figure 35: Représentation des différents termes du potentiel d'interactions intramoléculaires.

Concernant le potentiel non-liés, il contient deux contributions majeures :

$$V_{non-liés} = V_{VDW} + V_{Elec} \quad (2.70)$$

Elles représentent respectivement le potentiel de Van der Waals et le potentiel électrostatique coulombien :

$$V_{VDW} = \sum_{i=1}^M \sum_{j>i}^M \left[ \left( \frac{A_{ij}}{r_{ij}} \right)^{12} - \left( \frac{B_{ij}}{r_{ij}} \right)^6 \right] \quad (2.71)$$

$$V_{Elec} = \sum_{i=1}^M \sum_{j>i}^M \frac{q_i q_j}{4\pi \epsilon_0 r_{ij}} \quad (2.72)$$

Ces deux contributions représentent les interactions de répulsion-dispersion et les interactions électrostatiques entre les noyaux du système.

Les constantes de forces  $\{k_{ij}, k_{ijk}, k_{ijkl}, k'_{i,j,k,l}\}$ , les paramètres à l'équilibre  $\{d_{ij}^0, \theta_{ijk}^0, \phi_{ijkl}^0\}$  ainsi que les paramètres du potentiel de Van der Waals  $\{A_{ij}, B_{ij}\}$  ont été ajustés sur des résultats expérimentaux ou obtenus par mécanique quantique. Un champ de force est par définition une librairie qui contient tous ces paramètres, pour un nombre donné d'atomes.

Dans cette thèse, nous utilisons le champ de force AMBER, notamment au travers de ses deux composantes GLYCAM06 et LIPID14 développés pour la modélisation des polysaccharides et des lipides, respectivement. Parmi ses particularités notables, ce champ de force comprend un terme additionnel de type VDW pour la définition des liaisons hydrogène<sup>27</sup>.

De ce fait, là où la mécanique quantique nécessite la prise en compte des  $n$  électrons présents sur un noyau, la dynamique moléculaire s'affranchit de ces degrés de liberté, en les incluant dans le calcul de l'énergie potentiel sous forme de paramètres empiriques. L'utilisation des champs de force permet donc de réaliser un très grand nombre de calculs sur des systèmes beaucoup plus grands ( $\sim 10^5$ - $10^6$  atomes) que ceux étudiés par des approches de mécanique quantique comme la DFT.

### c. Limites de la méthode

La dynamique moléculaire comporte cependant des failles importantes, dont une sera soulignée lors de la description de ce travail de thèse. Tout d'abord, les paramètres étant ajustés sur des résultats de mécanique quantique, ils sont dépendant du système étudié. Une des hypothèses sous-jacentes à cet ajustement réside dans la transférabilité des paramètres d'un système moléculaire à un autre, ce qui n'est pas toujours vrai. Un autre problème dont il faut avoir conscience lors de l'utilisation de ce type de méthode est la prise en compte de l'ergodicité du système. Pour décrire ce problème, nous allons nous concentrer sur une observable macroscopique quelconque  $A_{macro}$  décrite par l'équation suivante :

$$A_{macro} = \langle A \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(t) dt \quad (2.73)$$

D'après cette équation, pour obtenir une représentation correcte de l'observable macroscopique, il faut l'échantillonner sur une durée infinie. En pratique nous nous contentons d'échantillonner le système sur une durée supérieure, de plusieurs ordres de grandeur, au temps caractéristique de la propriété étudiée. Si en général, cela reste suffisant pour obtenir une bonne approximation de  $A_{macro}$ , ceci implique de manière inévitable un échantillonnage lacunaire de la surface d'énergie potentielle. Il existe cependant des méthodes permettant un échantillonnage amélioré, telles que le recuit simulé ou l'échange de répliques, que nous ne développerons pas plus.

## Références

- (1) Pauli, W. Zur Frage Der Theoretischen Deutung Der Satelliten Einiger Spektrallinien Und Ihrer Beein-Flussung Dutch Magnefisehe Felder. *Naturwissenschaften* **1924**, 741–743.
- (2) Rabi, I. I.; Millman, S.; Kusch, P.; Zacharias, J. R. The Magnetic Moments of Li-6, Li-7 and F-19. *Phys. Rev.* **1938**, 53 (1933), 495.
- (3) Rabi, I. I.; Zacharias, J. R.; Millman, S.; Kusch, P. A New Method of Measuring Intensities of Magnetization. *Phys. Rev.* **1939**, 55, 318. [https://doi.org/10.1016/S0031-8914\(39\)80004-0](https://doi.org/10.1016/S0031-8914(39)80004-0).
- (4) Bloch, F. Nuclear Induction. *Phys. Rev.* **1946**, 70 (7–8), 460–474. <https://doi.org/10.1103/PhysRev.70.460>.
- (5) Purcell, E. M.; Torrey, H. C.; Pound, R. V. Resonance Absorption by Nuclear Magnetic Moments in a Solid. *Phys. Rev* **1946**, 69, 37–38.
- (6) Proctor, W. G.; Yu, F. C. The Dependence of a Nuclear Magnetic Resonance Frequency upon Chemical Compound. *Phys. Rev.* **1950**, 77, 717.
- (7) Dickinson, W. C. Dependence of the F19 Nuclear Resonance Position on Chemical Compound [32]. *Phys. Rev.* **1950**, 77 (5), 736–737. <https://doi.org/10.1103/PhysRev.77.736.2>.
- (8) Andrew, E. R.; Bradbury, A.; Eades, R. G. Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation. *Nature* **1959**, 183, 1802–1803.
- (9) Ernst, R. R.; Anderson, W. A. Application of Fourier Transform Spectroscopy to Magnetic Resonance. *Rev. Sci. Instrum.* **1966**, 37 (1), 93–102. <https://doi.org/10.1063/1.1719961>.
- (10) Goldman, M. Histoire de La RMN: Autour de Ses Premiers Acteurs... et Après. *Actual. Chim.* **2004**, 1956 (273), 57–60.
- (11) Wüthrich, K.; Wagner, G. NMR Investigations of the Dynamics of the Aromatic Amino Acid Residues in the Basic Pancreatic Trypsin Inhibitor. *FEBS Lett.* **1975**, 50 (2), 265–268. [https://doi.org/10.1016/0014-5793\(75\)80504-7](https://doi.org/10.1016/0014-5793(75)80504-7).
- (12) Kumar, A.; Ernst, R. R.; Wuthrich, K. A Two-Dimensional Nuclear Overhauser Enhancement (2D NOE) Experiment for the Elucidation of Complete Proton-Proton Cross-Relaxation Networks in Biological Macromolecules. *Biochem. Biophys. Res. Commun.* **1980**, 95 (1), 1–6.
- (13) Haeberlen, U. *Advances in Magnetic Resonance*; Waugh, J. S., Ed.; Academic Press, 1976.
- (14) Mehring, M. *Principles of High Resolution NMR in Solids*; Springer Verlag, 1983.
- (15) Spiess, H. W. *NMR Basic Principles and Progress*; Diehl, P., Fluck, E., Kosfeld, R., Eds.; Springer Verlag, 1978.

- (16) Hartmann, S. R.; Hahn, E. L. Nuclear Double Resonance in the Rotating Frame. *Phys. Rev.* **1962**, *128* (5), 2042–2053. <https://doi.org/10.1103/PhysRev.128.2042>.
- (17) Kaupp, M.; Bühl, M.; Malkin, V. G. *Calculation of NMR and EPR Parameters*; 2004. <https://doi.org/10.1002/3527601678>.
- (18) Kutzelnigg, W. Theory of Magnetic Susceptibilities and NMR Chemical Shifts in Terms of Localized Quantities. *Isr. J. Chem.* **1980**, *19* (1–4), 193–200. <https://doi.org/10.1002/ijch.198000020>.
- (19) Ditchfield, R. Molecular Orbital Theory of Magnetic Shielding and Magnetic Susceptibility. *J. Chem. Phys.* **1972**. <https://doi.org/10.1063/1.1677088>.
- (20) Wolinski, K.; Hinton, J. F.; Pulay, P. Efficient Implementation of the Gauge-Independent Atomic Orbital Method for NMR Chemical Shift Calculations. *J. Am. Chem. Soc.* **1990**. <https://doi.org/10.1021/ja00179a005>.
- (21) Pickard, C. J.; Mauri, F. All-Electron Magnetic Response with Pseudopotentials: NMR Chemical Shifts. *Phys. Rev. B* **2001**, *63* (24), 1–13. <https://doi.org/10.1103/physrevb.63.245101>.
- (22) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. Iterative Minimization Techniques for Ab Initio Total-Energy Calculations: Molecular Dynamics and Conjugate Gradients. *Rev. Mod. Phys.* **1992**, *64* (4), 1045–1097. <https://doi.org/10.1103/RevModPhys.64.1045>.
- (23) Pask, J. Linearized Augmented Planewave (LAPW) Method.
- (24) Sturdy, Y. K.; Clary, D. C. Torsional Anharmonicity in the Conformational Analysis of Tryptamine. *Phys. Chem. Chem. Phys.* **2007**, *9* (17), 2065–2074. <https://doi.org/10.1039/b615660f>.
- (25) Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ, United States, 2007. <https://doi.org/10.1177/072551369203300107>.
- (26) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103* (11), 4613–4621. <https://doi.org/10.1063/1.470648>.
- (27) Coupry, D. E.; Addicoat, M. A.; Heine, T. Explicit Treatment of Hydrogen Bonds in the Universal Force Field: Validation and Application for Metal-Organic Frameworks, Hydrates, and Host-Guest Complexes. *J. Chem. Phys.* **2017**, *147* (16). <https://doi.org/10.1063/1.4985196>.
- (28) Levitt, M. . *Spin Dynamics: Basics of Nuclear Magnetic Resonance*; 2000. <https://doi.org/10.1002/cmr.a.20130>.
- (29) Andrew, E. R.; Hinshaw, W. S.; Tiffen, R. S. NMR in Rapidly Rotated Metallic Aluminium and Cadmium. *J. Magn. Reson.* **1974**, *15* (2), 191–195. [https://doi.org/10.1016/0022-2364\(74\)90070-5](https://doi.org/10.1016/0022-2364(74)90070-5).

## **Travaux réalisés**



# Chapitre 1. Etudes théoriques de complexes d'amylose.

## 1. Contexte et objectifs du chapitre

Afin d'étudier les interactions polysaccharide-lipides d'un système biologique complexe, nous aurons besoin d'une méthodologie théorique solide. Pour s'en assurer, il est nécessaire de mettre en place une stratégie de calcul sur des systèmes simplifiés. Les complexes d'inclusion de type amylose-acide gras paraissent être un système de choix, en raison de l'ampleur de la littérature à leur sujet.

Comme nous l'avons précisé en introduction, l'amylose, polysaccharide essentiellement composé d'unités glucoses branchées en  $[\alpha 1 \rightarrow 4]$ , présente la particularité de se structurer différemment en milieu aqueux en fonction de la présence ou non d'autres molécules dans son environnement. En absence d'autre molécules, nous observerons une forme de double hélice appelée polymorphe A ou B selon les conditions de cristallisation, et en présence de petites molécules hydrophobes, ou de chaînes lipidiques, ce polysaccharide forme des hélices simples qui encapsulent ces dernières. Cette structure est appelée polymorphe V (Figure 36).

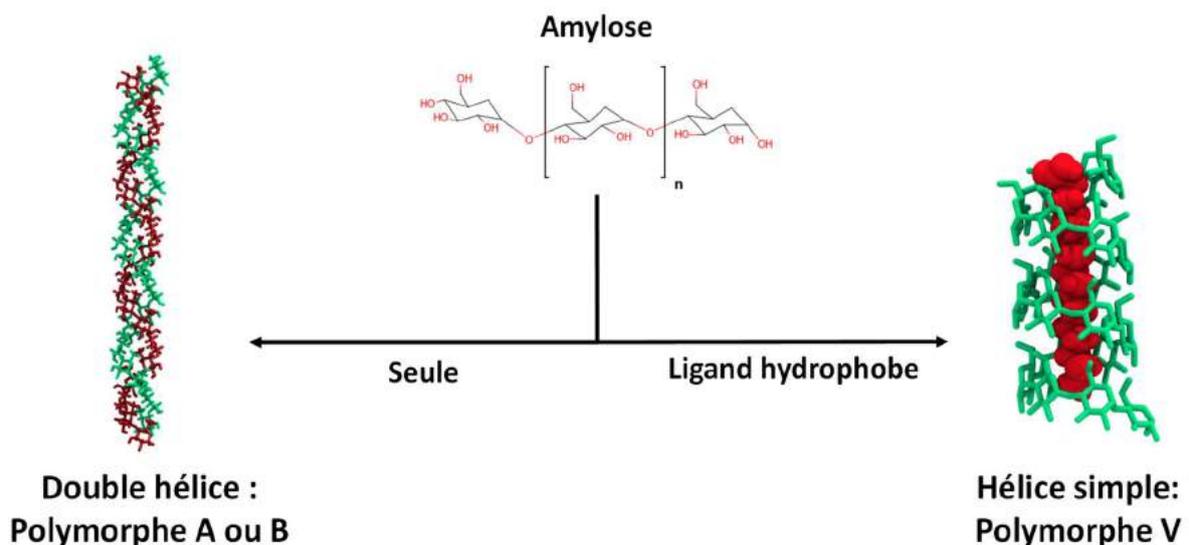


Figure 36: Structures formées par l'amylose en milieux aqueux, en absence ou présence d'un ligand hydrophobe.

Ces deux polymorphes peuvent être différenciés selon leur spectre RMN  $^{13}\text{C}$  respectifs, notamment au niveau du déplacement chimique des carbones formant la liaison glycosidique : le  $\text{C}_1$  et le  $\text{C}_4$  (Figure 37). En effet le signal du  $\text{C}_1$  est un pic unique présent à 102.5 ppm dans le polymorphe V et un double pic présent à 99.3 - 100.2 ppm dans le polymorphe B. Le signal du  $\text{C}_4$  est un pic unique à 82 ppm dans le polymorphe V et est compris dans le massif entre 70 et 76 ppm dans le cas du polymorphe B. La RMN est la méthode expérimentale utilisée dans le reste de ce travail de thèse, et nous souhaitons apporter des interprétations structurales supplémentaires grâce à l'utilisation de méthodes théoriques.

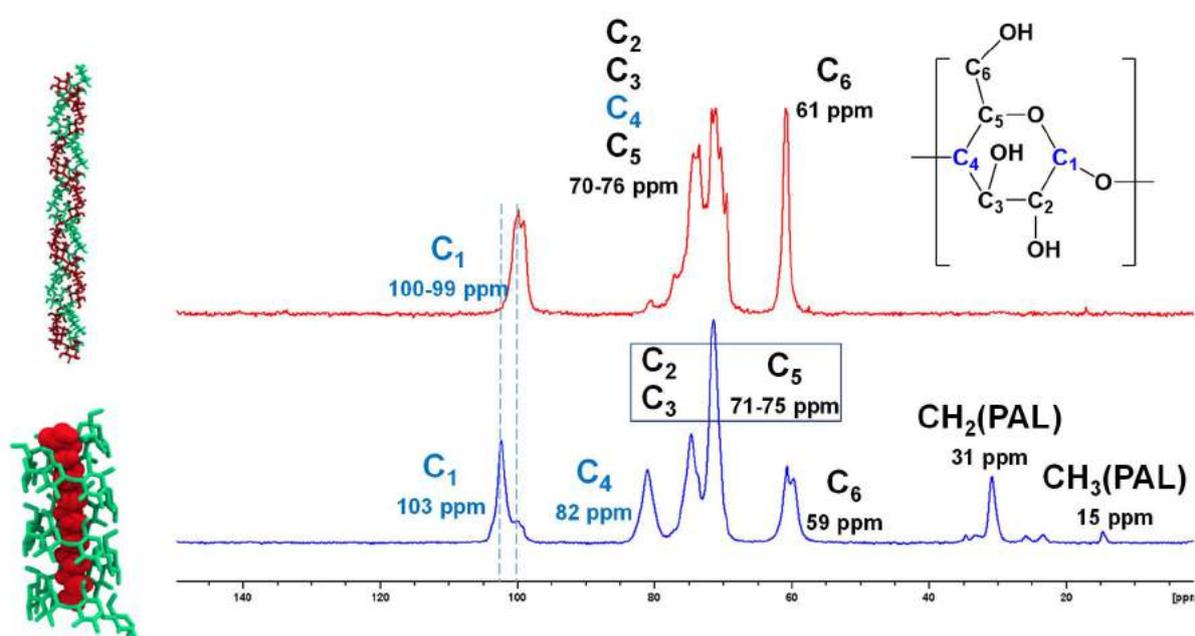


Figure 37: Exemple de spectres RMN  $^{13}\text{C}$  obtenus sur les polymorphes B (rouge) et V (bleu).

Nos objectifs étaient donc les suivants : établir un modèle chimique permettant d'étudier le comportement dynamique de ces complexes et réussir à reproduire les spectres RMN  $^{13}\text{C}$  expérimentaux déjà publiés sur ces deux structures.

Parmi les outils théoriques à notre disposition, il était nécessaire d'utiliser des méthodes de modélisation transférables sur des systèmes contenant plusieurs milliers d'atomes. Concernant le comportement de ces complexes, la dynamique moléculaire classique est ici un outil de choix car les logiciels et champs de force déjà développés ont fait preuve de leur robustesse, et la puissance de calcul disponible ne cesse

d'augmenter, permettant d'effectuer des calculs sur des systèmes de plusieurs centaines de milliers d'atomes.

Pour la description des paramètres des polysaccharides et des lipides, nous avons décidé d'utiliser les champs de force GLYCAM06 et LIPID14 respectivement. Nous avons tout d'abord vérifié la robustesse de ces derniers en réalisant une trajectoire de calcul dont la structure de départ du complexe présentait une chaîne d'amylose de 19 résidus repliée sous la forme d'un polymorphe V autour d'une molécule d'acide palmitique. En utilisant la version originale de GLYCAM06, la structure hélicoïdale se déforme au cours de la simulation, comme en atteste les conformations des résidus 5 et 11 (Figure 38, A). Un effet similaire avait déjà été observé concernant le champs de force GROMOS et avait engendré une modification de ce dernier. Les auteurs décidèrent d'augmenter le moment dipolaire des groupes hydroxyles libres en modifiant les charges partielles présentes sur les atomes concernés. Nous avons de ce fait réalisé des modifications similaires sur le champ de force GLYCAM06 en diminuant les charges des oxygènes des groupes hydroxyles et en ajustant celles des hydrogènes correspondants de manière à ce que l'amylose reste électriquement neutre. Suite à cela, nous avons pu observer une stabilité accrue du polymorphe V en milieu aqueux, les liaisons hydrogène intra-amylose étant ainsi renforcées au détriment de celles formées avec le solvant (Figure 38, B).

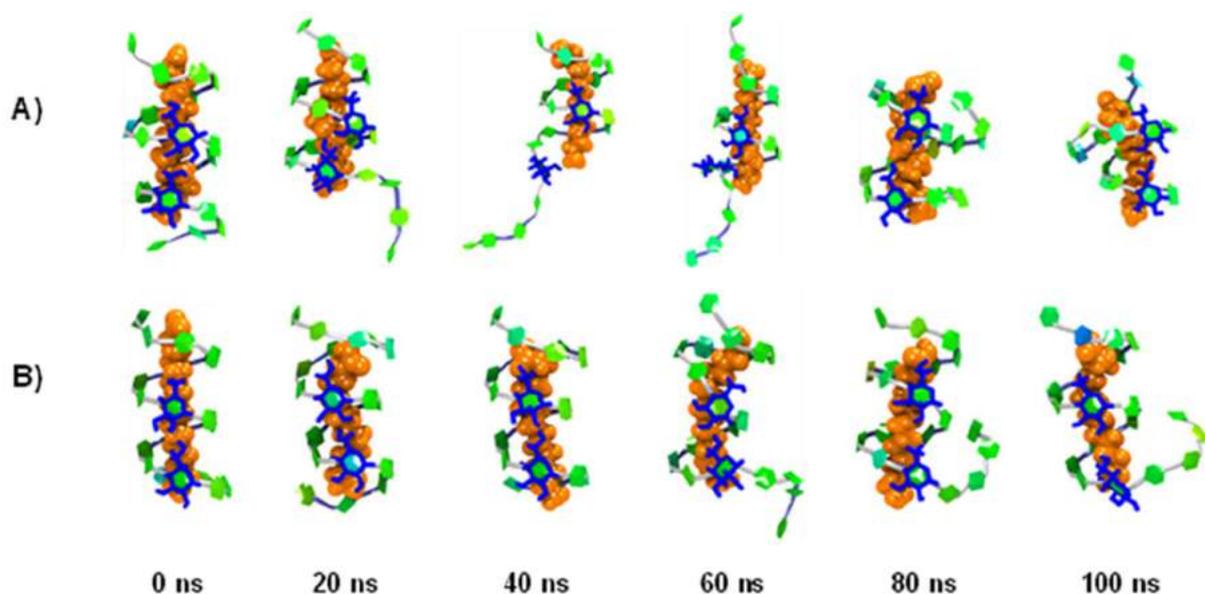


Figure 38: A) Simulation du complexe avec le champ de force GLYCAM06 original. B) Simulation du complexe avec la version modifiée de GLYCAM06. Orange : Acide palmitique. Polygones verts : résidus d'amylose. Bleu marine : résidus 5 et 11.

L'utilisation de cette version modifiée du champ de force nous a par la suite permis de mettre en évidence l'effet de la taille de l'amylose sur la stabilité du polymorphe V en solution.

Concernant le calcul de spectres RMN  $^{13}\text{C}$ , les méthodes de DFT présentées dans la partie méthode de ce manuscrit sont aussi connues pour leur robustesse, et sont dorénavant applicables sur des systèmes contenant des milliers d'atomes.

Le processus de calcul de paramètres RMN par ces méthodes est généralement divisé en deux parties : une optimisation de géométrie est tout d'abord réalisée sur le système et le calcul de paramètres RMN est ensuite réalisé sur la structure obtenue. En utilisant cette méthode, nous avons pu démontrer que l'effet de taille n'impactait pas seulement la stabilité du polymorphe V, mais impactait aussi le spectre RMN résultant du fait de la formation de nouvelles liaisons hydrogènes intra-amylose. Cependant, nous avons remarqué que si ces calculs permettaient une reproduction des déplacements chimiques des  $\text{C}_1$ ,  $\text{C}_4$  et  $\text{C}_6$ , les déplacements chimiques des  $\text{C}_2$ ,  $\text{C}_3$  et  $\text{C}_5$  étaient quant à eux légèrement surestimés. Nous avons émis l'hypothèse que le caractère dynamique du polymorphe pouvait être responsable de cet effet, comme cela avait déjà été démontré précédemment par notre groupe concernant le cholestérol. Nous avons donc décidé de réintroduire la dynamique du complexe dans nos calculs de RMN en les réalisant sur des structures extraites des trajectoires de dynamique moléculaire. Il en résulte une meilleure reproduction des déplacements chimiques des  $\text{C}_2$ ,  $\text{C}_3$  et  $\text{C}_5$ , et nous avons pu mettre en évidence qu'une partie du spectre RMN obtenu n'était pas due à des résidus localisés dans un polymorphe V, mais à ceux localisés dans une partie de l'amylose présentant des structures différentes.

Tous ces résultats ont été publiés dans le premier article de ce chapitre et soulèvent l'importance de l'aspect dynamique de ces complexes.

Cette méthodologie a également été appliquée à l'étude du polymorphe B qui est caractérisé par deux signaux RMN  $^{13}\text{C}$  caractéristiques de résidus glucoses dans deux environnements différents. Ces deux signaux ont pu être mis en évidence seulement sur un échantillon de haute cristallinité, et il a été postulé que les conditions cristallines pourraient être responsable de deux environnements différents. Dans le but d'éprouver notre méthodologie, nous avons effectué des calculs de RMN par DFT sur des

structures issues d'optimisations de géométrie et de trajectoires de dynamique moléculaire. Dans les deux cas, il nous a été impossible de différencier deux signaux RMN. Nous avons alors inclus les conditions périodiques dans notre calcul en utilisant l'approche GIPAW, brièvement décrite dans la partie méthode de ce manuscrit. L'utilisation de cette méthode nous a permis de différencier deux types signaux RMN distincts au sein des chaînes d'amylose. La corrélation a ensuite été réalisée avec leur positionnement dans le cristal : les résidus présents à l'interface entre deux doubles hélices au sein du cristal présentent un signal RMN calculé différent de ceux qui ne sont pas sur ces interfaces.

Ces résultats sont soumis pour publication et présentés dans le second article de ce chapitre et soulèvent l'importance des interactions entre hélices dans ces échantillons.

## 2. Article n°1

L'article intitulé « Structures and NMR spectra of short amylose-lipid complexes. Insight using molecular dynamics and DFT quantum chemical calculations » est présenté sous sa version finale pour publication dans *Carbohydrate Polymers* avant édition par le journal. Les informations supplémentaires peuvent être trouvées dans l'annexe n°2 de ce manuscrit.

# Structures and NMR spectra of short amylose-lipid complexes. Insight using molecular dynamics and DFT quantum chemical calculations

*Adrien Schahl<sup>1,2</sup>, Valérie Réat<sup>2</sup>, Franck Jolibois<sup>1</sup>*

1) LPCNO, Université Fédérale de Toulouse Midi-Pyrénées, INSA-CNRS-UPS, 135 avenue de Rangueil, 31077 Cedex 4 Toulouse, France

2) Institut de Pharmacologie et de Biologie Structurale, UMR 5089, CNRS - Université de Toulouse, UPS BP 64182, 205 route de Narbonne, 31077 Toulouse, Cedex 04, France

## **Introduction.**

A member of the family of polysaccharides, amylose is an essentially linear polymer composed of [ $\alpha$ -1 $\rightarrow$ 4]-linked glucose residues (Takeda, Shitaozono, & Hizukuri, 1990). It has been the subject of much attention and serves as a model for the study of more complex polysaccharides. The analysis of its interactions with other molecules is of great importance in many scientific fields: food chemistry, biological systems, pharmaceutical sciences (Carbinatto, Ribeiro, Colnago, Evangelista, & Cury, 2016). Amylose adopts specific conformations, including double helical and V-type helical forms, or unstructured conformations, depending on its environment (Burghammer et al., 2009; Gessler et al., 1999; Godet, Buléon, Tran, & Colonna, 1993; Horii, Yamamoto, Hirai, & Kitamaru, 1987; Le Bail, Rondeau, & Buléon, 2005; Nuessli, Putaux, Le Bail, & Buléon, 2003; Takahashi, Kumano, & Nishikawa, 2004). In particular, it has been demonstrated that amylose chains fold around guest hydrophobic molecules and in such complexes adopts V-helix conformations (Brisson, Chanzy, & Winter, 1991; Godet et al., 1993; Nuessli et al., 2003; Rappenecker & Zugenmaier, 1981; Whittam et al., 1989).

An experimental description of the structure of amylose and of its interactions with other molecules relies on X-ray diffraction and NMR studies. NMR experiments revealed specific  $^{13}\text{C}$  chemical shifts that are characteristic of the V-helix form of amylose complexes (Gidley & Bociek, 1988; Veregin, Fyfe, & Marchessault, 1987; Zabar, Lesmes, Katz, Shimoni, & Bianco-Peled, 2009). Numerous efforts have sought to characterise the structure of amylose using molecular dynamics approaches (Cheng et al., 2018; Cheng, Zhu, Hamaker, Zhang, & Campanella, 2019; Feng et al., 2015, 2017; López, De Vries, & Marrink, 2012; Tusch, Krüger, & Fels, 2011). However, the theoretical link between this specific V-helix conformation and its associated chemical shifts has not previously been explored. The precise calculation of NMR parameters using quantum chemistry methods is now possible for large molecular systems. The

calculation of chemical shifts provides theoretical support to experimental data, facilitating resonance assignment or the prediction of other parameters. In addition, the atomistic level of calculation coupled with the ability to include solvent effects and/or to take into account specific interactions with other molecules, allows experimental results to be interpreted in terms of conformational parameters and the effects of interactions.

To the best of our knowledge, the study of polysaccharides using quantum chemistry methods has focused mainly on small systems (Toukach & Ananikov, 2013). These may provide insight into the variation of chemical shifts in terms of amylose structure but conformational properties alone are not sufficient to fully understand the changes in chemical shifts that occur when amylose adopts a specific form such as the V-helix. In this case, a sufficiently large system is needed to include all the factors that influence the calculation of chemical shielding. In addition, since the amylose helix undergoes fluctuations in structure as function of time (Cheng et al., 2018; Feng et al., 2015; López et al., 2012; Tusch et al., 2011) we must take into account the influence of such dynamics on NMR spectra.

We present here a detailed study of the structure and NMR parameters of amylose chains of increasing size and their interactions with palmitic acid. Analysis of the stability and complexation process of these molecular structures was undertaken using molecular dynamics (MD) simulations. In addition,  $^{13}\text{C}$  chemical shifts were calculated using Density Functional Theory (DFT) methods. In order to establish a computational strategy applicable to polysaccharide-host complexes of various sizes, the choice of structures used for the calculation of the NMR parameters was analyzed: either a structure calculated by geometry optimization at the DFT level or a set of 50 structures extracted from MD simulations was used. We show that longer polymers tend to stabilize the V-helix formation and their calculated spectra are closer to the experiment. Moreover, our calculations show that the incorporation of flexibility

(on nanosecond time scale explored) is essential for the accurate reproduction of experimental spectra.

## 2. Computational details.

### 2.1. Classical molecular dynamics.

The starting structure of palmitic acid was created using Avogadro molecular software (Hanwell et al., 2012). Two types of amylose chain were constructed. The first was constructed as starting point for studying the stability of amylose-fatty acid complex and consisted of a 19-residue chain folded around a molecule of palmitic acid, in order to obtain a structure similar to that determined crystallographically (Rappenecker & Zugenmaier, 1981). The second type was constructed to analyse the complexation process according to the size of the amylose and consisted of an unfolded amylose chain of 7, 9, 11, 13, 15, 17 or 19 residues in interaction with palmitic acid. The C<sub>8</sub> carbon of palmitic acid was positioned at 5.0 Å from the centre of mass of the 4<sup>th</sup> or 10<sup>th</sup> residue of the amylose chain of less than or greater than 13 residues, respectively. In both cases, the LEaP program of AMBER suite (Case et al., 2018) was used. For the folded amylose, dihedral angles  $\phi$  (O5-C1-O4'-C4') and  $\psi$  (C1-O4'-C4'-C3') were set to  $107.5^\circ \pm 4.5^\circ$  and  $110.8^\circ \pm 5.5^\circ$ , respectively (Figure 1; primes denote atoms of the preceding residue). For the unfolded structures, dihedral angles  $\phi$  and  $\psi$  angles were arbitrary set to  $93.4^\circ \pm 0.1^\circ$  and  $93.8^\circ \pm 0.1^\circ$ , respectively. The amylose-palmitic acid systems were solvated in a box of water such that the solute was placed set at 20 Å from the box boundaries. Box sizes ranged from 60x68x63 Å<sup>3</sup> (6,900 water molecules) to 86x85x86 Å<sup>3</sup> (17,340 water molecules) for amylose of 7 and 19 residues, respectively.

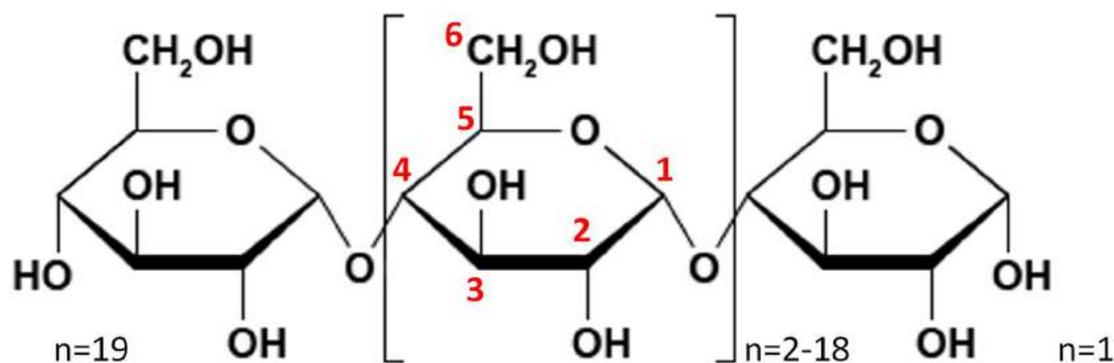


Figure 1: Carbon atom numbering of glucose residues.

All MD simulations were performed using NAMD software (Phillips et al., 2005). Langevin dynamics were used to control the temperature and, for simulations within the isothermal-isobaric (NPT) ensemble, Nose-Hoover Langevin piston was employed to maintain the pressure at a constant value of 1 atmosphere. The propagation of the atoms was carried out using the velocity Verlet integration method with a time step of 1 fs. The energy gradients required to solve the equations of motion were computed using GLYCAM-06 (Kirschner et al., 2008) and LIPID-14 (Dickson et al., 2014) force fields for the description of amylose and palmitic acid, respectively. A cut-off of 15 Å and a switching function starting at 12 Å were applied for non-bonded interactions. The Particle Mesh Ewald (Darden et al., 1993) method was used for the complete electrostatic calculation using periodic boundary conditions. TIP3P model (Jorgensen, Chandrasekhar, Madura, Impey, & Klein, 1983) was used to describe water molecules. All MD simulations followed the same computational protocol: After a short 1000 cycle energy minimization procedure using a conjugate gradient algorithm, a slow increase in temperature from 5 K to 295 K in increments of 10 K was carried out with a 5 ps simulation at each temperature. The MD equilibration was allowed to continue at 300 K, for 800 ps under the NPT ensemble. The effective production step was then performed for at least 95 ns in the canonical ensemble (NVT) by saving atomic coordinates every picosecond.

## 2.2. Quantum chemical calculations.

Quantum chemistry calculations were performed using the Gaussian 09 suite of programs (Frisch et al., 2009). Where necessary, geometry optimization was performed at the DFT level using the B3LYP hybrid functional (Becke, 1993; Lee, Yang, & Parr, 1988) associated to the Pople double-zeta basis set augmented by a polarisation function on all atoms, namely 6-31G(d,p) (Krishnan, Binkley, Seeger, & Pople, 1980; McLean & Chandler, 1980).

$^{13}\text{C}$  isotropic chemical shieldings were computed using the GIAO approach (Ditchfield, 1972; Wolinski, Hinton, & Pulay, 1990) at the B3LYP DFT level using the 6-31G(d,p) basis set. These were converted to chemical shifts using a value of 191.89 ppm for the chemical shielding of TMS (Jolibois, Soubias, Réat, & Milon, 2004). Structures obtained by full geometry optimisation at the DFT level or extracted from MD simulations were used for NMR computation. In the latter case, 50 structures were extracted every 200 fs from the last 10 to 20 ns of each trajectory depending on the size of the amylose under study.

The left-handed helix of V-amylose has been intensively characterised (Eliasson & Krog, 1985; Godet et al., 1993; Sarko & Winter, 1974). While a right-handed form may be considered from structural considerations, it has rarely been observed experimentally, to the best of our knowledge (Bulpin, Welsh, Morris, & England, 1982). The comparison of the relative energies of left and right handed  $V_6$ -type helix and  $V_7$ -type left handed helix (with 6 or 7 glucoses per turn, respectively) was performed at the DFT level for helices containing 19 units. These calculations confirmed that left-handed forms are the most stable (see Supplementary information) and only these were considered further.

## 2.3. Simulation of NMR spectra.

In order to simulate  $^{13}\text{C}$  NMR spectra, a summation procedure was developed.

For the complete quantum chemistry calculations (geometry optimisation and NMR calculation), the calculated chemical shift for each carbon was replaced by a Gaussian function

centered on the chemical shift value  $\delta_i$ . The whole spectrum was then obtained by summing these Gaussian functions:

$$I_{QM}(\delta_{spec}) = \sum_i^N e^{-\frac{(\delta_{spec}-\delta_i)^2}{\Delta}} \quad \text{Eq.1}$$

where N is the number of carbons and  $\Delta$  was set to 2.25 ppm in order to obtain the required full width at half height at 2.5 ppm, estimated from experimental data (Gidley & Bociek, 1988).

Where conformations were extracted from MD simulations, we first calculated the average value of the chemical shift of each carbon along the trajectory  $\langle \delta_i \rangle$ . This value was used instead of  $\delta_i$  in the previous equation in order to calculate the entire spectrum.

### 3. Results and discussion.

#### 3.1 Force field adaptation and validation in stability and complexation analyses.

##### 3.1.1. Stability analysis.

Atoms	Original Force Fields		Modified Force Fields		TIP3P
	GROMOS	GLYCAM 06	GROMOS	GLYCAM 06	
O <sub>2</sub>	-0.642	-0.713	-0.842	-0.842	-0.843
O <sub>3</sub>	-0.642	-0.699	-0.842	-0.842	
O <sub>6</sub>	-0.642	-0.682	-0.842	-0.842	
H <sub>2</sub>	0.410	0.437	0.610	0.566	0.417
H <sub>3</sub>	0.410	0.427	0.610	0.570	
H <sub>6</sub>	0.410	0.418	0.610	0.578	

Table 1: Comparison of original GROMOS and GLYCAM partial charges of selected hydroxyl group atoms with modified GROMOS, GLYCAM and with TIP3P partial charges (See figure 1 for atom numbering).

In order to validate force field parameters for MD calculations, an initial simulation was performed to analyse the stability of the left-handed V<sub>6</sub> helix form of a 19-residue amylose chain folded around a palmitic acid molecule. Starting from a folded configuration, the helical conformation is destabilised in less than 60 ns with the default parameters and charges in

GLYCAM06 and LIPID14. These parameters seemed unsuitable for reproducing both the experimental data (Godet et al., 1993) and the results of previous calculations (López et al., 2012). It has been demonstrated that, for the GROMOS force field, the dipolar moments of the H–O<sub>2</sub>, H–O<sub>3</sub> and H–O<sub>6</sub> bonds are underestimated (López et al., 2012). A modification of the partial charges for the atoms of these hydroxyl groups has been proposed in order to maintain the strength of the hydrogen bond network between (HO)<sub>6</sub> of residue [n] and (HO)<sub>2</sub> and (HO)<sub>3</sub> of residue [n+6]. The authors increased the charge difference between H and O from 1.05 e to 1.45 e by modifying the initial partial charges of these atoms, leading to a slightly higher charge difference than in the TIP3P water model (1.25 e (Jorgensen et al., 1983)). As the GLYCAM hydroxyl groups charge difference is similar to that in GROMOS (1.15 e and 1.05 e, respectively - see Table 1), we applied the same strategy and increased the partial charges on O<sub>2</sub>, O<sub>3</sub> and O<sub>6</sub> to -0.842. The partial charges of H<sub>2</sub>, H<sub>3</sub> and H<sub>6</sub> were adjusted to maintain the electro-neutrality of the molecule (see Table 1). The resulting system appears better “balanced” insofar as the competition for hydrogen bond formation either between glucose units or with the surrounding water molecules.

For a simulation carried out with these new partial charges, we first analysed the distribution of rotameric pairs around the glycosidic bond, using dihedral angles  $\varphi$  (O<sub>5</sub>-C<sub>1</sub>-O<sub>4</sub>'-C<sub>4</sub>') and  $\psi$  (C<sub>1</sub>-O<sub>4</sub>'-C<sub>4</sub>'-C<sub>3</sub>'). It has been shown experimentally that, in the case of a V-type helix, these angles have values of  $103.6^\circ \pm 12^\circ$  and  $115^\circ \pm 16^\circ$ , respectively (Gessler et al., 1999). For each set of atomic coordinates recorded along the trajectory, the 18 pairs of angles for the 19-residue chain were analysed (Figure 2a).

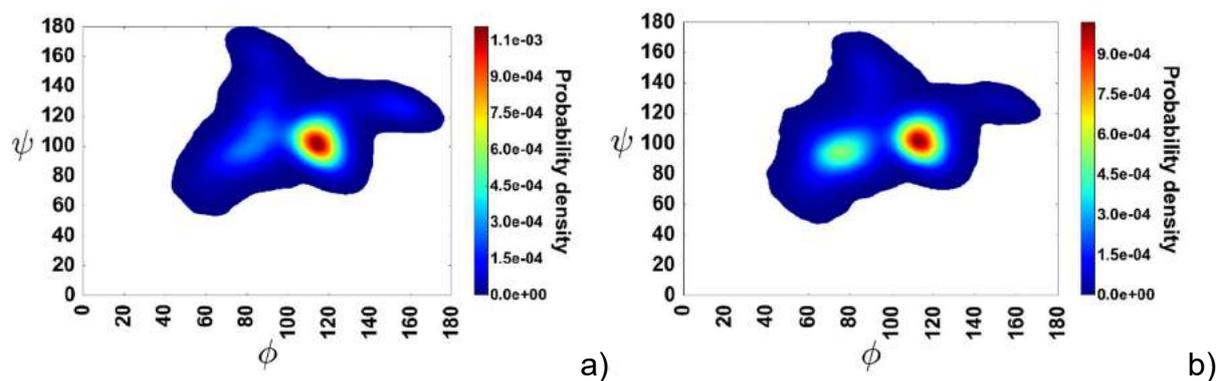


Figure 2. Ramachandran-like plot of  $\phi$  and  $\psi$  dihedral angles (in degrees) of a 19-residue amylose-palmitic acid complex. a) Analysis of the stability of V-helix type conformation. b) Analysis of the complexation process.

We show that these angles are mainly located in a region around  $\phi = 115^\circ$ ,  $\psi = 105^\circ$  in good agreement with previously reported data for this helical conformation (Cheng et al., 2018; Gessler et al., 1999; Tusch et al., 2011). The region of the Ramachandran diagram explored by amylose during the molecular dynamics simulations is extended and reflects the dynamic nature of the helix.

This is consistent with fluctuations of the  $V_6$ -pitch (i.e. variation of the distance between the centres of mass of residues  $[n]$  and  $[n+6]$  along the chain) along the trajectory (Figure S2). Experimentally, the pitch measured on the crystalline structure of the V-helix form of amylose is  $8.05 \pm 0.1 \text{ \AA}$  (Gessler et al., 1999; Murphy, Zaslów, & French, 1975; Rappenecker & Zugenmaier, 1981; Sarko & Zugenmaier, 2009). Depending on the position of the residues along the amylose chain, simulations show a more or less significant variation in  $V_6$ -pitch compared to experiment. This clearly indicates that the  $V_6$ -helix conformation is not fully preserved throughout the simulation. In order to obtain further insight into the stability of the structure, we have calculated the mean values of the  $V_6$ -pitch (Figure 3) for the entire trajectory of 100 ns and for the first and last 10 ns. Error bars were obtained by determining the standard deviations around these mean values. Changes in  $V_6$ -pitch imply changes in the position of 7

residues along the amylose chain and rely on extended fluctuations. They indicate that the entire amylose chain is in a  $V_6$ -helix conformation at the beginning of the simulation (Figures 3b–Red bars and 4a), but that during the last 10 ns, this  $V_6$ -helix conformation is broken, such that the average values of several  $V_6$ -itches are far larger than those of the experiment. This indicates that the nature of the helix evolves during time (as is it shown in figure 4b). Transitions between  $V_6$ -helix to another kind of helical conformation are allowed and distance variations reflect the flexibility of a molecule that is not constrained by a crystal packing.

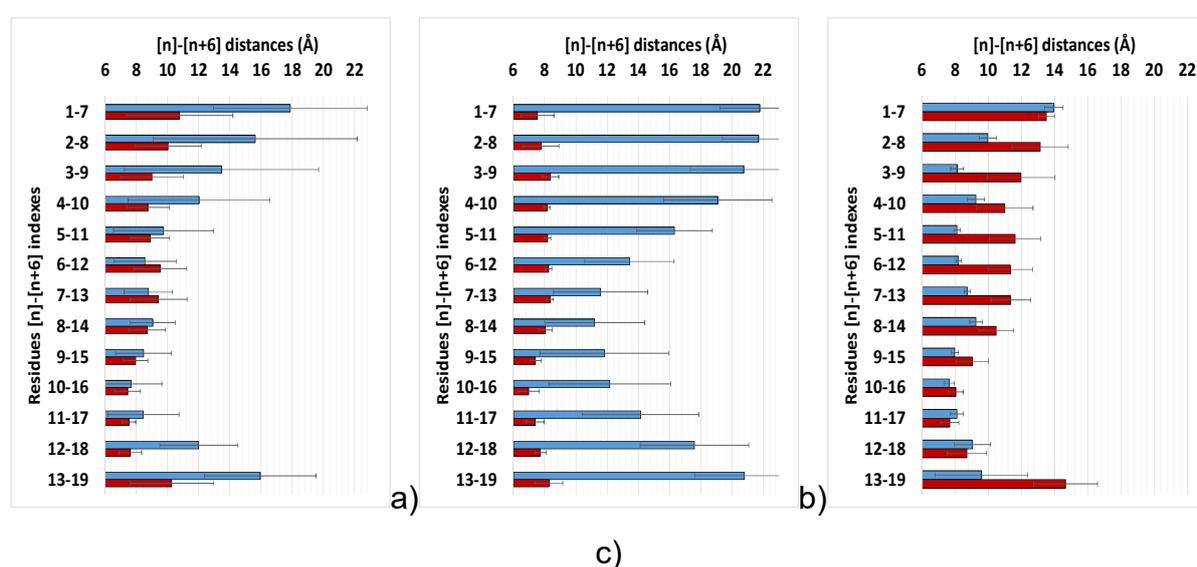


Figure 3: Mean values of the  $V_6$ -itches (angstroms) as a function of the position inside the amylose chain. Full NVT trajectory (a), first 10 ns of NVT trajectory (b) and last 10 ns of NVT trajectory (c). Red bars: stability process and Blue bars: complexation process.

### 3.1.2. Complexation analysis.

In studies of complexation, the starting conformation consists of an unfolded 19-residue chain of amylose located close to a molecule of palmitic acid. In less than 100 ns, complexation is observed with the folding of the amylose into a V-helix around the fatty acid, as in previous studies (Cheng et al., 2018; López et al., 2012). An analysis of the distribution of the  $V_6$ -itches as a function of simulation time (Figure S3) shows that while the amylose is unfolded at the

beginning of the dynamics trajectory (Figure 3b–Blue bars), the V-helix conformation is unambiguously characterized for the last 10 ns (Figures 3c–Blue bars and 4c).

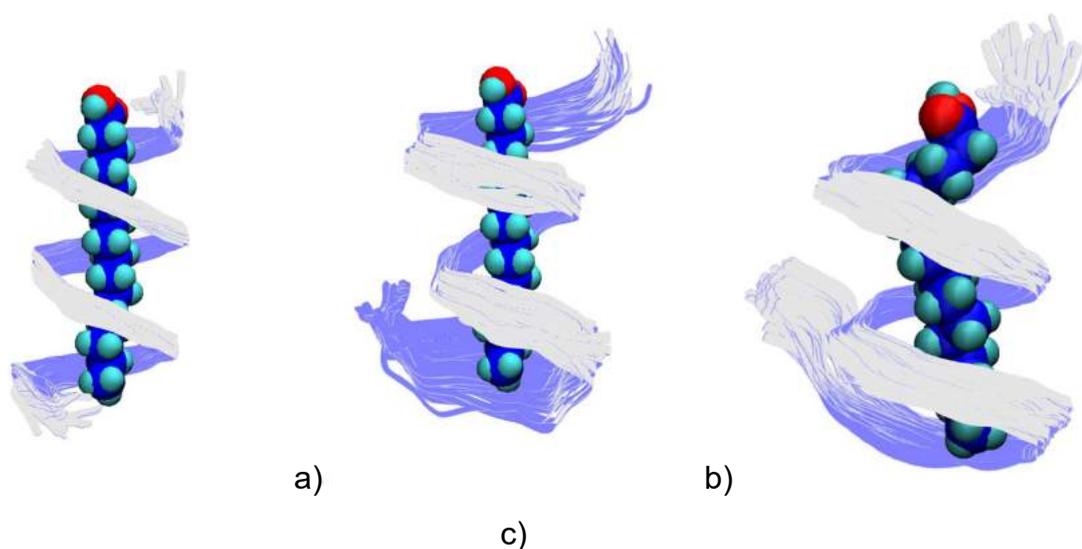


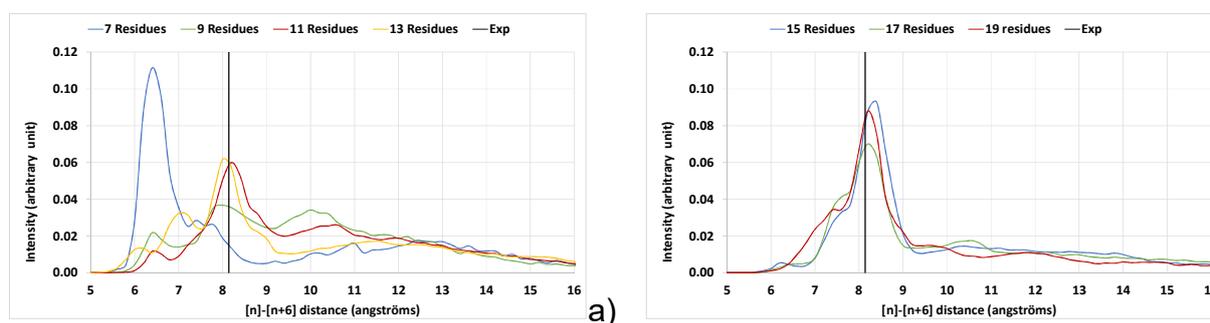
Figure 4: Superposition of 50 structures extracted from molecular dynamic simulations at regular intervals. a) First 10 ns of stability analysis; b) Last 10 ns of stability analysis; c) Last 10 ns of complexation analysis. Residue n=1 is at the bottom of each figure.

In the case of amylose complexed with palmitic acid, it appears that a minimum size of amylose is necessary to envelop the hydrophobic part of the fatty acid. It can be seen that residues at the ends of the helix are more flexible and unfold in solution. This suggests that the new dipolar momentum on the H–O bonds, introduced to stabilize the helix around palmitic acid is not strong enough to trap the glucose units in a V-helix conformation, in absence of a lipidic neighbourhood.

### 3.2 Size effect and stabilisation of V-helix conformation.

In order to identify the minimum size of amylose necessary to maintain the stability of the V-helix around the palmitic acid, the behaviour of chains of different lengths in the presence of fatty acid and in solution was investigated. The starting point of each simulation was an unfolded amylose chain (with 7, 9, 11, 13, 15, 17 residues) placed near the palmitic acid. For all molecular complexes,  $V_6$ -pitches distributions were plotted over the entire simulation

(Figure 5). For small amyloses, there is no clear evidence of the formation of a V-helix along the polymer (Figure 5a). With only 7 residues, the amylose exhibits  $V_6$ -itches that are too short compared to experiment ( $\sim 6.4 \text{ \AA}$  vs  $8.05 \pm 0.1 \text{ \AA}$ ). As the number of residues increases, this peak at  $6.4 \text{ \AA}$  disappears and a maximum appears around  $8.0 \text{ \AA}$  with increasing intensities as the amylose length increases. These distributions indicate that, above a minimum value of 11 residues, the amylose tends to structure locally in a helical conformation (Figures 5a and b) with the characteristic pitch of a  $V_6$ -helix.



b)

Figure 5:  $V_6$ -pitch distribution calculated for the whole NVT trajectory as a function of amylose length. (a) 7 to 13 residues; (b) 15 to 19 residues.

The difference in behaviour as a function of the length of amylose can be interpreted as a subtle balance between the need to achieve two distinct objectives: to envelop most of the hydrophobic portion of the lipid and to form an intramolecular hydrogen bond network that stabilizes the V-helix form. If the chain is small, these two objectives cannot be achieved simultaneously. For a larger number of glucose units (15 or more residues), the hydrophobic interactions and the formation of a hydrogen bonds network are both satisfied, making it possible to conserve the helix in a more sustainable manner.

### 3.3 Computation of $^{13}\text{C}$ NMR spectra from simulated amylose structures.

Since the structure of amylose often relies on  $^{13}\text{C}$  solid-state NMR, computation of these spectra was carried out for amylose-palmitic acid complexes as a function of amylose length.

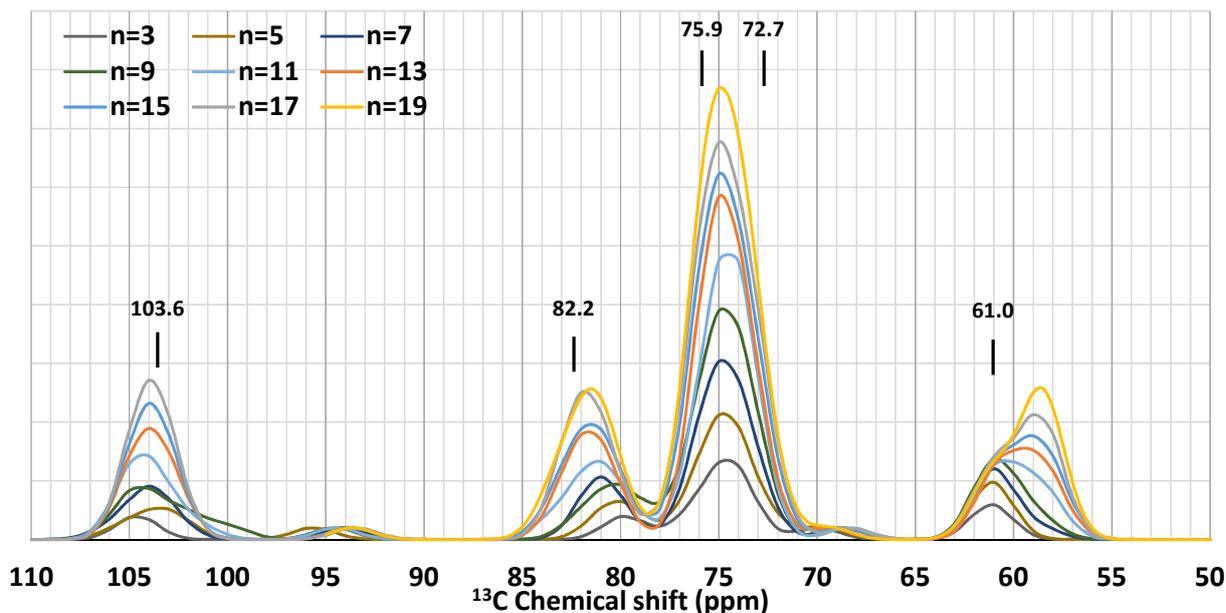


Figure 6:  $^{13}\text{C}$  theoretical NMR spectra of amylose in amylose-palmitic acid molecular complex as a function of amylose size (n number of glucose units). Reported values are experimental chemical shifts (Gidley & Bociek, 1988).

A full quantum chemistry approach was employed to validate the computational strategy and to highlight the structural parameters (conformation, H-bonding ...) that can help for the analysis of spectra. A set of molecular systems with amylose length varying from 3 to 19 glucose residues were investigated. From the 19-residue amylose folded around a palmitic acid molecule constructed for simulations, shorter amyloses were obtained by removing residues in pairs (1 residue from each side). Geometry optimizations followed by chemical shielding calculations were performed and  $^{13}\text{C}$  NMR spectra were simulated according to the summation procedure. The theoretical  $^{13}\text{C}$  spectra of amylose in the complex are plotted as a function of length (Figure 6). The increase in length results in a variation of the calculated chemical shifts from 104.6 to 103.8 ppm for anomeric- $\text{C}_1$  and from 80.0 to 81.7 ppm for  $\text{C}_4$  carbons, converging towards the values obtained experimentally for amylose/lipid complex (103.6 ppm and 82.2 ppm, respectively) (Gidley & Bociek, 1988). It should be noted that the  $\text{C}_1$  and  $\text{C}_4$  chemical shifts of the terminal residues are computed to lie between 94.5 and 93.6 ppm and 68.4 and 69.6

ppm, respectively. The C<sub>6</sub> carbon exhibits particular behaviour depending on whether or not its hydroxyl group, (HO)<sub>6</sub>, is involved in an intramolecular hydrogen bond. In our calculations, there is no solvent effect in order to model water. It has been shown that explicit water molecules would be necessary to reproduce such effects (Cuny, Jolibois, & Gerber, 2018) but this would increase computational costs significantly and induce convergence difficulties. As a result, several C<sub>6</sub> carbons belonging to residues at the reducing end of the amylose helix do not experience hydrogen bonding. Their chemical shift is greater (~61.0 ppm) than that of C<sub>6</sub> carbons for which (HO)<sub>6</sub> is engaged in the intramolecular hydrogen bonding network (~59.0 ppm) (Figure 6). In the region of C<sub>2</sub>, C<sub>3</sub> and C<sub>5</sub> chemical shifts where two characteristic peaks are present in experimental spectra (Gidley & Bociek, 1988), we observe a single broad peak centred at ~75 ppm. Reducing  $\Delta$  in Eq.1 reveals two peaks but with intensities inverted with respect to experiment: for the complex with 19-residue amylose, the chemical shifts of C<sub>2</sub> and C<sub>3</sub> are close to 75.4 ppm and only that associated with C<sub>5</sub> is about 73.8 ppm. A common method for correcting theoretical values with respect to experimental data is to use a linear fit to account for corrections arising from the slope and origin (essentially environmental effects that are not taken into account in the computational model) (Jolibois et al., 2004). For the complex with 19-residue amylose, such a fit performed using the chemical shifts of the unambiguously assigned carbons (C<sub>1</sub>, C<sub>4</sub> and C<sub>6</sub>) gave the following equation:  $\delta_{\text{corrected}}=0.9424*\delta_{\text{theo}}+5.4892$  with a  $R^2=0.9999$ . Correction of the entire spectrum resulted in an increase in the chemical shifts of C<sub>2</sub>, C<sub>3</sub> and C<sub>5</sub> carbons away from experimental values (Figure 7). Even with this correction, the entire <sup>13</sup>C NMR spectrum is not correctly reproduced suggesting that the computational strategy needs to be improved. While the method for computing NMR parameters is suitable for this type of system, our initial strategy fails to take into account the flexibility of the amylose chain. This can be introduced into our theoretical analysis via MD simulations. Carbon chemical shifts were calculated for each of 50 structures extracted from the last 10 ns of simulations performed

for the analysis of the complexation process and an averaged  $^{13}\text{C}$  NMR spectrum was simulated. The resulting carbon chemical shifts are all overestimated compare to the experiment. Anomeric  $\text{C}_1$  carbons resonate around 110.0 ppm; the chemical shift ranges for carbons  $\text{C}_2$ ,  $\text{C}_3$  and  $\text{C}_5$  and for  $\text{C}_4$  and  $\text{C}_6$  carbons are located around 80.0, 88.0 and 66.0 ppm, respectively. These discrepancies can be attributed to: 1) the force field parameters that produce structures different from the optimized structure; 2) the environmental effects; 3) unsuitability of the calculated reference (TMS) chemical shielding for chemical shift calculations using MD structures (Cuny et al., 2018; Gerber & Jolibois, 2015). As above, this spectrum may be corrected using a linear fit ( $\delta_{\text{corrected}}=0.8963*\delta_{\text{theo}}+2.7044$  with a  $R^2=0.9993$ ), leading to an overall improvement, in particular in the upfield shift of the region corresponding to  $\text{C}_2$ ,  $\text{C}_3$  and  $\text{C}_5$  atoms, toward that observed experimentally (i.e. between 70 and 79 ppm with a maximum at 73 ppm). This suggests that the description of the hydrogen bond network involving the  $\text{C}_2$ ,  $\text{C}_3$  and  $\text{C}_6$  carbons of residues  $[\text{n}]$  and  $[\text{n}+6]$  is better reproduced using a dynamical model than a single optimised conformation.

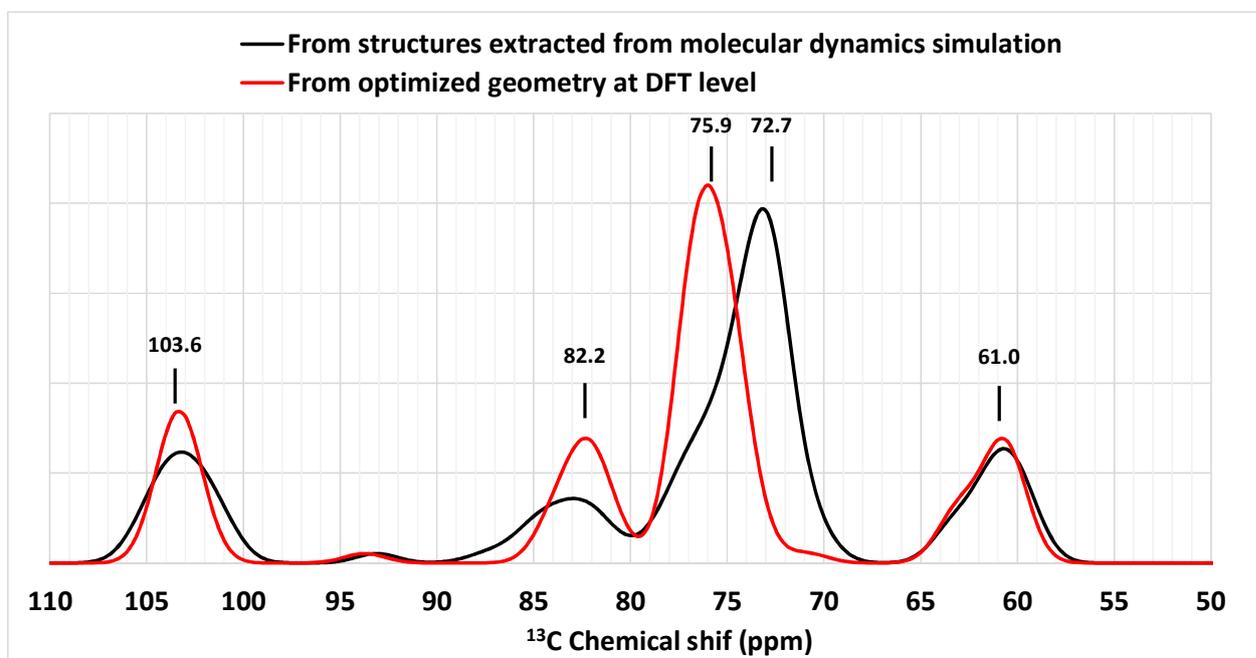


Figure 7: Corrected theoretical  $^{13}\text{C}$  NMR spectra of amylose-palmitic acid complexes with a 19 sugars amylose. Red line: calculated using optimized geometry at the DFT level; Black line:

calculated using 50 structures extracted from the last 10 ns of complexation analysis trajectory. Reported values are experimental chemical shifts (Gidley & Bociek, 1988).

Since we observed a different conformational behaviour during studies of the stability and the complexation process of amylose, the NMR spectra of 19-residue amylose in complex was simulated, using 50 structures extracted from the first and the last 10 ns of stability analysis molecular simulation. The resulting spectra do not show significant differences to that associated with the complexation process (Figure S4). The shoulder that appears around 76-77 ppm (Figure 7 and Figure S4) might be attributed to the resonance observed at 75.4 ppm in experimental spectra.

Spectral decomposition was undertaken by dividing the amylose into helical and non-helical portions using the conformational analysis performed previously (Figures 8, S5 and S6). If we consider the helical parts of amylose (residues 5 to 16, 3 to 10 and 6 to 16 for the last 10 ns of complexation, the first 10 ns of stability and the last 10 ns of stability analysis MD simulation, respectively), there is no major contribution of C<sub>2</sub>, C<sub>3</sub> or C<sub>5</sub> carbons to the spectrum above 75 ppm (peak maximum centred at 73 ppm). The theoretical resonance close to 76 ppm is however associated with carbon atoms 2 to 4 of the first four residues (Figure 8, blue line) for which a non-helical conformational has been highlighted in the case of the last 10 ns of complexation trajectory (Figure 6c). The experimental resonance at 75.4 ppm seems therefore not to be a specific feature related to V-helix conformation associated to C<sub>3</sub> carbon. The experimental assignments of the C<sub>2</sub>, C<sub>3</sub> and C<sub>5</sub> chemical shifts of amylose-lipid complexes have always been performed by analogy with cyclodextrin crystals and/or amylose solubilized in (CD<sub>3</sub>)<sub>2</sub>SO at high temperature (Gidley & Bociek, 1988). Some models have been postulated showing unstructured parts of the amylose in amorphous lamellas (Gelders, Goesart, & Delcour, 2005). Because these unstructured parts induce the modification of the hydrogen bonds network present in the V-helix, the 76 ppm resonance may be the fingerprint of another type of

interactions coming either from conformational modifications or from the presence of intermolecular interactions between V-type helices within a crystal lattice.

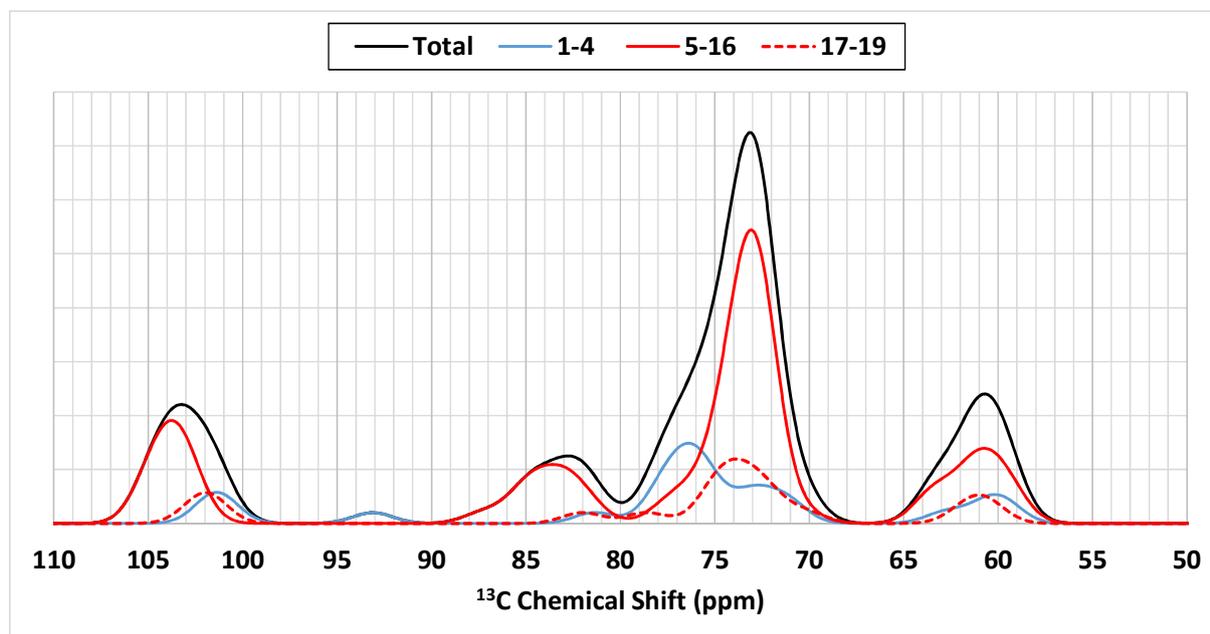


Figure 8: Decomposition of theoretical  $^{13}\text{C}$  NMR spectra of amylose-palmitic acid complexes with 19-residue amylose obtained from 50 structures extracted from the last 10 ns of complexation analysis MD simulation. Black line: Spectrum calculated using the full set of 19 residues; contributions to the spectrum for residues 1 to 4, 5 to 16 and 17 to 19 are given using plain blue line, plain red line and dotted red line, respectively.

#### 4. Conclusion.

We report intensive calculations using MD simulations to analyse the conformation of small amylose molecules complexed with palmitic acid. We show that a minimum of 11 residues is sufficient to observe the transitional appearance of a V-helix structure, but that 15 residues are necessary for the amylose to fold around the palmitic acid in a well-established helical conformation. A DFT method has been employed to simulate NMR spectra of various amylose-palmitic acid complexes. We demonstrate that some carbon chemical shifts are affected by the amylose length, depending on the presence of intramolecular hydrogen bonds. Decomposition

of NMR spectra of 19-residue amylose calculated using a series of structures extracted from molecular dynamics has allowed the precise assignment of each of the characteristic resonances. We propose that the double peak that appears, experimentally, in the region between 73 and 76 ppm is not a characteristic of V-helix conformation. On the contrary, we postulate that these peaks correspond to the presence of different local conformations or of specific intermolecular interactions between V-type helices within a crystal lattice.

### **Acknowledgements**

A.S. was a recipient of a PhD scholarship from University of Toulouse and the “Région Occitanie”. This work was supported by the CNRS-MITI grant “Modélisation du vivant” 2019. This work was granted access to the HPC resources of CALMIP supercomputing center under the allocation 2018-p0758 and of CINES and IDRIS supercomputing center under the allocation 2019-A0060810816 made by GENCI. The authors acknowledge Anne Lemassu at IPBS-Toulouse for fruitful discussions and Andrew Atkinson at King’s College London for the meticulous corrections of the manuscript.

### **References.**

- Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7), 5648–5652.
- Brisson, J., Chanzy, H., & Winter, W. T. (1991). The crystal and molecular structure of VH amylose by electron diffraction analysis. *International Journal of Biological Macromolecules*, 13(1), 31–39.
- Bulpin, B. P. V, Welsh, E. J., Morris, E. R., & England, B. (1982). Physical characterization of amylose-fatty acid complexes in starch granules and in solution. *Starch-Stärke*, 10, 335–

- Burghammer, M., Chanzy, H., Potocki-Véronèse, G., Putaux, J.-L., Montesanti, N., Popov, D., ... Buléon, A. (2009). Crystal structure of A-amylose: A revisit from synchrotron microdiffraction analysis of single crystals. *Macromolecules*, *42*(4), 1167–1174.
- Carbinatto, F. M., Ribeiro, T. S., Colnago, L. A., Evangelista, R. C., & Cury, B. S. F. (2016). Preparation and characterization of amylose inclusion complexes for drug delivery applications. *Journal of Pharmaceutical Sciences*, *105*(1), 231–241.
- Case, D. A., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, III, T. E., Cruzeiro, V. W. D., ... Kollman, P. A. (2018). AMBER 2018. San Francisco: University of California,.
- Cheng, L., Feng, T., Zhang, B., Zhu, X., Hamaker, B., Zhang, H., & Campanella, O. (2018). A molecular dynamics simulation study on the conformational stability of amylose-linoleic acid complex in water. *Carbohydrate Polymers*, *196*, 56–65.
- Cheng, L., Zhu, X., Hamaker, B. R., Zhang, H., & Campanella, O. H. (2019). Complexation process of amylose under different concentrations of linoleic acid using molecular dynamics simulation. *Carbohydrate Polymers*, *216*, 157–166.
- Cuny, J., Jolibois, F., & Gerber, I. C. (2018). Evaluation of gas-to-liquid  $^{17}\text{O}$  chemical shift of water: A test case for molecular and periodic approaches. *Journal of Chemical Theory and Computation*, *14*(8), 4041–4051.
- Darden, T., York, D., Pedersen, L., Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald : An  $N \log ( N )$  method for Ewald sums in large systems Particle mesh Ewald : An  $N - \log ( N )$  method for Ewald sums in large systems. *J. Chem. Phys.*, *100*89.
- Dickson, C. J., Madej, B. D., Skjerveik, Å. A., Betz, R. M., Teigen, K., Gould, I. R., & Walker, R. C.

- (2014). Lipid14: The amber lipid force field. *Journal of Chemical Theory and Computation*, *10*(2), 865–879.
- Ditchfield, R. (1972). Molecular orbital theory of magnetic shielding and magnetic susceptibility. *The Journal of Chemical Physics*, *56*(11), 5688–5691.
- Eliasson, A. C., & Krog, N. (1985). Physical properties of amylose-monoglyceride complexes. *Journal of Cereal Science*, *3*(3), 239–248.
- Feng, T., Li, M., Zhou, J., Zhuang, H., Chen, F., Ye, R., ... Fang, Z. (2015). Application of molecular dynamics simulation in food carbohydrate research - A review. *Innovative Food Science and Emerging Technologies*, *31*, 1–13.
- Feng, T., Wang, K., Zhuang, H., Bhopatkar, D., Carignano, M. A., Park, S. H., & Bing, F. (2017). Molecular dynamics simulation of amylose-linoleic acid complex behavior in water. *Journal of Nanoscience and Nanotechnology*, *17*(7), 4724–4732.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., ... Fox, D. J. (2009). Gaussian 09, Revision A.02. Wallingford CT: Gaussian, Inc.
- Gelders, G. G., Goesaert, H., & Delcour, J. A. (2005). Potato phosphorylase catalyzed synthesis of amylose-lipid complexes. *Biomacromolecules*, *6*(5), 2622–2629.
- Gerber, I. C., & Jolibois, F. (2015). Theoretical gas to liquid shift of  $^{15}\text{N}$  isotropic nuclear magnetic shielding in nitromethane using ab initio molecular dynamics and GIAO/GIPAW calculations. *Physical Chemistry Chemical Physics*, *17*(18), 12222–12227.
- Gessler, K., Uson, I., Takaha, T., Krauss, N., Smith, S. M., Okada, S., ... Saenger, W. (1999). V-Amylose at atomic resolution: X-ray structure of a cycloamylose with 26 glucose residues (cyclomaltohexacosaoose). *Proceedings of the National Academy of Sciences*, *96*(8),

4246–4251.

Gidley, M. J., & Bociek, S. M. (1988).  $^{13}\text{C}$  CP/MAS NMR studies of amylose inclusion complexes, cyclodextrins, and the amorphous phase of starch granules: Relationships between glycosidic linkage conformation and solid-state  $^{13}\text{C}$  chemical shifts. *Journal of the American Chemical Society*, *110*(13), 3820–3829.

Godet, M. C., Buléon, A., Tran, V., & Colonna, P. (1993). Structural features of fatty acid-amylose complexes. *Carbohydrate Polymers*, *21*(2–3), 91–95.

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, *4*(8), 1–17.

Horii, F., Yamamoto, H., Hirai, A., & Kitamaru, R. (1987). Structural study of Amylose polymorphs by cross-polarization-magic-angle spinning  $^{12}\text{C}$ -NMR spectroscopy. *Carbohydrate Research*, *160*, 29–40.

Jolibois, F., Soubias, O., Réat, V., & Milon, A. (2004). Understanding sterol-membrane interactions part I: Hartree-Fock versus DFT calculations of  $^{13}\text{C}$  and  $^1\text{H}$  NMR isotropic chemical shifts of sterols in solution and analysis of hydrogen-bonding effects. *Chemistry - A European Journal*, *10*(23), 5996–6004.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935.

Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., & Woods, R. J. (2008). GLYCAM06: A generalizable biomolecular force field.

- carbohydrates. *Journal of Computational Chemistry*, 29(4), 622–655.
- Krishnan, R., Binkley, J. S., Seeger, R., & Pople, J. A. (1980). Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *The Journal of Chemical Physics*, 72(1), 650–654.
- Le Bail, P., Rondeau, C., & Buléon, A. (2005). Structural investigation of amylose complexes with small ligands: Helical conformation, crystalline structure and thermostability. *International Journal of Biological Macromolecules*, 35(1–2), 1–7.
- Lee, C., Yang, W., & Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2), 785–789.
- López, C. A., De Vries, A. H., & Marrink, S. J. (2012). Amylose folding under the influence of lipids. *Carbohydrate Research*, 364, 1–7.
- McLean, A. D., & Chandler, G. S. (1980). Contracted gaussian basis sets for molecular calculations. I. Second row atoms, Z=11-18. *The Journal of Chemical Physics*, 72(10), 5639–5648.
- Murphy, V. G., Zaslow, B., & French, A. D. (1975). The Structure of V amylose dehydrate : A combined X-Ray and stereochemical approach. *Biopolymers*, 14, 1487–1501.
- Nuessli, J., Putaux, J. L., Le Bail, P., & Buléon, A. (2003). Crystal structure of amylose complexes with small ligands. *International Journal of Biological Macromolecules*, 33(4–5), 227–234.
- Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., ... Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16), 1781–1802.

- Rappenecker, G., & Zugenmaier, P. (1981). Detailed refinement of the crystal structure of Vh-Amylose. *Carbohydrate Polymers*, *89*, 11–19.
- Sarko, A., & Winter, W. T. (1974). Crystal and molecular structure of the amylose-DMSO complex. *Biopolymers*, *13*(7), 1461–1482.
- Sarko, A., & Zugenmaier, P. (2009). Crystal structures of amylose and its derivatives. A review. In A. D. French & K. H. Gardner (Eds.), *Fiber Diffraction Methods - ACS Symposium series* (pp. 459–482).
- Takahashi, Y., Kumano, T., & Nishikawa, S. (2004). Crystal structure of B-amylose. *Macromolecules*, *37*(18), 6827–6832.
- Takeda, Y., Shitaozono, T., & Hizukuri, S. (1990). Structures of sub-fractions of corn amylose. *Carbohydrate Research*, *199*, 207–214.
- Toukach, F. V., & Ananikov, V. P. (2013). Recent advances in computational predictions of NMR parameters for the structure elucidation of carbohydrates: Methods and limitations. *Chemical Society Reviews*, *42*(21), 8376–8415.
- Tusch, M., Krüger, J., & Fels, G. (2011). Structural stability of V-amylose helices in water-DMSO mixtures analyzed by molecular dynamics. *Journal of Chemical Theory and Computation*, *7*(9), 2919–2928.
- Veregin, R. P., Fyfe, C. A., & Marchessault, R. H. (1987). Investigation of the crystalline “V” amylose complexes by high-resolution <sup>13</sup>C CP/MAS NMR spectroscopy. *Macromolecules*, *20*(12), 3007–3012. <https://doi.org/10.1021/ma00178a010>
- Whittam, M. A., Orford, P. D., Ring, S. G., Clark, S. A., Parker, M. L., Cairns, P., & Miles, M. J. (1989). Aqueous dissolution of crystalline and amorphous amylose-alcohol complexes.

*International Journal of Biological Macromolecules*, 11(6), 339–344.

Wolinski, K., Hinton, J. F., & Pulay, P. (1990). Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations. *Journal of the American Chemical Society*, 112(23), 8251–8260.

Zabar, S., Lesmes, U., Katz, I., Shimoni, E., & Bianco-Peled, H. (2009). Studying different dimensions of amylose-long chain fatty acid complexes: Molecular, nano and micro level characteristics. *Food Hydrocolloids*, 23(7), 1918–1925.

### 3. Article n°2

L'article intitulé « Diversity of hydrogen bond network and its impact on NMR parameters of amylose B polymorph. A study using molecular dynamics and DFT calculations within periodic boundary conditions » est présenté sous sa version finale pour soumission dans *Journal of Physical Chemistry A*. Les informations supplémentaires peuvent être trouvées dans l'annexe n°3 de ce manuscrit.

# Diversity of hydrogen bond network and its impact on NMR parameters of amylose B polymorph. A study using molecular dynamics and DFT calculations within periodic boundary conditions.

*Adrien Schahl<sup>a,b</sup>, Iann C. Gerber<sup>a</sup>, Valérie Réat<sup>b</sup> and Franck Jolibois<sup>a\*</sup>*

(a) A.Schahl, I.C.Gerber and F.Jolibois, LPCNO, CNRS UMR 5215, Université de Toulouse-INSA-UPS, 135 av. Rangueil, F-31077, Toulouse, France,  
franck.jolibois@univ-tlse3.fr

(b) A.Schahl and V.Réat Institut de Pharmacologie et Biologie Structurale, UMR 5089, CNRS–Université de Toulouse-UPS BP 64182, 205 route de Narbonne, 31077 Toulouse, Cedex 04, France

KEYWORDS Starch • Amylose • DFT • Molecular dynamics • Periodic Boundary Conditions

## **ABSTRACT.**

Classical molecular dynamics simulations have been combined to quantum (DFT) calculations of <sup>13</sup>C NMR parameters in order to relate the experimental spectrum of the double-helix form

of the amylose B-polymorph in highly crystalline conditions not only to its 3D structure but also to the arrangement of atoms in the crystal lattice. Structures obtained from these simulations or from geometry optimizations procedure at the DFT level have shown the presence of hydrogen bond networks between sugars of the same helix or between residues of the two chains of the double helix.  $^{13}\text{C}$  NMR parameter calculations have revealed the impact of such network on the chemical shifts of carbon atoms. In addition, DFT calculations using periodic boundary conditions were compulsory to highlight the presence of two types of sugar within the crystal sample. It allows us to confirm theoretically, the experimental hypothesis that the existence of two distinct sugar types in the NMR spectrum is a consequence of crystal packing.

## **INTRODUCTION.**

Amylose, one of the major constituents of starch with amylopectin, is an essentially linear polymer composed of glucose residues linked in  $[\alpha\text{-}1\rightarrow 4]$  and is often used to model the crystalline parts of starch. Most of the experimental structures are originated from X-ray studies<sup>1-4</sup> and it is well known that it can adopt different helical conformations, dependent of its environment. The single helical structure called the V-polymorph, has been proven to encapsulate guest molecules of various sizes, such as Iodine<sup>5</sup>, aroma molecules<sup>6,7</sup> or even fatty acids<sup>8-11</sup>. The double helical structures are called A or B-polymorphs depending on the crystal packing conditions and are known to be present in high crystalline material amongst starch. The method of choice used for the study of amylose structure is the  $^{13}\text{C}$  solid-state NMR<sup>9,10,12-14</sup>, allowing to differentiate between these polymorphs. It is especially known that chemical shifts of the  $\text{C}_1$  and the  $\text{C}_4$ , are the most affected by the dihedral angles of the glycosidic linkage. At the same time, chemical shift of the other carbon atoms ( $\text{C}_2$ ,  $\text{C}_3$ ,  $\text{C}_5$  and  $\text{C}_6$ ) are mainly affected by the presence and the nature of hydrogen bonds involving the hydroxyl groups and by the associated dihedral angles<sup>15,16</sup>.

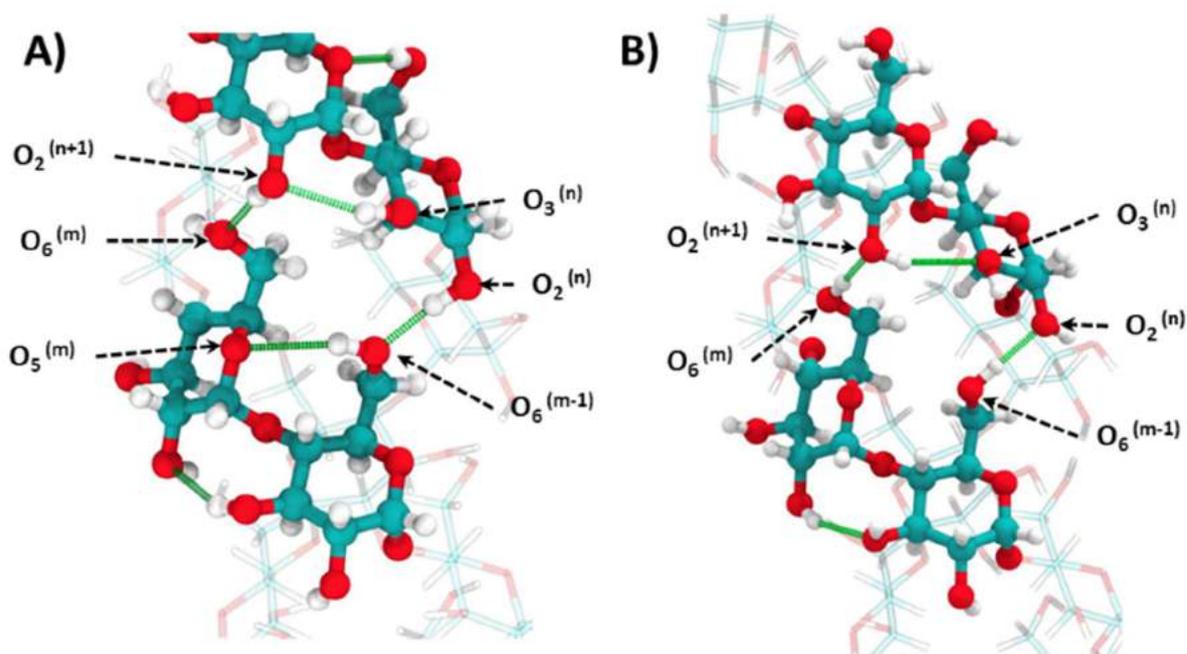
If the general spectrum of the B-polymorph has been known for a long time, it is only since 2006 that the precise NMR spectrum assignment has been performed on a sample uniformly  $^{13}\text{C}$  labelled and presenting a high crystallinity<sup>17</sup>. Particularly, it reveals the presence of two different signals for the C<sub>1</sub>, C<sub>2</sub>, C<sub>4</sub> and C<sub>5</sub> carbons that witness the presence of two different glucose residues. This observation of two different monomers characteristic of this amylose polymorph was related to the crystal packing conditions. However, the link between the molecular structure and the environment leading to two different NMR signatures remains unknown.

Theoretical studies of the structures of amylose have been undertaken but they are mainly based on the use of classical molecular dynamics (MD). The development of force fields<sup>18–20</sup> dedicated to carbohydrates has greatly facilitated such studies. They have helped to study changes in the conformation of amylose in different environments<sup>21,22</sup>, as well as the complexation process occurring between fatty acids and amylose polymers of different degree of polymerization (DP)<sup>23–26</sup>. Computational studies using quantum chemical methods for the calculations of NMR parameters have mainly been focused on relatively small systems<sup>27</sup>. It is only recently that larger systems have been studied, such as cellulose clusters<sup>28,29</sup> or medium size amylose-lipids complexes<sup>30</sup>. In the past decades, *ab initio* or DFT based quantum chemical approaches using Gauge Including Atomic Orbital (GIAO)<sup>31</sup> method have been widely used for the calculations of  $^{13}\text{C}$  chemical shielding tensors. In the context of polysaccharides, these methods are especially useful as they allow precise reproduction of experimental  $^{13}\text{C}$  chemical shifts and correlation between 3D structural changes and evolution of chemical shielding tensors<sup>27,30</sup>. However, if this method behaves successfully in reproducing NMR parameters of molecular systems, important effects due to packing conditions are missing if crystalline systems are considered. A common way to circumvent this problem is to take into account crystal effects using DFT calculations within periodic boundary conditions (PBC). For the

special case of carbon chemical shielding tensor calculations, the Gauge Including Projected Augmented Wave (GIPAW)<sup>32,33</sup> is the method of choice that has already proved its robustness either for small molecules<sup>34,35</sup>, or for bulk calculations<sup>36</sup>. To the best of our knowledge, this method has never been used for the calculation of NMR parameters of polysaccharides composed systems.

Here we present a study that relates <sup>13</sup>C chemical shifts of the B-polymorph of amylose to its 3d structure and its major interactions. We used a combination of DFT and classical molecular dynamics calculations to generate structures for the theoretical computation of NMR parameters. We show that the use of molecular methods is not sufficient to capture the main features of the experimental <sup>13</sup>C chemical shifts, thus confirming that this effect is not due to structural differences inside the complex. However, this computational study highlights the impact of the hydrogen bond network (HBN) on the calculations of NMR parameters of such systems. As we refer to experimental data obtained on a crystalline material, we have intensively used the GIPAW methodology. It allowed us to differentiate between two different types of residue placed alternately along each amylose chain and to link their position in the crystal to the two distinct glucose NMR signals characterized experimentally.

## **COMPUTATIONAL DETAILS.**



**Figure 1.** A) Hydrogen bond network 1 (HBN\_1) B) Hydrogen bond network 2 (HBN\_2). The two structures presented are resulting from geometry optimizations of all the atoms positions.

**Classical molecular dynamics.** Molecular dynamics simulations were performed using the NAMD software<sup>37</sup>. The initial structure of the double chain of amylose has been constructed using the leap program of the AMBER suite<sup>38</sup>. Two chains of 20 residues have been generated with phi ( $\varphi = \text{O5-C1-O4'-C4'}$ ) and psi ( $\psi = \text{C1-O4'-C4'-C3'}$ ) dihedral angles set to  $91.2^\circ \pm 3.5^\circ$  and  $91.8^\circ \pm 3.2^\circ$ , respectively. One of these chains was then rotated by  $180^\circ$  around its principal axis of inertia using the Avogadro software<sup>39</sup> in order to obtain a double helix close to the B-polymorph X-ray crystallographic structure. The procedure for solvation and molecular dynamics was the same as recently described<sup>30</sup>. Amylose was solvated in a box of water such that it was placed at 20 Å from the box boundaries. A modified version of GLYCAM06 force field<sup>30</sup> and the TIP3P model were used for the description of the amylose and the water molecules respectively. The production run was carried out during 55 ns in the canonical ensemble (NVT) with atomic coordinates frame saved every picosecond.

**Quantum chemical calculations – Molecular approach.** All calculations were performed using Gaussian09 suite of programs<sup>40</sup>. Geometry optimisations and spectroscopic properties

calculations have been performed at the DFT level using B3LYP<sup>41,42</sup> hybrid functional associated to the 6-31G(d,p) basis set<sup>43,44</sup>. For the computation of isotropic chemical shielding, the GIAO method<sup>45</sup> has been employed for the numerous advantages it offers. Calculations of NMR parameters were carried out considering a molecular approach. In a first case, <sup>13</sup>C chemical shielding calculations were performed using a full quantum chemical computational strategy i.e. NMR calculations on structures obtained by geometry optimisations. First, full geometry optimisation (optimisation of all atomic positions) has been performed. In this special case, implicit solvent effects (water solvent) were taken into account using the Integral Equation Formalism variant of the Polarizable Continuum Model (IEFPCM<sup>46</sup>). This was compulsory to maintain the 4 terminal residues in a conformation close to the B-polymorph. The starting geometry was obtained from the “Centre de Recherche sur les Macromolécules Végétales (CERMAV)” website<sup>47,48</sup>. One double helix was extracted from the crystallographic data set. It is made of two chains of six glucose residues. Hydrogens of hydroxyl groups have been added with the automatic tool of the Avogadro software<sup>39</sup>. It is well known that the nature and orientation of hydrogen bonds have an impact over calculations of NMR parameters in the case of carbohydrate systems<sup>15</sup>. Consequently, two starting structures with different hydrogen bond network were considered for geometry optimisations. One (noted HBN\_1) is formed by hydrogen bonds between H<sub>O3</sub><sup>(n)</sup> and O<sub>2</sub><sup>(n+1)</sup>, H<sub>O2</sub><sup>(n+1)</sup> and O<sub>6</sub><sup>(m)</sup>, H<sub>O6</sub><sup>(m)</sup> and O<sub>5</sub><sup>(m+1)</sup>, with n being the numbering of considered residues of one chain and m being the numbering of residues composing the other chain (figure 1: A). The other (noted HBN\_2) is formed by hydrogen bonds between H<sub>O6</sub><sup>(m)</sup> and O<sub>2</sub><sup>(n+1)</sup> and H<sub>O2</sub><sup>(n+1)</sup> and O<sub>3</sub><sup>(n)</sup> (Figure1: B). In addition to full geometry optimisation procedure, constrained geometry optimisations have also been undertaken to keep the crystallographic position of heavy atoms, as the NMR experimental data refer to a crystalline material. Under these conditions, only the positions of the hydrogen atoms have been optimized for the two different hydrogen bond networks.

In a second stage, because it is known that molecular dynamics have an impact on structures and spectroscopic parameters<sup>30</sup>, calculations of NMR parameters have been performed on 100 structures extracted from the molecular dynamics trajectory. To avoid border effects, 24 residues over 40 have been selected in the middle part of B-amylose in order to obtain two chains of 12 residues folded in a double helix fashion. The 4 terminal residues have been completed with hydrogen atoms linked either to the O<sub>4</sub> or the O<sub>1</sub> oxygen atoms.

Chemical shieldings of carbon atoms have been compared with experimental data and linear fit performed to correct first theoretical values by mean of the slope and the intercept. The slope of this linear correlation allows to correct our data from weakness of the theoretical model (mainly absence of environment effects) and the intercept corresponds to the reference shielding used to transform chemical shielding into chemical shift.

***Quantum chemical calculations - periodic boundary conditions.*** Periodic DFT calculations were carried out using the Vienna ab initio simulation package VASP<sup>49–52</sup>. The code uses the full-potential projector augmented wave (PAW) framework<sup>53,54</sup>. Exchange-correlation effects have been approximated using the PBE functional<sup>55</sup> and completed by Grimme D3 scheme<sup>56</sup> to account for van der Waals forces. A kinetic-energy cut-off of 500 eV was found to be sufficient to achieve a total-energy convergence within several meV, considering a k-point sampling of (2x2x3) for the structure's optimization, in conjunction with a gaussian smearing with a width of 0.05 eV. Force convergence criteria have been set to 0.01 eV/Å on each relaxed atom. Periodic calculations of nuclear shielding have been performed within the GIPAW approach<sup>32,33</sup> as implemented in the VASP package<sup>57</sup>, using a (1x1x1) k-point grid for the integration of the Brillouin zone, since the use of denser grids has a very limited impact on the result's accuracy. For these calculations, the molecular system consisted in a 3d cell constituted with two left-handed, parallel-stranded double helices. The cell parameters (hexagonal space group with a=b=18.5 Å and c=10.4 Å) and atomic coordinates (252 atoms) were obtained and generated

according to the X-Ray crystallographic structure<sup>48</sup>. As in the previous molecular case, theoretical shieldings were transposed into chemical shifts using the parameters of the linear correlation between theoretical and experimental data.

## RESULTS AND DISCUSSION.

**Standard molecular approach.** In a first attempt to reproduce carbon experimental chemical shifts of the B-type double helix structure of amylose, calculations have been performed using a standard approach i.e. spectroscopic parameters calculated on optimized geometry of isolated molecular system (see computational details). To account for dynamical effects on spectroscopic parameters, carbon chemical shielding averaged over a set of structures extracted from molecular dynamics have also been considered.

All values obtained for each selected dihedral angle are displayed in tables S1 - S3 for geometry optimisations and in table S4 and S5 for MD simulations. From a structural point of view, the average value of selected dihedral angles was calculated along the glucose chains and, when necessary, along the MD simulation. This was done to identify the hydrogen bond networks and the structural features of amylose chains in B-polymorph.

Regarding the angles that characterise the anomeric bond (Table 1), the optimised  $\varphi$  angles are on average  $10^\circ$  larger than in the crystallographic structure, while the corresponding angles  $\psi$  are almost not affected. In the case of MD simulation, even if these two angles are close to the crystallographic ones,  $\varphi$  angles exhibits a broader dynamic. According to this, it appears that the nature of the hydrogen bond network has little influence on the values of these glycosidic dihedral angles. The same conclusion can be reached for the dihedral angles around the C<sub>5</sub>-C<sub>6</sub> bonds (Table 1), all glucose residues mainly adopting a gg conformation as already observed experimentally either on crystal structure<sup>58</sup> or by mean of <sup>1</sup>H NMR spectroscopy<sup>59</sup>.

	Opt <sub>HBN_1</sub>	Opt <sub>HBN_2</sub>	Opt <sub>HBN_1</sub> (H)	Opt <sub>HBN_2</sub> (H)
--	----------------------	----------------------	--------------------------	--------------------------

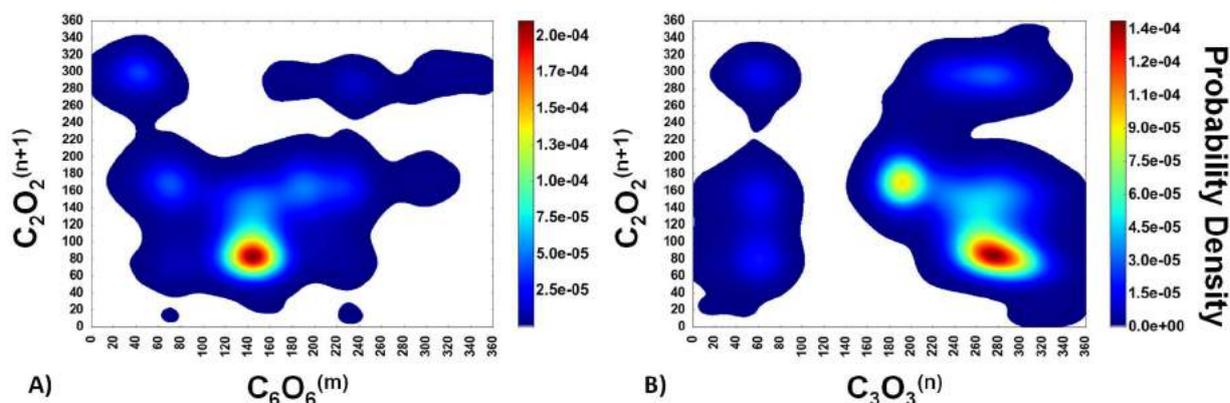
O <sub>5</sub> -C <sub>1</sub> -O <sub>4</sub> '-C <sub>4</sub> '	94.7 ± 4.6	95.3 ± 6.1	84.0 ± 0.1	84.0 ± 0.1
C <sub>1</sub> -O <sub>4</sub> '-C <sub>4</sub> '-C <sub>3</sub> '	97.8 ± 6.1	94.7 ± 4.0	94.2 ± 0.1	94.2 ± 0.1
C <sub>4</sub> -C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub>	48.6 ± 8.7	57.7 ± 3.0	59.6 ± 8.2	59.6 ± 8.2
C <sub>1</sub> -C <sub>2</sub> -O <sub>2</sub> -H <sub>O2</sub>	227.9 ± 25.4	27.3 ± 3.8	278.8 ± 1.7	43.1 ± 5.2
C <sub>2</sub> -C <sub>3</sub> -O <sub>3</sub> -H <sub>O3</sub>	188.7 ± 9.3	307.7 ± 3,6	202.3 ± 12.2	300.1 ± 7.7
C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub> -H <sub>O6</sub>	284.9 ± 3.9	154.3 ± 8.7	284.3 ± 2.3	140.2 ± 4.8

**Table 1.** Selected averaged dihedral angles (degrees) of the full optimisations resulting in the HBN\_1 (Opt<sub>H\_BN1</sub>) and in the HBN\_2 (Opt<sub>H\_BN2</sub>) structures and of the optimisations of the hydrogen positions only resulting in the Opt<sub>H\_BN1</sub>(H) and in the HBN\_2 Opt<sub>H\_BN2</sub>(H) structure. O<sub>5</sub>-C<sub>1</sub>-O<sub>4</sub>'-C<sub>4</sub>' and C<sub>1</sub>-O<sub>4</sub>'-C<sub>4</sub>'-C<sub>3</sub>' are the  $\phi$  and  $\psi$  angles, respectively. See Figure 1 for the definition of the hydrogen bond networks HBN\_1 and 2.

The dihedral angles associated to the hydrogen atoms of the three hydroxyl groups obviously correspond to structural features that can be used to differentiate the two hydrogen bond networks.

Geometry optimizations including all atoms or restrained to hydrogens do not have a drastic impact on the values of the characteristic dihedral angles for each hydrogen bonds network (Table 1). According to the description of each hydrogen bond network (Figure 1), regions of the configuration space can be highlighted. HBN\_1 can be associated to the dihedral angles C<sub>1</sub>-C<sub>2</sub>-O<sub>2</sub>-H<sub>O2</sub>, C<sub>2</sub>-C<sub>3</sub>-O<sub>3</sub>-H<sub>O3</sub> and C<sub>5</sub>-C<sub>6</sub>-O<sub>6</sub>-H<sub>O6</sub> ranging from 200° to 280°, 180° to 215° and 280° to 290°, respectively. For HBN\_2, another region of the configuration space is characterized with dihedral angles ranging from 25° to 50°, 290° to 310° and 135° to 165° for C<sub>1</sub>-C<sub>2</sub>-O<sub>2</sub>-H<sub>O2</sub>, C<sub>2</sub>-C<sub>3</sub>-O<sub>3</sub>-H<sub>O3</sub> and C<sub>5</sub>-C<sub>6</sub>-O<sub>6</sub>-H<sub>O6</sub>, respectively. The average values of the three dihedral angles were obtained while excluding the values of the residues 1, 2, 3 and 1' for the C<sub>1</sub>-C<sub>2</sub>-O<sub>2</sub>-H<sub>O2</sub> and C<sub>2</sub>-C<sub>3</sub>-O<sub>3</sub>-H<sub>O3</sub> angles and 6, 4', 5' and 6' for the C<sub>5</sub>-C<sub>6</sub>-O<sub>6</sub>-H<sub>O6</sub> (see details in Table S4). In our initial model, these particular hydroxyl groups have no residues facing them. Their conformations are therefore not representative of hydrogen bonding networks. For the MD

simulations, the Ramachandran-type plots between the dihedral angles  $C_1-C_2-O_2-H_{O_2}^{(n+1)}$  and  $C_2-C_3-O_3-H_{O_3}^{(n)}$  and the dihedral angles  $C_5-C_6-O_6-H_{O_6}^{(m)}$  and  $C_1-C_2-O_2-H_{O_2}^{(n)}$  (Figure 2) show maximum values densities of  $C_1-C_2-O_2-H_{O_2}^{(n+1)}$ ,  $C_2-C_3-O_3-H_{O_3}^{(n)}$  and  $C_5-C_6-O_6-H_{O_6}^{(m)}$  around  $80^\circ$ ,  $275^\circ$  and  $145^\circ$ , respectively. Despite substantial differences with the static quantum chemical structures (Table 1), these values are closer to the representative angles of the hydrogen bond network HBN\_2. Furthermore, in the Ramachandran plot, the  $C_5-C_6-O_6-H_{O_6}^{(m)}$  dihedral angle values only are representative of the HBN\_2 network (Table 1) attesting the formation of the  $H_{O_6}^{(m)} \cdots O_2^{(n+1)}$  hydrogen bond (Figure 1-B, Figure 2-A). This may indicate that this network is mainly present during the molecular dynamic simulation. The presence of a second maximum on the  $C_1-C_2-O_2-H_{O_2}^{(n+1)}$  and  $C_2-C_3-O_3-H_{O_3}^{(n)}$  Ramachandran map at  $170^\circ$  and  $190^\circ$  may be an indicator, to a lesser extent, of the presence of the HBN\_1 network. Yet, if it attests the formation of  $H_{O_3}^{(n)} \cdots O_2^{(n+1)}$  hydrogen bonds (Figure 1-B), the location of the  $C_1-C_2-O_2-H_{O_2}^{(n+1)}$  at  $170^\circ$  prevents the formation of  $H_{O_2}^{(n+1)} \cdots O_6^{(m)}$  hydrogen bond which is a characteristic of HBN\_1. The presence of a non-negligible population of  $C_2-C_3-O_3-H_{O_3}^{(n)}$  located around  $190^\circ$  (Figure 2-B) associated to  $C_1-C_2-O_2-H_{O_2}^{(n+1)}$  located around  $170^\circ$  proves, however, the ability of the system to start forming the HBN\_1 network according to our simulation conditions (Table 1).



**Figure 2.** A) Ramachandran map of the  $C_2O_2^{(n+1)}$  ( $C_1-C_2-O_2-H_{O2}$  of residue  $n+1$ ) and  $C_6O_6^{(m)}$  ( $C_5-C_6-O_6-H_{O6}$  of residue  $m$ ) dihedral angles (degrees). B) Ramachandran map of the  $C_3O_3^{(n)}$  ( $C_2-C_3-O_3-H_{O3}$  of residue  $n$ ) and  $C_2O_2^{(n+1)}$  dihedral angles (degrees).

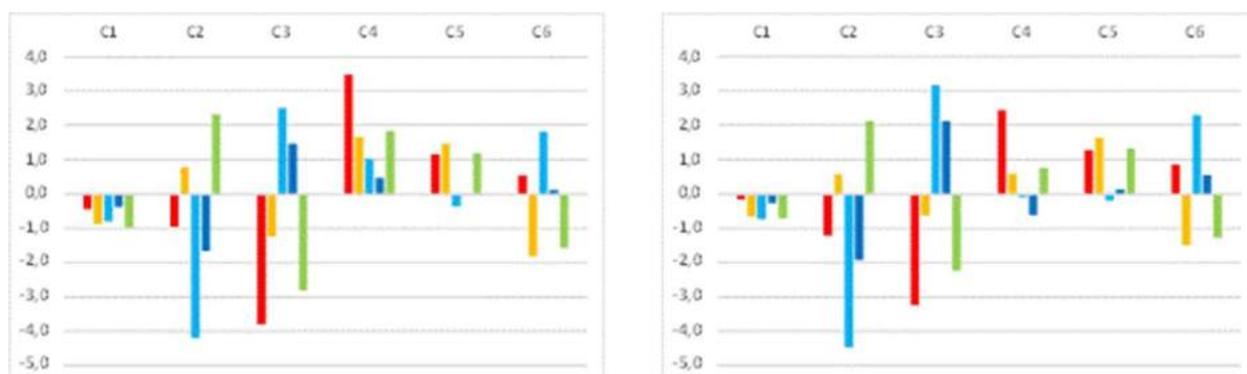
Based on this analysis of conformational properties, molecular dynamics simulations of isolated B-type double helix of amylose indicate that even if one hydrogen bond network (HBN\_2) is mainly present, the second might also exist. The existence of these two hydrogen bond networks might be responsible of the two distinct sets of chemical shifts observed by NMR experiments and that indicate the existence of two types of sugars within the polymer.

Using all the previously described molecular structures,  $^{13}C$  chemical shift calculations were carried out (theoretical chemical shielding given in Table S6 and S7 for optimised HBN\_1 and 2 structures) and compared to experimental chemical shifts (Table S8). Correlation between theoretical and experimental data has failed to distinguish two sugars, as expected from experiment. Indeed, according to the distribution of the chemical shift values for each carbon (see Table S6 and S7), no specific pattern has appeared that can allow to highlight two different sugars. As a consequence, the average chemical shieldings have been calculated for each carbon and compared to experiment. For  $C_2$ ,  $C_3$ ,  $C_5$  and  $C_6$ , all residues have been considered for the average estimates while, in the case of  $C_1$  and  $C_4$ , carbons attached to a hydroxyl group (sugars at the extremities of each chain) have not been taken into account. This set of 6 mean values of carbon chemical shieldings was finally compared to the carbon chemical shifts characterizing one of the 2 types of glucose (noted Sugar 1 and 2) identified experimentally. Linear correlations between theoretical and experimental data have been performed and chemical shifts have been determined using the slope and the intercept of each linear fit (data reported in Table 2 and S8). Finally, the root mean square differences (RMSD) have been calculated in order to give another criterion of computational accuracy. For each correlation, results ( $R^2$  and

RMSD) are given in Table 2. In addition, the relative errors compare to experiment are displayed in Figure 3.

	Sugar 1	Opt <sub>HBN_1</sub>	Opt <sub>HBN_2</sub>	Opt <sub>HBN_1</sub> (H)	Opt <sub>HBN_2</sub> (H)	Mol. Dyn.
C1	<b>99.3</b>	98.9 (± 1.2)	98.4 (± 0.7)	98.5 (± 0.7)	98.9 (± 0.9)	98.3(±0.4)
C2	<b>71.3</b>	70.4 (± 0.6)	72.1 (± 0.6)	67.1 (± 0.3)	69.6 (± 0.3)	73.6(±0.5)
C3	<b>75.1</b>	71.3 (± 1.2)	73.9 (± 1.6)	77.6 (± 0.8)	76.6 (± 1.1)	72.3(±0.3)
C4	<b>72.6</b>	76.1 (± 1.8)	74.3 (± 1.1)	73.6 (± 0.3)	73.1 (± 0.6)	74.4(±0.7)
C5	<b>69.8</b>	70.9 (± 1.1)	71.3 (± 1.1)	69.4 (± 1.2)	69.8 (± 1.2)	71.0(±0.5)
C6	<b>61.3</b>	61.8 (±1.2)	59.5 (± 0.7)	63.1 (± 1.3)	61.4 (± 1.3)	59.7(±0.5)
R <sup>2</sup>		0.965	0.987	0.965	0.994	0.974
RMSD		2.2	1.4	2.2	0.9	1.9
	Sugar 2	Opt <sub>HBN_1</sub>	Opt <sub>HBN_2</sub>	Opt <sub>HBN_1</sub>	Opt <sub>HBN_2</sub>	Mol. Dyn.
C1	<b>100.2</b>	100.1 (± 1.2)	99.5 (± 0.7)	99.5 (± 0.7)	99.9 (± 0.9)	99.5 (± 0.4)
C2	<b>72.1</b>	70.9 (± 0.6)	72.7 (± 0.6)	67.6 (± 0.3)	70.2 (± 0.3)	74.2 (± 0.5)
C3	<b>75.1</b>	71.9 (± 1.2)	74.5 (± 1.6)	78.3 (± 0.8)	77.2 (± 1.1)	72.9 (± 0.3)
C4	<b>74.3</b>	76.7 (± 1.8)	74.9 (± 1.1)	74.2 (± 0.7)	73.7 (± 0.6)	75.1 (± 0.7)
C5	<b>70.2</b>	71.5 (± 1.1)	71.8 (± 1.1)	70.0 (± 1.2)	70.3 (± 1.2)	71.5 (± 0.5)
C6	<b>61.3</b>	62.2 (± 1.2)	59.8 (± 0.7)	63.6 (± 1.7)	61.8 (± 1.3)	60.0 (± 0.5)
R <sup>2</sup>		0.976	0.993	0.958	0.990	0.984
RMSD		1.8	1.0	2.4	1.2	1.5

**Table 2.** Corrected values of averaged theoretical chemical shifts (ppm). Correlation coefficient and RMSD (ppm) between the corrected averaged theoretical chemical shifts and the experimental chemical shifts of Sugar 1 and Sugar 2. Opt<sub>HBN\_1</sub>, Opt<sub>HBN\_2</sub> or Opt<sub>HBN\_1</sub> (H only), Opt<sub>HBN\_2</sub> (H only) refer to geometry optimisations on all atoms or on only hydrogen atoms for the HBN\_1 and HBN\_2 structures, respectively. Mol. Dyn. refers to calculation performed on structures extracted from molecular dynamics simulations. Errors (standard deviations) are given in parenthesis.



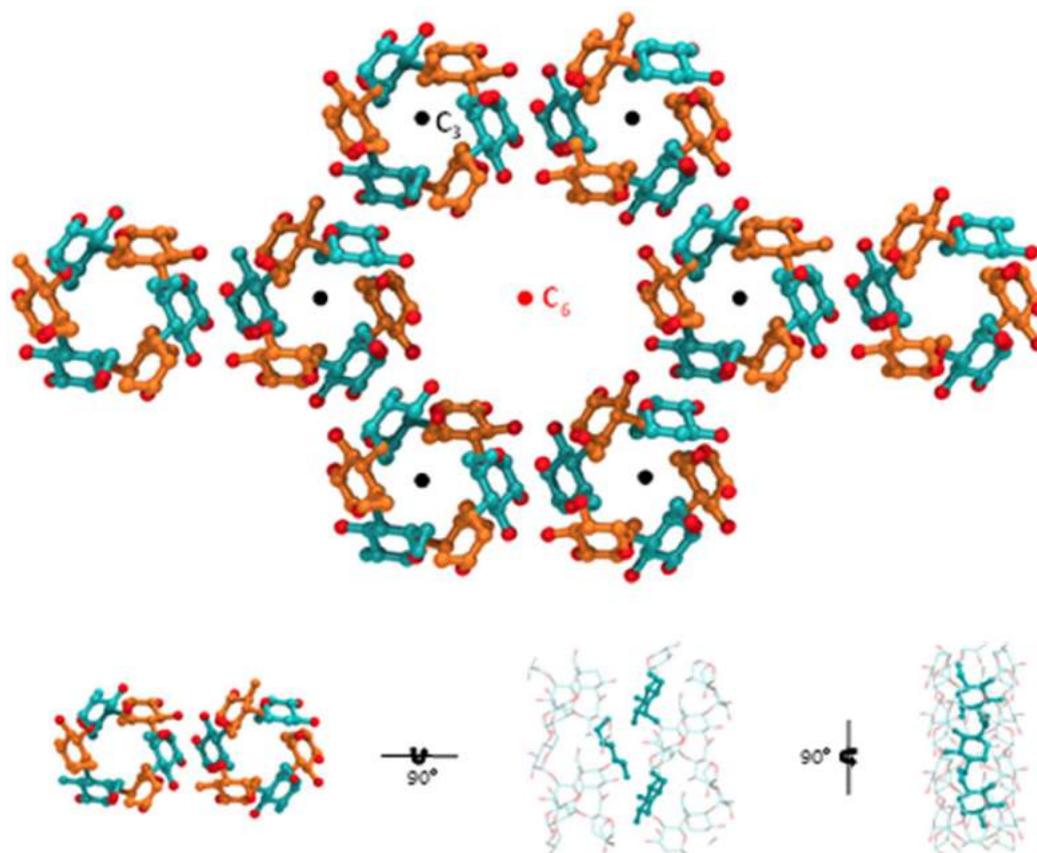
**Figure 3.** Relative errors (ppm) for each carbon atom of the glucose unit in comparison with Sugar 1 (left) and Sugar 2 (right). Red: Opt<sub>HBN\_1</sub>. Yellow: Opt<sub>HBN\_2</sub>. Light blue: Opt<sub>HBN\_1</sub> (H only). Dark blue: Opt<sub>HBN\_2</sub>(H only). Green: Mol. Dyn. Opt<sub>HBN\_1</sub> , Opt<sub>HBN\_2</sub> or Opt<sub>HBN\_1</sub> (H only), Opt<sub>HBN\_2</sub> (H only) refer to geometry optimisations on all atoms or on only hydrogen atoms for the HBN\_1 and HBN\_2 structures, respectively. Mol. Dyn. refers to calculation performed on structures extracted from molecular dynamics simulations.

Among all calculations, the HBN\_2 structure appears to be the best if correlation coefficient and RMSD are considered (Figure 3 – yellow bars). It has been shown recently that the addition of dynamical effects by mean of molecular dynamic simulations can improve the calculation of NMR spectra<sup>30</sup>. Therefore, this approach was applied to a B-polymorph model of 2 x 20 residues (see computational details). Adding dynamical effects was not sufficient to allow the differentiation of two specific sugars (Table S1). As previously done, linear fits of the average values were performed with either Sugar 1 or Sugar 2 experimental values (Table 2 – “Mol. Dyn.” column and Figure 3 - Green bar). Despite a correct correlation with the Sugar 1 and Sugar 2 experimental values ( $R^2 = 0.974$  and  $0.984$ , RMSD= 1.9 and 1.5 ppm, respectively), introduction of dynamical effects did not allow to improve the previous NMR parameters calculated on optimized geometries. Because the relative atomic positions may have a significant impact on the value of chemical shifts and may allow to discriminate two types of sugars within the B-type amylose, calculation of NMR parameters has been considered on the geometry obtained experimentally. In this case, solely the positions of hydrogen atoms have been optimised because their position is neither well defined experimentally nor defined at all in the case of hydroxyl groups. As previously observed, keeping the heavy atoms frozen to their crystallographic positions does not allow to differentiate two types of sugar regarding <sup>13</sup>C chemical shifts (Table S7). However, while the structure associated to the HBN\_1 network presents a poor correlation with the experimental values (Table 2 “Opt<sub>HBN\_1</sub>(H)”,  $R^2=0.965$  and

0.958, RMSD=2.2 and 2.4 ppm for Sugar 1 and Sugar 2, respectively), the other one, associated to the HBN\_2 network gives best results (Table 2 “Opt<sub>HBN\_2</sub>(H)”,  $R^2=0.994$  and  $0.990$ , RMSD=0.9 and 1.2 ppm for Sugar 1 and Sugar 2 respectively). Even if this last structure seems to be in better agreement with experiment than the other one, it doesn't manage to reproduce correctly all individual carbon chemical shifts (in particular for C<sub>2</sub> and C<sub>3</sub> with relative errors between 1.5 and 2 ppm in absolute value). Furthermore, the fact that the correlation is good with the two types of sugar remains also questionable.

At this stage of our study, two conclusions can be underlined: 1) Despite a good overall correlation between calculated values and experimental ones, molecular methods are not sufficient to clearly differentiate clearly two types of sugar in such systems. Thus, this difference may not be explained neither by a dynamical effect nor an intrinsic structural effect. 2) The hydrogen bond network inside one double helix has a non-negligible impact on the calculation of NMR parameters. Here, molecular methods suggest that the HBN\_2 structure is responsible for the experimental chemical shift as far as there is no packing effect. Thus, the packing effect observed experimentally has to be somehow considered in the NMR calculation to allow this differentiation.

***Periodic boundary conditions approach.*** Experimentally, it has been postulated that the presence of two NMR signals for glucose residues is due to a packing effect in the B-polymorph amylose crystal<sup>17,48</sup>. The complete attribution, realised by Rondeau-Mouro *et al.*, has been done on a highly crystalline material. In order to take this periodicity property into account, the determination of the theoretical NMR parameters was carried out under PBC using the GIPAW method.



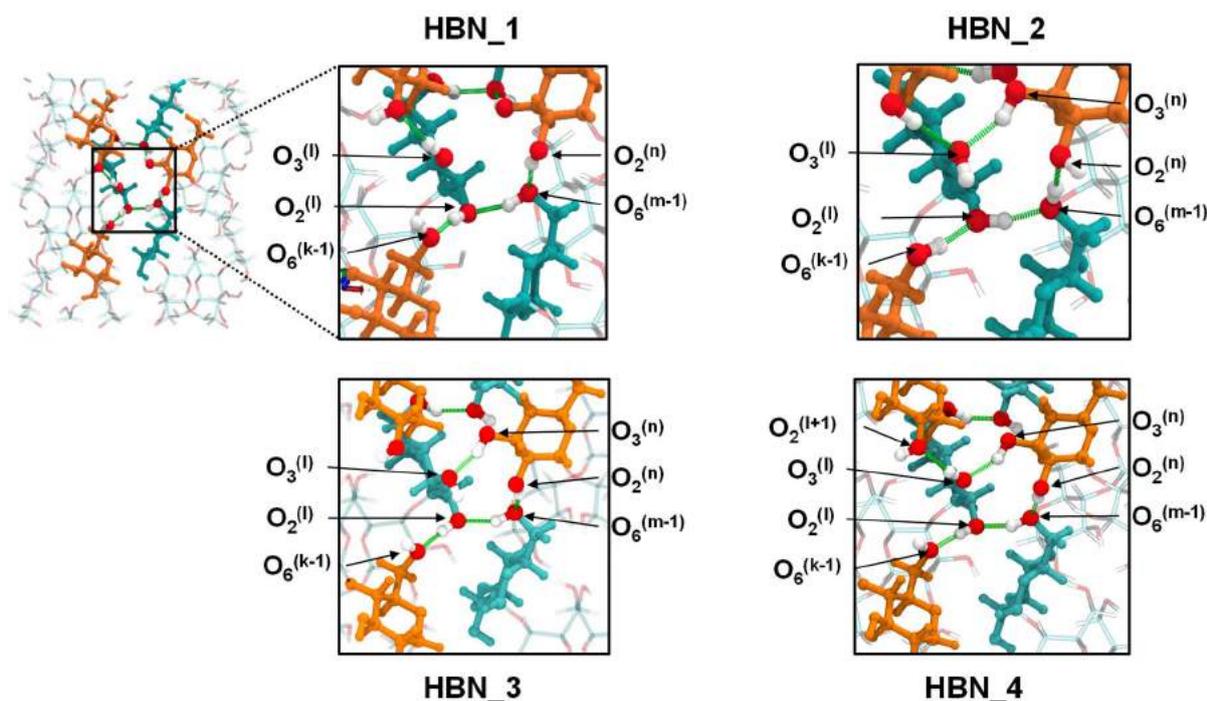
**Figure 4.** Up: Localisation of  $\text{sugar}_{\text{interface}}$  (blue) and  $\text{sugar}_{\text{solvent}}$  (orange) in the extended crystal mesh. In the crystal unit,  $C_3$  is the rotation- translation axis  $3_1$  and the inertial axis of a double-helix,  $C_6$  is the rotation- translation axis  $6_1$  and director axis of the crystal solvent channel.

	HBN_1	HBN_2	HNB_3	HNB_4
$O_5-C_1-O_4'-C_4'$	<b>81.7</b>	90.8	87.4	79.7
	90.6	88.9	82.8	92.9
$C_1-O_4'-C_4'-C_3'$	97.8	96.4	92.8	95.7
	94.5	98.0	97.4	96.6
$C_4-C_5-C_6-O_6$	51.4	54.9	57.3	59.2
	67.9	60.5	62.2	61.1
$C_1-C_2-O_2-H_{02}$	249.8	33.5	281.3	267.7
	247.5	<b>120.1</b>	288.7	265.2
$C_2-C_3-O_3-H_{03}$	192.1	294.9	313.6	<b>189.3</b>

C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub> -H <sub>o6</sub>	177.5	308.1	300.6	292.8
	<b>17.4</b>	129.3	16.9	14.3
	280.5	147.0	290.0	295.1

**Table 3.** Selected averaged dihedral angles (degrees) of the full optimisations, using PBC, resulting in the HBN\_1, HBN\_2, HBN\_3 and HBN\_4 structures (See text for details). O5-C1-O4'-C4' and C1-O4'-C4'-C3' are the  $\phi$  and  $\psi$  angles, respectively. For each angle, the first value corresponds to the sugar<sub>interface</sub>, the second to the sugar<sub>solvent</sub>. In bold, dihedral angle values out of range of values defined by molecular approach.

As in our previous analysis on molecular systems, two types of conformation have been considered to model the two hydrogen bonds networks (HBN\_1 and HBN\_2) previously highlighted. For PBC calculations, geometry optimizations were performed considering the entire set of atoms. In this case, constrained geometry optimizations of hydrogen atoms only have not been considered because PBC induce crystal packing constraints. According to the 3d structure (Figure 4), two pseudo-axis of third (rotation- translation axis 3<sub>1</sub>) and sixth (rotation-translation axis 6<sub>1</sub>) order rotational symmetry determine the repartition of each helix within the crystal. One double helix has three close neighbors distributed according to the third order rotational axis. Thus, two populations of residues can be differentiated along amylose chains regarding their positions inside the crystal (Figure 4). One type of sugar is located at the interface between two amylose double helices (sugar in blue color in Figure 4 and noted sugar<sub>interface</sub>) while the other corresponds to sugars that point toward the sixth order rotational axis (sugar in orange color in Figure 4 and noted sugar<sub>solvent</sub>). These last sugars are in contact with structural water molecules that have been characterized in the solvent channel within the crystal structure<sup>48</sup>. For our calculations, these water molecules have not been taken into account considering the difficulty to unambiguously place their hydrogen atoms. The principal dihedral angles for both structures and optimization strategies are presented in Table 3.



**Figure 5.** Details of HBN\_1, HBN\_2, HBN\_3 and HBN\_4 structures resulting from full geometry optimisations in periodical conditions.

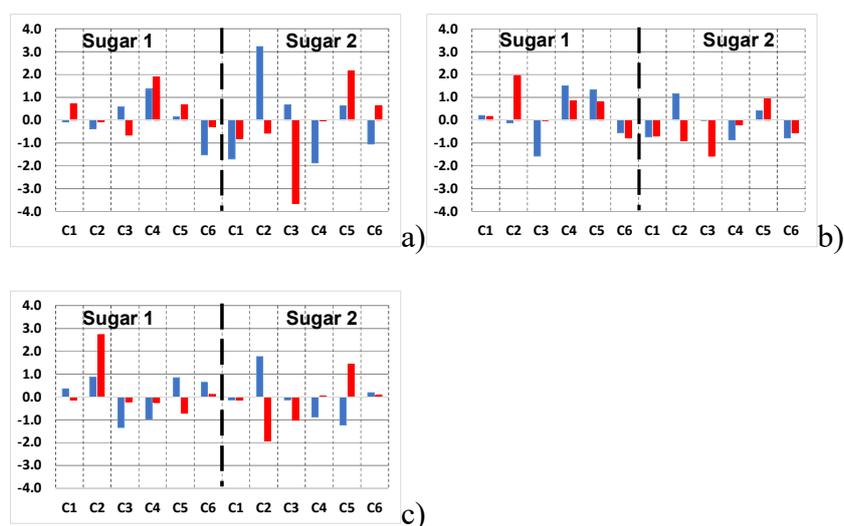
As for the molecular approach, the nature of the hydrogen bond network has little influence on the values of the dihedral angles that defined the conformation around the C<sub>5</sub>-C<sub>6</sub> bonds (mainly *gg* type). For the glycosidic angles  $\varphi$  and  $\psi$ , a similar conclusion is reached, even if a difference of 10° is observed for the  $\varphi$  angle of the residue at the interface when one goes from HBN\_1 to HBN\_2 (81.7° to 90.8°, respectively). When others structural parameters are considered, two sets of values are obtained for all dihedral angles reflecting this differentiation between residues. Moreover, if compared with the previous hydrogen bond networks (molecular approach-HBN\_1, characterised by C<sub>1</sub>-C<sub>2</sub>-O<sub>2</sub>-H<sub>O2</sub>, C<sub>2</sub>-C<sub>3</sub>-O<sub>3</sub>-H<sub>O3</sub>, C<sub>5</sub>-C<sub>6</sub>-O<sub>6</sub>-H<sub>O6</sub> ranging from 200° to 280°, 180° to 215° and 280° to 290°; molecular approach-HBN\_2, characterised by C<sub>1</sub>-C<sub>2</sub>-O<sub>2</sub>-H<sub>O2</sub>, C<sub>2</sub>-C<sub>3</sub>-O<sub>3</sub>-H<sub>O3</sub>, C<sub>5</sub>-C<sub>6</sub>-O<sub>6</sub>-H<sub>O6</sub> ranging from 25° to 50°, 290° to 310° and 135° to 165°), while there are similarities for dihedral angles, the ones resulting from PBC calculations present differences (Figure 5 and Table 3) due to interactions between helices.

In the case of HBN\_1, difference between the two types of sugar is observed for the dihedral angle around the C<sub>6</sub>-O<sub>6</sub> bonds, while for HBN\_2, this difference is observed for the conformation around the C<sub>2</sub>-O<sub>2</sub> bond. In both cases, the difference in dihedral angle values is around 90 degrees. This can be related to the formation of slightly different hydrogen bond networks compare to molecular approach. The HBN\_1 is now formed by hydrogen bonds between H<sub>O</sub><sub>2</sub><sup>(n)</sup> and O<sub>6</sub><sup>(m-1)</sup>, H<sub>O</sub><sub>6</sub><sup>(m-1)</sup> and O<sub>2</sub><sup>(l)</sup>, H<sub>O</sub><sub>2</sub><sup>(l)</sup> and O<sub>6</sub><sup>(k-1)</sup>, with m and n referring to residues of different chains situated in the same double helix, k and l referring to residues of different chains situated in another double helix (Figure 5 HBN\_1). The hydrogen bonds formed between H<sub>O</sub><sub>3</sub><sup>(n or l)</sup> and O<sub>2</sub><sup>(n+1 or l+1)</sup> are conserved along the helix with no variation dependant on the position of the considered residue. The HBN\_2 is formed by hydrogen bonds between H<sub>O</sub><sub>6</sub><sup>(k-1)</sup> and O<sub>2</sub><sup>(l)</sup>, H<sub>O</sub><sub>2</sub><sup>(l)</sup> and O<sub>6</sub><sup>(m-1)</sup>, H<sub>O</sub><sub>6</sub><sup>(m-1)</sup> and O<sub>2</sub><sup>(n)</sup>, H<sub>O</sub><sub>2</sub><sup>(l+1)</sup> and O<sub>3</sub><sup>(l)</sup>, H<sub>O</sub><sub>3</sub><sup>(n)</sup> and O<sub>3</sub><sup>(l)</sup> (Figure 5 HBN\_2). In the latter case, the O<sub>3</sub> of sugars<sub>solvent</sub> are only hydrogen bond donor (Figure 5 HBN\_2, O<sub>3</sub><sup>(n)</sup>) whereas O<sub>3</sub> of sugars<sub>interface</sub> are acceptors of two hydrogen bonds (Figure 5 HBN\_1, O<sub>3</sub><sup>(k)</sup>). Thus, HO<sub>3</sub> hydroxyl groups can be differentiated in the HBN\_2 from sugar<sub>interface</sub> to sugar<sub>solvent</sub> using their intrinsic number of hydrogen bonds. This differentiation is not possible in the HBN\_1 structure.

		Exp.	HBN_1	HBN_2	Average 1	Average 2	HBN_3	HBN_4
Sugar 1	C1	99.3	99.2	100.0	99.5	99.5	99.7	99.2
	C2	71.3	70.9	71.2	71.2	73.2	72.2	74.0
	C3	75.1	75.7	74.4	73.5	75.1	73.8	74.9
	C4	72.6	73.9	74.5	74.1	73.4	71.6	72.3
	C5	69.8	70.0	70.5	71.1	70.6	70.6	69.1
	C6	61.3	59.8	61.0	60.7	60.5	61.9	61.4
Sugar 2	C1	100.2	98.5	99.4	99.5	99.5	100.1	100.1
	C2	72.1	75.3	71.5	73.2	71.2	73.9	70.2
	C3	75.1	75.8	71.5	75.1	73.5	75.0	74.1
	C4	74.3	72.4	74.3	73.4	74.1	73.4	74.3

	C5	70.2	70.8		72.4		70.6		71.1		69.0		71.6					
	C6	61.3	60.3		61.9		60.5		60.7		61.5		61.4					
<b>R<sup>2</sup></b>	<b>0.986</b>			<b>0.985</b>			<b>0.994</b>			<b>0.994</b>			<b>0.994</b>			<b>0.991</b>		
<b>RMSD</b>	<b>1.4</b>			<b>1.4</b>			<b>0.9</b>			<b>0.9</b>			<b>0.9</b>			<b>1.1</b>		

**Table 4.** Best calculated chemical shifts (ppm) for HBN\_1, HBN\_2 HBN\_3 and HBN\_4 theoretical structures after correction with experimental data using the coefficients of the linear interpolation (experiment vs theory). Chemical shifts (Average 1 and 2) obtained after averaging of  $\text{sugar}_{\text{interface}}$  or  $\text{sugar}_{\text{solvent}}$  chemical shieldings of HBN\_1 and HBN\_2 optimized structures and correction with experimental data using the coefficients of the linear interpolation are also calculated.  $R^2$  and RMSD (in ppm) are given in order to indicate an estimation of errors.



**Figure 6.** Relative errors (ppm) resulting from the best theory/experiment correlations for (a) HBN\_1 (Blue) and HBN\_2 (Red), (b) for average\_1 (Blue) and average\_2 (Red) and (c) HBN\_3 (Blue) and HBN\_4 (Red).

It has been demonstrated in monosaccharide systems that the number and nature of hydrogen bond formed on a  $C_i$  carbon have a direct impact on the calculated chemical shift of carbons  $C_{i+1}$  and  $C_{i-1}$ <sup>15</sup>. As a consequence, chemical shieldings have been calculated for the sugar carbon atoms in the crystal structures using PBC and compared to experiment. First, two populations

of residues were differentiated in terms of chemical shielding (Figure S1) for both HBN\_1 and HBN\_2 geometrically optimized structures. One is associated to sugars located at the interface between double-helices ( $\text{sugar}_{\text{interface}}$ ) while the other corresponds to the other family of sugar facing the solvent channel ( $\text{sugar}_{\text{solvent}}$ ). This observation seems to confirm the experimental hypothesis about the differentiation of two sugars due to crystal packing. The next question is whether these two sets of  $^{13}\text{C}$  chemical shieldings ( $\text{sugar}_{\text{interface}}$  and  $\text{sugar}_{\text{solvent}}$ ) calculated for the HBN structures can be directly associated to the two sets of experimental data (Sugar 1 and Sugar 2). Using the 4 possible combinations ( $[\text{sugar}_{\text{interface}}, \text{sugar}_{\text{solvent}}]$  vs  $[\text{sugar 1}, \text{sugar 2}]$ ) for the geometrical structures HBN\_1 and 2, see Table S9), it was not possible to find a global correlation (i.e. using simultaneously the 12 experimental carbon chemical shifts) between the theoretical and experimental data. The best correlations are displayed in Table 4 and Figure 6 ( $R^2 = 0.985-986$ ; RMSD=1.4 ppm). On the basis of the structures HBN\_1 and HBN\_2 that we have defined, it is not then possible to associate directly the structural position of sugar (facing a double helix or solvent) with one type of sugar identified experimentally. However, if we consider the relative errors associated with the twelve carbon chemical shifts (Figure 6), we observe that the  $\text{sugar}_{\text{interface}}$  with HBN\_1 hydrogen bonds (Figure 6.a, Sugar 1, blue bars) presents lower absolute values ( $< 1.5$  ppm) than when they are involved with HBN\_2 hydrogen bonds (Figure 6.a, Sugar 2, red bars,  $< 3.2$  ppm). On the contrary,  $\text{sugar}_{\text{solvent}}$  presents smaller absolute relative error values in the HBN\_2 structure (Figure 6.a, Sugar 1, red bars) than in HBN\_1 structure (Figure 6.a, Sugar 2, blue bars) with maximum at 1.9 ppm compared to 3.6 ppm, respectively. These results suggest that one homogeneous hydrogen bond orientation replicated along the entire amylose chains does not form a viable model for the calculation of NMR parameters of such complexes. Because fast exchange between the two hydrogen bond networks can be present, dynamical effects must be somehow taken into account. In order to statically model the latter, the averaging of chemical shifts has been considered. Mean values

have been calculated for each carbon by averaging chemical shifts of HBN\_1 and HBN\_2. Comparison with experimental data shows that the new sets of 12 chemical shifts (average\_1 and average\_2 in Table 4) substantially improved our theoretical results with excellent correlation coefficient ( $R^2=0.994$ ) and RMSD ( $< 1$  ppm). In addition, the errors are more evenly distributed over all carbons with values below 2 ppm (Figure 6.b). However, the discrimination between  $\text{sugar}_{\text{interface}}$  and  $\text{sugar}_{\text{solvent}}$  remains impossible (see for example all  $C_1$  values for average\_1 and 2 in Table 4). This average of HBN\_1 and HBN\_2 can be considered as a static model of a synchronous dynamic of both hydrogen bond networks. i.e. the whole hydrogen bond network moves from HBN\_1 to HBN\_2.

An asynchronous dynamic can be envisaged for which only the hydrogen bond network of one type of sugar is modified (either  $\text{sugar}_{\text{interface}}$  or  $\text{sugar}_{\text{solvent}}$ ). To confirm this hypothesis, we constructed two other hydrogen bond networks: The first, noted HBN\_3, presents HBNs for the  $\text{sugar}_{\text{interface}}$  and  $\text{sugar}_{\text{solvent}}$  close to HBN\_2 and HBN\_1, respectively. The second, noted HBN\_4 on the contrary, presents HBNs for the  $\text{sugar}_{\text{interface}}$  and  $\text{sugar}_{\text{solvent}}$  close to HBN\_1 and HBN\_2, respectively. The two resulting structures were subjected to geometry optimization and chemical shielding were calculated and compared to experiment (see Table 4). The carbonated skeletons exhibit no specific alteration for the two new considered structures if  $\varphi$ ,  $\psi$  and dihedral angles around  $C_5-C_6$  are considered (Table 3). Concerning the hydrogen bond networks, only the dihedral angles around  $C_6-O_6$  can discriminate the two sugars (interfacial and facing to solvent) for HNB\_3 while for HNB\_4, dihedral angles around  $C_6-O_6$  and  $C_3-O_3$  allow to differentiate the two positions. As already described in the previous paragraphs, these reflect difference of hydrogen bonding between helices both in intra and inter double-helix, within the crystal lattice and can be at the origin of different chemical shifts repartition calculated for these structures.

For HBN\_3 and HBN\_4 structures, global correlation of NMR chemical shielding has been envisaged by correlating at the same time  $\text{sugar}_{\text{interface}}$  with sugar 1 and  $\text{sugar}_{\text{solvent}}$  with sugar 2 or the opposite (Table S10). The best correlations are displayed in Table 4. For HBN\_3, the best results are obtained by assigning chemical shifts of  $\text{sugar}_{\text{interface}}$  and  $\text{sugar}_{\text{solvent}}$  to Sugar 1 and 2 respectively, while for HBN\_4 the opposite has been found,  $\text{sugar}_{\text{interface}}$  and  $\text{sugar}_{\text{solvent}}$  being correlated to Sugar 2 and 1, respectively. Even though several little errors are still presents (Figure 6.c. error < 1.8 and 2.7 ppm for HBN\_3 and 4 structures, respectively), these values yield also a very good agreement compared to experiment. The corresponding correlation coefficients (0.994 and 0.991) together with RMSD around 1 ppm calculated using the whole set of chemical shifts are a proof of this correctness. However, for one of these structures, Sugar 1 is attributed to sugars at the interface between double-helices while for the other, the same experimental sugar is attributed to sugars facing the solvent channel.

## CONCLUSIONS

In conclusion, we demonstrate that it is possible to achieve qualitative good agreement between experiment and theoretical  $^{13}\text{C}$  NMR chemical shifts in the case of a highly crystalline sample of the Amylose B-polymorph. To do so, it is necessary to correctly model the chemical environment using calculations at the DFT level within periodic boundary conditions. This was necessary to confirm, for the first time, the experimental hypothesis that the presence of two types of sugar in the NMR spectrum is a consequence of crystal packing.

However, neither of the theoretical structures allow us to discriminate the allocation of NMR spectra with respect to the position of the sugars. It is necessary to remember that no water molecules are included in our calculations, especially those that have been characterised in a specific channel created by the six double-helices distributed around the  $\text{C}_6$  pseudo axis. These water molecules are crucial for the crystal structure of the B-polymorph of amylose<sup>48</sup> and

obviously may play an important role in the evaluation of  $^{13}\text{C}$  NMR chemical shifts. They may create new dynamic hydrogen bonds with the hydroxyl groups of the residues located in this channel and thus influence the value of chemical shielding of carbons  $\text{C}_2$ ,  $\text{C}_3$  and  $\text{C}_5$ . Because, the position of the hydrogen atoms of water is not characterised by X-ray scattering, it is impossible to correctly model these molecules in our computational strategy without resorting to *ab initio* molecular dynamics, which is outside the scope of this study.

## **Supporting**

## **Information.**

The supporting information is available free of charge. Table S1:  $\phi$  and  $\psi$  angles after geometry optimizations and for the crystal structure; Table S2: Selected dihedral angles for full (all atoms) geometry optimizations; Table S3: Selected dihedral angles for constrained (Hydrogen atoms only) geometry optimizations; Table S4:  $\phi$  and  $\psi$  angles averages extracted from MD simulation; Table S5: Selected dihedral angles averages extracted from MD simulation; Table S6: Chemical shielding obtained for full (all atoms) geometry optimizations methods; Table S7: Chemical shielding obtained for constrained (hydrogen atoms only) geometry optimizations method; Table S8: Molecular approach. Experimental chemical shift and average theoretical chemical shielding used for theory vs experience comparison and parameters of the linear fit; Table S9: Periodic Boundary conditions approach. Experimental chemical shift and theoretical chemical shielding used for theory vs experience comparison; Table S10: Periodic Boundary conditions approach. Experimental chemical shift and theoretical chemical shielding used for theory vs experience comparison; Figure S1: Carbon chemical Shielding obtained after GIPAW calculations on HBN\_1 and HBN\_2 optimized structures;

## **Corresponding Author**

\*Franck Jolibois. E-Mail: franck.jolibois@univ-tlse3.fr

### **Author Contributions**

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

### **ACKNOWLEDGMENT**

A.S. was a recipient of a PhD scholarship from University of Toulouse and the “Région Occitanie”. This work was supported by the CNRS-MITI grant “Modélisation du vivant” 2019. This work was granted access to the HPC resources of CALMIP supercomputing centre under the allocation 2018-p0758 and of CINES and IDRIS supercomputing centre under the allocation 2019-A0060810816 made by GENCI.

### **REFERENCES**

(Word Style “TF\_References\_Section”).

- (1) Rappenecker, G.; Zugenmaier, P. Detailed Refinement of the Crystal Structure of Vh-Amylose. *Carbohydr. Polym.* **1981**, *89*, 11–19.
- (2) Winter, W. T.; Sarko, A. Crystal and Molecular Structure of the Amylose-DMSO Complex. *Biopolymers* **1974**, *13* (7), 1461–1482.
- (3) Takahashi, Y.; Kumano, T.; Nishikawa, S. Crystal Structure of B-Amylose. *Macromolecules* **2004**, *37* (18), 6827–6832.
- (4) Murphy, V. G.; Zaslav, B.; French, A. D. The Structure of V Amylose Dehydrate :

- A Combined X-Ray and Stereochemical Approach. *Biopolymers* **1975**, *14*, 1487–1501.
- (5) Immel, S.; Lichtenthaler, F. W. The Hydrophobic Topographies of Amylose and Its Blue Iodine Complex. *Starch - Stärke* **2002**, *52* (1), 1–8.
  - (6) Nuessli, J.; Putaux, J. L.; Le Bail, P.; Buléon, A. Crystal Structure of Amylose Complexes with Small Ligands. *Int. J. Biol. Macromol.* **2003**, *33* (4–5), 227–234.
  - (7) Tapanapunnitikul, O.; Chaiseri, S.; Peterson, D. G.; Thompson, D. B. Water Solubility of Flavor Compounds Influences Formation of Flavor Inclusion Complexes from Dispersed High-Amylose Maize Starch. *J. Agric. Food Chem.* **2008**, *56* (1), 220–226.
  - (8) Eliasson, A. C.; Krog, N. Physical Properties of Amylose-Monoglyceride Complexes. *J. Cereal Sci.* **1985**, *3* (3), 239–248.
  - (9) F. Horii, H. Yamamoto, A. Hirai, R. K. Structural Study of Amylose Polymorphs by Cross-Polarization-Magic-Angle Spinning, <sup>13</sup>C-N.M.R. Spectroscopy. *Carbohydr. Res.* **1987**, *160*, 29–40.
  - (10) Le Bail, P.; Buleon, A.; Shiftan, D.; Marchessault, R. H. Mobility of Lipid in Complexes of Amylose-Fatty Acids by Deuterium and <sup>13</sup>C Solid State NMR. *Carbohydr. Polym.* **2000**, *43* (4), 317–326.
  - (11) Gelders, G. G.; Vanderstukken, T. C.; Goesaert, H.; Delcour, J. A. Amylose-Lipid Complexation: A New Fractionation Method. *Carbohydr. Polym.* **2004**, *56* (4), 447–458.
  - (12) J., G. M. Quantification of the Structural Features of Starch Polysaccharides by

- NMR Spectroscopy. *Carbohydr. Res.* **1985**, *139*, 85–93.
- (13) Snape, C. E.; Morrison, W. R.; Maroto-Valer, M. M.; Karkalas, J.; Pethrick, R. A. Solid State  $^{13}\text{C}$  NMR Investigation of Lipid Ligands in V-Amylose Inclusion Complexes. *Carbohydr. Polym.* **1998**, *36* (2–3), 225–237.
- (14) Duus, J.; Gotfredsen, C. H.; Bock, K. Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations. *Chem. Rev.* **2000**, *100* (12), 4589–4614.
- (15) Suzuki, S.; Horii, F.; Kurosu, H. Theoretical Investigations Of  $^{13}\text{C}$  Chemical Shifts in Glucose, Cellobiose, and Native Cellulose by Quantum Chemistry Calculations. *J. Mol. Struct.* **2009**, *921* (1–3), 219–226.
- (16) Jolibois, F.; Soubias, O.; Réat, V.; Milon, A. Understanding Sterol-Membrane Interactions Part I: Hartree-Fock versus DFT Calculations of  $^{13}\text{C}$  and  $^1\text{H}$  NMR Isotropic Chemical Shifts of Sterols in Solution and Analysis of Hydrogen-Bonding Effects. *Chem. - A Eur. J.* **2004**, *10*, 5996–6004.
- (17) Rondeau-Mouro, C.; Veronese, G.; Buléon, A. High-Resolution Solid-State NMR of B-Type Amylose. *Biomacromolecules* **2006**, *7*, 2455–2460.
- (18) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. *Carbohydrates. J. Comput. Chem.* **2008**, *29* (4), 622–655.
- (19) Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W.; MacKerell, A. D. CHARMM Additive All-Atom

- Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate-Protein Modeling. *J. Chem. Theory Comput.* **2011**, 7 (10), 3162–3180.
- (20) Plazinski, W.; Lonardi, A.; Hünenberger, P. H. Revision of the GROMOS 56A6 CARBO Force Field: Improving the Description of Ring-Conformational Equilibria in Hexopyranose-Based Carbohydrates Chains. *J. Comput. Chem.* **2016**, 37 (3), 354–365.
- (21) Tusch, M.; Krüger, J.; Fels, G. Structural Stability of V-Amylose Helices in Water-DMSO Mixtures Analyzed by Molecular Dynamics. *J. Chem. Theory Comput.* **2011**, 7 (9), 2919–2928.
- (22) Koneru, J. K.; Zhu, X.; Mondal, J. Quantitative Assessment of the Conformational Heterogeneity in Amylose across Force Fields. *J. Chem. Theory Comput.* **2019**, 15 (11), 6203–6212.
- (23) López, C. A.; De Vries, A. H.; Marrink, S. J. Amylose Folding under the Influence of Lipids. *Carbohydr. Res.* **2012**, 364, 1–7.
- (24) Feng, T.; Li, M.; Zhou, J.; Zhuang, H.; Chen, F.; Ye, R.; Campanella, O.; Fang, Z. Application of Molecular Dynamics Simulation in Food Carbohydrate Research - A Review. *Innov. Food Sci. Emerg. Technol.* **2015**, 31, 1–13.
- (25) Cheng, L.; Feng, T.; Zhang, B.; Zhu, X.; Hamaker, B.; Zhang, H.; Campanella, O. A Molecular Dynamics Simulation Study on the Conformational Stability of Amylose-Linoleic Acid Complex in Water. *Carbohydr. Polym.* **2018**, 196, 56–65.
- (26) Cheng, L.; Zhu, X.; Hamaker, B. R.; Zhang, H.; Campanella, O. H. Complexation

- Process of Amylose under Different Concentrations of Linoleic Acid Using Molecular Dynamics Simulation. *Carbohydr. Polym.* **2019**, *216*, 157–166.
- (27) Toukach, F. V.; Ananikov, V. P. Recent Advances in Computational Predictions of NMR Parameters for the Structure Elucidation of Carbohydrates: Methods and Limitations. *Chem. Soc. Rev.* **2013**, *42* (21), 8376–8415.
- (28) Kubicki, J. D.; Mohamed, M. N.-A.; Watts, H. D. Quantum Mechanical Modeling of the Structures, Energetics and Spectral Properties of I $\alpha$  and I $\beta$  Cellulose. *Cellulose* **2013**, *20* (1), 9–23.
- (29) Yang, H.; Wang, T.; Oehme, D.; Petridis, L.; Hong, M.; Kubicki, J. D. Structural Factors Affecting <sup>13</sup>C NMR Chemical Shifts of Cellulose: A Computational Study. *Cellulose* **2018**, *25* (1), 23–36.
- (30) Schahl, A.; Réat, V.; Jolibois, F. Structures and NMR Spectra of Short Amylose-Lipid Complexes. Insight Using Molecular Dynamics and DFT Quantum Chemical Calculations. *Carbohydr. Polym.* **2020**, *235* (January), 115846.
- (31) Wolinski, K.; Hinton, J. F.; Pulay, P. Efficient Implementation of the Gauge-Independent Atomic Orbital Method for NMR Chemical Shift Calculations. *J. Am. Chem. Soc.* **1990**, *112* (23), 8251–8260.
- (32) Pickard, C. J.; Mauri, F. All-Electron Magnetic Response with Pseudopotentials: NMR Chemical Shifts. *Phys. Rev. B* **2001**, *63* (24), 245101–245113.
- (33) Yates, J. R.; Pickard, C. J.; Mauri, F. Calculation of NMR Chemical Shifts for Extended Systems Using Ultrasoft Pseudopotentials. *Phys. Rev. B* **2007**, *76* (2), 024401–024411.

- (34) Marín-Luna, M.; Alkorta, I.; Elguero, J. A Theoretical NMR Study of Selected Benzazoles: Comparison of GIPAW and GIAO-PCM (DMSO) Calculations. *Magn. Reson. Chem.* **2018**, *56* (3), 164–171.
- (35) Gerber, I. C.; Jolibois, F. Theoretical Gas to Liquid Shift of  $^{15}\text{N}$  Isotropic Nuclear Magnetic Shielding in Nitromethane Using Ab Initio Molecular Dynamics and GIAO/GIPAW Calculations. *Phys. Chem. Chem. Phys.* **2015**, *17* (18), 12222–12227.
- (36) Charpentier, T.; Menziani, M. C.; Pedone, A. Computational Simulations of Solid State NMR Spectra: A New Era in Structure Determination of Oxide Glasses. *RSC Adv.* **2013**, *3* (27), 10550–10578.
- (37) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.
- (38) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688.
- (39) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* **2012**, *4* (8), 1–17.
- (40) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort; Fox, D. J. Gaussian 09, Revision A.02. *Gaussian, Inc.*,

Wallingford CT **2016**, 2016.

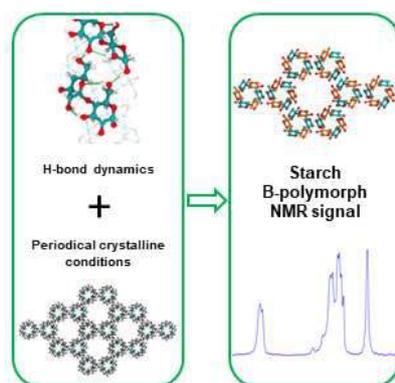
- (41) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.
- (42) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789.
- (43) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *J. Chem. Phys.* **1980**, *72* (1), 650–654.
- (44) McLean, A. D.; Chandler, G. S. Contracted Gaussian Basis Sets for Molecular Calculations. I. Second Row Atoms,  $Z=11-18$ . *J. Chem. Phys.* **1980**, *72* (10), 5639–5648.
- (45) Ditchfield, R. Molecular Orbital Theory of Magnetic Shielding and Magnetic Susceptibility. *J. Chem. Phys.* **1972**, *56* (11), 5688–5691.
- (46) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilizaion of AB Initio Molecular Potentials for the Prevision of Solvent Effects. *Chem. Phys.* **1981**, *55* (1), 117–129.
- (47) Sarkar, A.; Pérez, S. PolySac3DB: An Annotated Data Base of 3 Dimensional Structures of Polysaccharides. *BMC Bioinformatics* **2012**, *13* (1), 302.
- (48) Imberty, A.; Pérez, S. A Revisit to Three-Dimensional Structure of B-Type Starch. *Biopolymers* **1988**, *27*, 1205–1221.

- (49) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6* (1), 15–50.
- (50) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47* (1), 558–561.
- (51) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169–11186.
- (52) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metamorphous- Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49* (20), 14251–14269.
- (53) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50* (24), 17953–17979.
- (54) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59* (3), 1758–1775.
- (55) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.
- (56) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104–154119.
- (57) Vasconcelos, F.; De Wijs, G. A.; Havenith, R. W. A.; Marsman, M.; Kresse, G. Finite-Field Implementation of NMR Chemical Shieldings for Molecules: Direct

and Converse Gauge-Including Projector-Augmented-Wave Methods. *J. Chem. Phys.* **2013**, *139* (1), 014109–014116.

- (58) Marchessault, R. H.; Perez, S. Conformations of the Hydroxymethyl Group in Crystalline Aldohecopyranoses. *Biopolymers* **1979**, *18* (9), 2369–2374.
- (59) Nishida, Y.; Ohrui, H.; Meguro, H. <sup>1</sup>H-NMR Studies of (6R)- and (6S)-Deuterated D-Hexoses: Assignment of the Preferred Rotamers About C5-C6 Bond of D-Glucose AND D-Galactose Derivatives in Solutions. *Tetrahedron Lett.* **1984**, *25* (15), 1575–1578.

### Entry for the Table of Contents



## 4. Conclusion : Validité et limites de notre approche.

Les calculs que nous avons réalisés ont permis de mettre en évidence l'importance du choix et de la robustesse de la méthode de calcul ainsi que celle de la pertinence du modèle chimique utilisés dans l'étude de ces complexes.

Concernant le polymorphe V, une modification du champ de force a été nécessaire afin de stabiliser la conformation hélicoïdale de l'amylose lors des calculs de dynamique moléculaire. Par la suite, nous avons pu démontrer que notre modèle chimique devait inclure la dynamique intrinsèque du polymorphe afin de permettre une reproduction plus fidèle des spectres RMN  $^{13}\text{C}$  expérimentaux. Pour le polymorphe B, il a été nécessaire d'utiliser une méthode de calcul incluant les conditions périodiques afin de pouvoir différencier deux résidus selon leur déplacement chimique et effectuer la corrélation entre ce dernier et leur positionnement dans le cristal.

Il nous faut maintenant discuter des limites de ces modèles afin de pouvoir s'assurer qu'elles ne seront pas délétères pour la suite du travail de thèse.

Si l'inclusion de la dynamique dans le calcul de RMN nous a permis d'améliorer la reproduction des spectres RMN, les déplacements chimiques calculés des  $\text{C}_2$ ,  $\text{C}_3$  et  $\text{C}_5$  ne permettent pas de différencier deux pics, comme cela est observé expérimentalement. Ceci peut être expliqué par deux effets différents, l'un étant dû à la méthode de calcul, l'autre au modèle chimique que nous utilisons.

Les déplacements chimiques de ces carbones sont répartis expérimentalement sur une plage de 4 ppm. La précision de notre méthode de calcul est de l'ordre de quelques ppm pour les déplacements chimiques du carbone, et nous avons décidé d'utiliser une largeur à mi-hauteur de 2.5 ppm car celle-ci correspond à celle observée expérimentalement. Il est donc possible que la séparation de ces trois carbones selon leur déplacement chimique ne soit pas possible en raison de la limite de précision de notre méthode de calcul.

En contrepartie, les calculs que nous avons réalisés sur le polymorphe B ont démontré que les interactions entre deux hélices influencent le calcul des déplacements chimiques de ces trois carbones. Cela nous permet d'introduire la limite principale de notre modèle chimique : les liaisons hydrogènes intermoléculaires avec le solvant (ici

H<sub>2</sub>O). Il s'agit d'une interaction que nous avons décidé de négliger dans la majorité de nos calculs de RMN. Or, ces interactions sont connues comme pouvant influencer sur le déplacement chimique des carbones portant le groupe hydroxyle concerné par la liaison hydrogène ainsi que sur les carbones voisins. Il est donc tout à fait envisageable que ces interactions influencent les déplacements chimiques des trois carbones cités ci-dessus.

Dans ce travail de thèse, notre objectif principal réside dans la caractérisation de complexes d'inclusion dans des échantillons contenant des lipides de structures primaires plus complexes que celles des acides gras en présence de polysaccharides linéaires ou branchés. Pour ceci, nous nous intéresserons aux déplacements chimiques des deux carbones formant la liaison glycosidique, le C<sub>1</sub> et le C<sub>4</sub>, car ils permettent de caractériser la présence ou l'absence de complexes d'inclusion de type polymorphe V. Nous avons pu démontrer que la méthode de calcul ainsi que le modèle chimique choisis ci-dessus permettaient de reproduire le déplacement chimique des ces deux carbones. Les limitations de notre stratégie de calcul ne seront donc pas un obstacle pour la suite de ce travail.

## Chapitre 2. Etude théorique et expérimentale de complexes d'amylose en présence de lipides divers

La méthodologie théorique ayant été détaillée, il est maintenant temps de décrire les résultats obtenus sur des systèmes comprenant des lipides qui possèdent une structure primaire plus complexe que l'acide palmitique. Rappelons que les deux polymorphes décrits dans le chapitre précédent présentent une signature RMN différente<sup>1,2</sup> (Figure 39). Deux carbones de l'unité glucose de l'amylose vont notamment nous servir de « sonde » : le C<sub>1</sub>, qui est caractérisé par un pic unique à ~102-103 ppm en polymorphe V et deux pics à 100 ppm et 99 ppm en polymorphe B et le C<sub>4</sub>, dont le déplacement chimique est de ~82 ppm en polymorphe V et est compris dans le massif entre 69 et 76 ppm en polymorphe B.

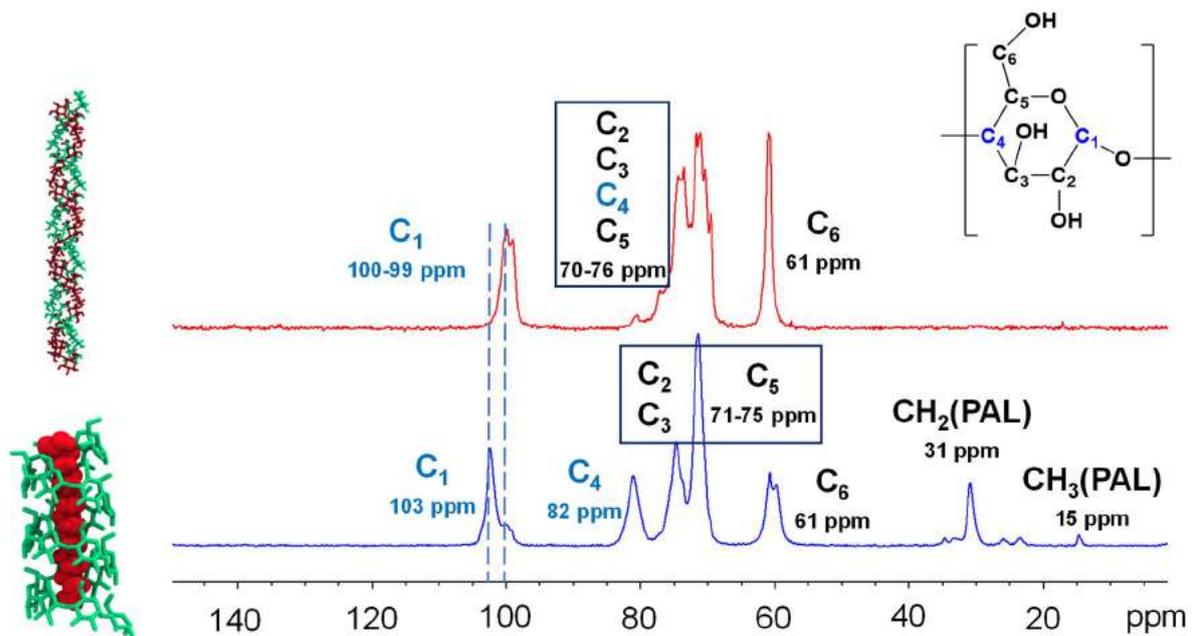


Figure 39: Spectres expérimentaux obtenus par RMN <sup>13</sup>C sur un échantillon de polymorphe B (rouge) et de polymorphe V (bleu). L'expérience RMN correspond à un spectre CP-MAS avec un échantillon en rotation à 10 kHz.

Dans un premier temps, il est nécessaire de s'attarder sur les méthodes de fabrication de ces complexes. En effet, certains paramètres tels que la longueur moyenne du polysaccharide utilisé ou le taux d'hydratation de l'échantillon sont connus pour influencer sur le spectre RMN obtenu<sup>2</sup>. Il était donc essentiel de les décrire avant de commencer des expériences sur des systèmes plus complexes. De plus, les matières premières utilisées ainsi que les méthodes d'obtention de ces dernières seront aussi décrites.

En parallèle de l'étude de ces paramètres, l'étude de complexes obtenus en présence d'acides gras et phospholipides a été réalisée.

Dans la littérature, les structures de l'amylose en polymorphe V ont rarement été caractérisées en présence de lipides contenant plus d'une chaîne lipidique. Sang et coauteurs ont réalisé des mesures des propriétés de viscosité et des propriétés thermiques de mélanges contenant de l'amidon de blé et des lipides extraits de jaune d'œuf de poule<sup>3</sup>. La présence d'un pic endothermique observé en DSC à 98.8°C a permis de suggérer la formation de polymorphe V en présence d'une faible quantité de lipides de type phosphatidylcholine. Concernant les caractéristiques structurales des complexes formés, la diffraction aux rayons X ainsi que la dynamique moléculaire ont été utilisées. Le profil de diffraction aux rayons X montre un pic présent à  $2\theta = 20.1^\circ$ , ce qui pourrait être le signe de la formation de polymorphes V. Cependant ce pic est aussi présent dans leur échantillon contrôle (contenant seulement l'amidon utilisé), et les pics observés de manière usuelle dans ce genre d'échantillon, à savoir ceux présents à  $2\theta = 13^\circ$  et  $2\theta = 7.6^\circ$ , ne sont eux pas observés. Il est donc difficile de conclure de manière univoque que la présence du motif polymorphe V est due à la présence de lipides. Les calculs de dynamique moléculaire présentent un polymorphe de 13 résidus replié de manière hélicoïdale autour d'une molécule de POPC cependant on observe une chaîne qui sort de l'hélice à travers la paroi, et le pas de l'hélice formée est de 13.76 Å ce qui n'est pas cohérent avec les données connues sur ce type de complexes (le pas d'hélice d'un polymorphe V étant attendu à environ 8 Å).

Nous avons réalisé des expériences de complexation sur deux acides gras et trois phospholipides différents et nous nous référerons à ces lipides comme étant des lipides « classiques » dans la suite de ce manuscrit, dans le sens où ceux-ci sont disponibles commercialement. La deuxième partie de ce chapitre vise à décrire les résultats obtenus par RMN du solide du <sup>13</sup>C ainsi que les calculs de dynamique moléculaire et de RMN par DFT réalisés sur ces complexes. Ces dernières permettront de proposer un modèle reliant la structure primaire des lipides, le protocole de synthèse utilisé et la propension du système à réaliser des complexes de type polymorphe V.

Enfin, nous décrirons les résultats des expériences obtenues en utilisant des lipides issus de *Mycobacterium tuberculosis* et de *Corynebacterium glutamicum*. Ces derniers présentent des structures similaires aux phospholipides utilisés dans la première partie de ce chapitre, mais peuvent contenir jusqu'à quatre chaînes lipidiques ainsi qu'une tête polaire différente. Nous verrons que les résultats de complexation peuvent s'inscrire dans le modèle antérieurement décrit, appuyant sa validité.

## 1. Avant-propos : méthodes d'études et matières premières

### I. Matières premières et modélisation des complexes

Le choix des matières premières utilisées est important dans l'étude de ces complexes, car il peut influencer sur la formation du polymorphe V. De ce fait, nous allons maintenant présenter les choix réalisés quant à la provenance de l'amylose et des lipides. De plus, si le protocole utilisé pour la modélisation de ces complexes est fortement similaire à celui décrit dans la partie précédente, nous avons effectué quelques modifications qui doivent être décrites.

#### a. Choix du type d'amylose

Plusieurs tests de complexation ont été tentés avec de l'amylose de pomme de terre commerciale, et deux échantillons d'amylose produite par une enzyme appelée amylosucrase issue de *Deinococcus geothermalis* (AsDg)<sup>4</sup>. Cette enzyme utilise du sucrose comme substrat et permet de faire varier le degré de polymérisation (DP) moyen de l'amylose produite, simplement en modifiant la concentration initiale de sucrose ajouté à la solution<sup>4,5</sup>. Deux échantillons d'amylose ont été synthétisés grâce à cette enzyme qui nous a été fournie par Florent Grimaud, du Toulouse Biotechnology Institute. Ces deux échantillons d'amylose, appelés A400 et A600, ont été produits en suivant des protocoles déjà publiés, et les concentrations initiales en sucrose dans le milieu réactif étaient de 400 et 600 mM, respectivement.

b. *Provenance des lipides :*

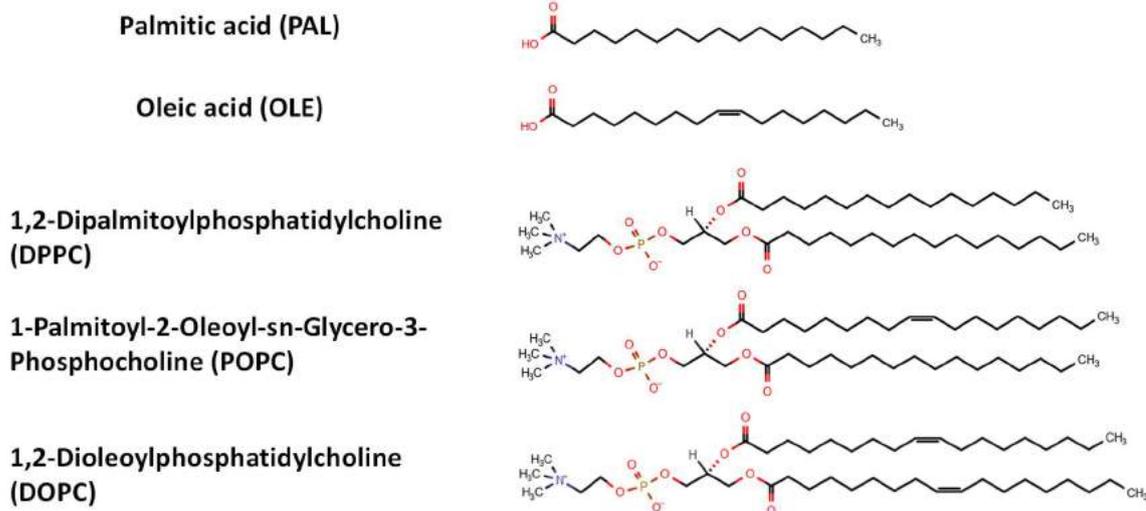


Figure 40: Structure des lipides "classiques" utilisés : acide palmitique (PAL), acide oléique (OLE), 1,2-dipalmitoylphosphatidylcholine (DPPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-dioleoylphosphatidylcholine (DOPC)

Les lipides classiques (Figure 40), disponibles commercialement, sont l'acide palmitique (PAL), l'acide oléique (OLE), le 1,2-Dipalmitoylphosphatidylcholine (DPPC), le 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) et le 1,2-Dioleoyl-sn-Glycero-3-Phosphocholine (DOPC). Ils ont été acquis chez Sigma-Aldrich.

Concernant les lipides mycobactériens (Figure 40), des acides 3-hydroxy-2-tetradecyloctadecanoïque, aussi appelés acides corynomycoliques (MYC), des monomycolates de tréhalose (MMT) ainsi que des dimycolates de tréhalose (DMT), nous ont été fournis par Maryelle Tropis de l'équipe « Mycobacterial Envelopes : Structures, Biosynthesis and Function » de l'Institut de Pharmacologie et Biologie Structurale (IPBS). Ces lipides issus de *Corinebacterium glutamicum* présentent une distribution d'insaturations (0 à 2) et de longueurs de chaîne acyle (en C32 à C36). Du 3-oxo-2-tetradecyloctadecanoate de méthyle aussi appelé acide  $\beta$ -cétomycolique (BET) nous a été fourni par A. Lemassu, de cette même équipe.

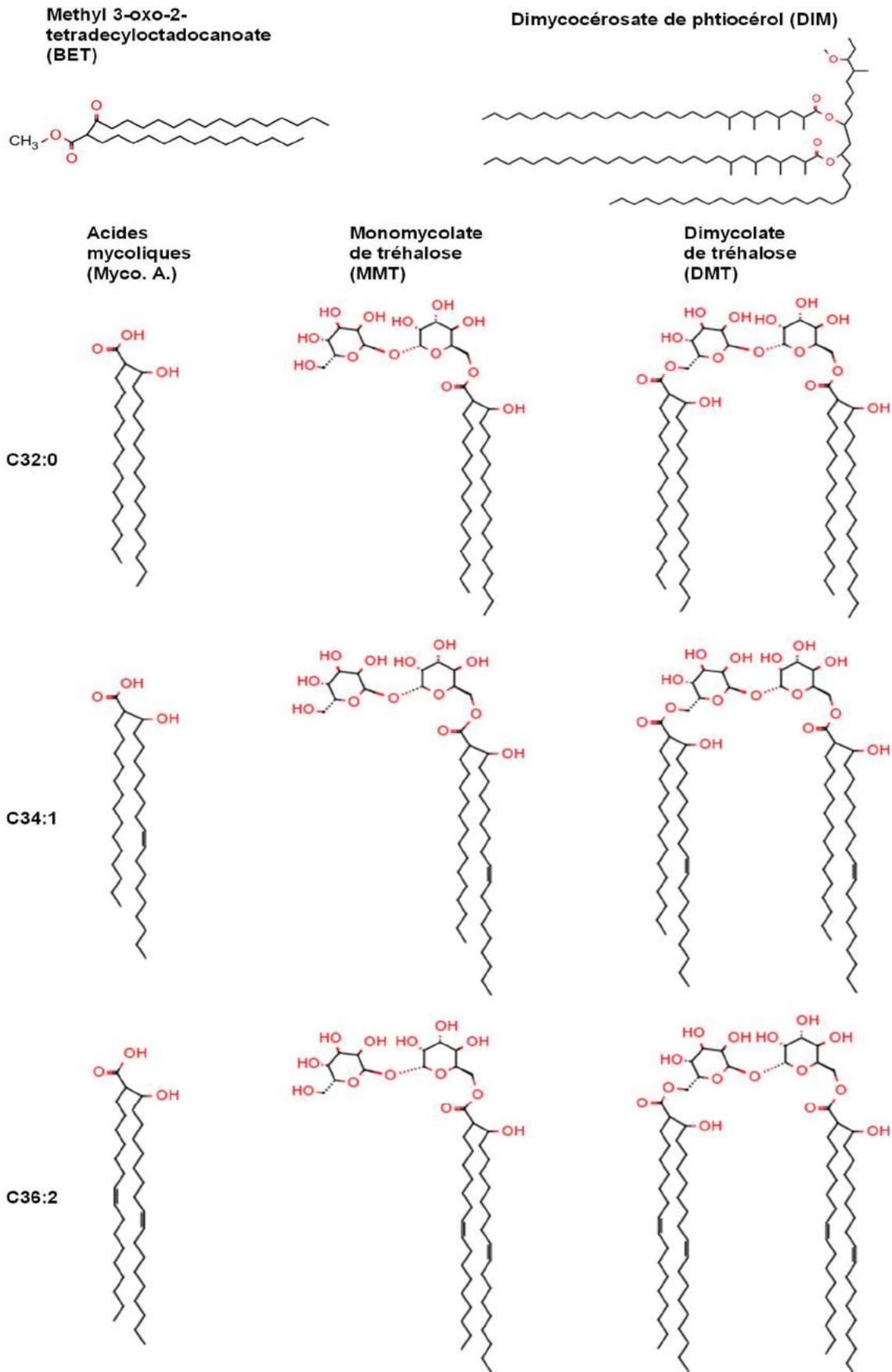


Figure 41 : Structure des lipides bactériens utilisés : 3-oxo-2-tetradecyloctadecanoate de méthyle (BET), dimycocérosate de phtiocérol (DIM), acides mycoliques (MYC), monomycolate de tréhalose (MMT) et dimycolate de tréhalose (DMT). Les échantillons de MYC, MMT et DMT utilisés sont des mélanges des structures C32:0, C34:1 et C36:2.

Les dimycocérosates de phtiocérol (DIM) ont été purifiés à partir de cultures de *M. tuberculosis* H37Rv cultivées 3 semaines sur milieu minimum 7H9 supplémenté en glycérol dans un laboratoire de niveau de sécurité 3 (LBS3). Les bactéries ont été mises en présence d'un mélange  $\text{CHCl}_3:\text{CH}_3\text{OH}$  (1:2 [vol :vol]) pour 48h avant d'être sorties du laboratoire LBS3. Après une deuxième extraction dans un mélange  $\text{CHCl}_3:\text{CH}_3\text{OH}$  (2:1 [vol :vol]) pendant 24h, les phases organiques ont été filtrées, rassemblées, puis lavées avec de l'eau et séchées. Environ 50mg de cet extrait lipidique brut a été dissous dans 3mL d'éther de pétrole (EP), et les PDIM ont ensuite été purifiés sur un système de chromatoflash équipé d'une colonne de silice Puriflash (4g 15 $\mu\text{m}$ ) en utilisant un débit de 20mL/min. Le dépôt a été élué successivement par 20 volumes de colonne d'EP, 10 volumes d'un mélange EP:éther diéthylique(E) (9:1 [vol :vol]), 23 volumes de 100% d'E, puis par du  $\text{CHCl}_3$  et du  $\text{CH}_3\text{OH}$  pour nettoyer la colonne. Des fractions de 2mL ont été collectées, concentrées et reprises dans environ 0,3mL de  $\text{CHCl}_3$ . La présence de PDIM a été testée par chromatographie sur couche mince de silice dans un solvant de migration EP:E (9:1 [vol :vol]), puis révélation à la primuline et comparaison avec des témoins. Les fractions contenant des PDIM, en général les fractions 15 à 20, ont été rassemblées, puis pesées.

### *c. Modèles de lipides et structures initiales pour la dynamique moléculaire*

Maintenant que les composants importants de la partie expérimentale ont été décrits, nous allons nous attarder sur la partie théorique de l'étude de ces complexes. Pour commencer, il faut préciser que les calculs réalisés ne peuvent malheureusement pas explorer entièrement l'espace des configurations disponibles pour ce type de systèmes. Cependant, en faisant varier les structures initiales utilisées pour générer nos trajectoires et en les analysant en regard avec des données expérimentales, nous espérons apporter des informations structurales complémentaires permettant une meilleure appréhension des interactions entre ces lipides et l'amylose.

Le modèle de polysaccharide reste le même que celui présenté dans le chapitre précédent. Les lipides classiques ont été décrits en utilisant le champ de force LIPID14<sup>6</sup>, par l'intermédiaire de la routine tleap<sup>7</sup>. Les paramètres concernant les MMT, DMT et DIM ont été déterminés par G. Czaplicki, de l'équipe de RMN biologique et

intégrative de l'IPBS et F. Jolibois, de l'équipe de modélisation physico-chimique du Laboratoire de Physique et Chimie des Nano-Objets, et ont été décrits précédemment<sup>8,9</sup>. Pour le BET, les paramètres utilisés sont ceux du General Amber Force Field (GAFF) et les charges ont été calculées au niveau AM1-BCC<sup>10</sup>. Le protocole utilisé pour les calculs de dynamique moléculaire est le même que celui décrit dans le chapitre 1 de cette thèse. En revanche, nous avons utilisé diverses structures initiales lors de nos calculs, pour illustrer différents effets liés à la structure primaire des lipides.

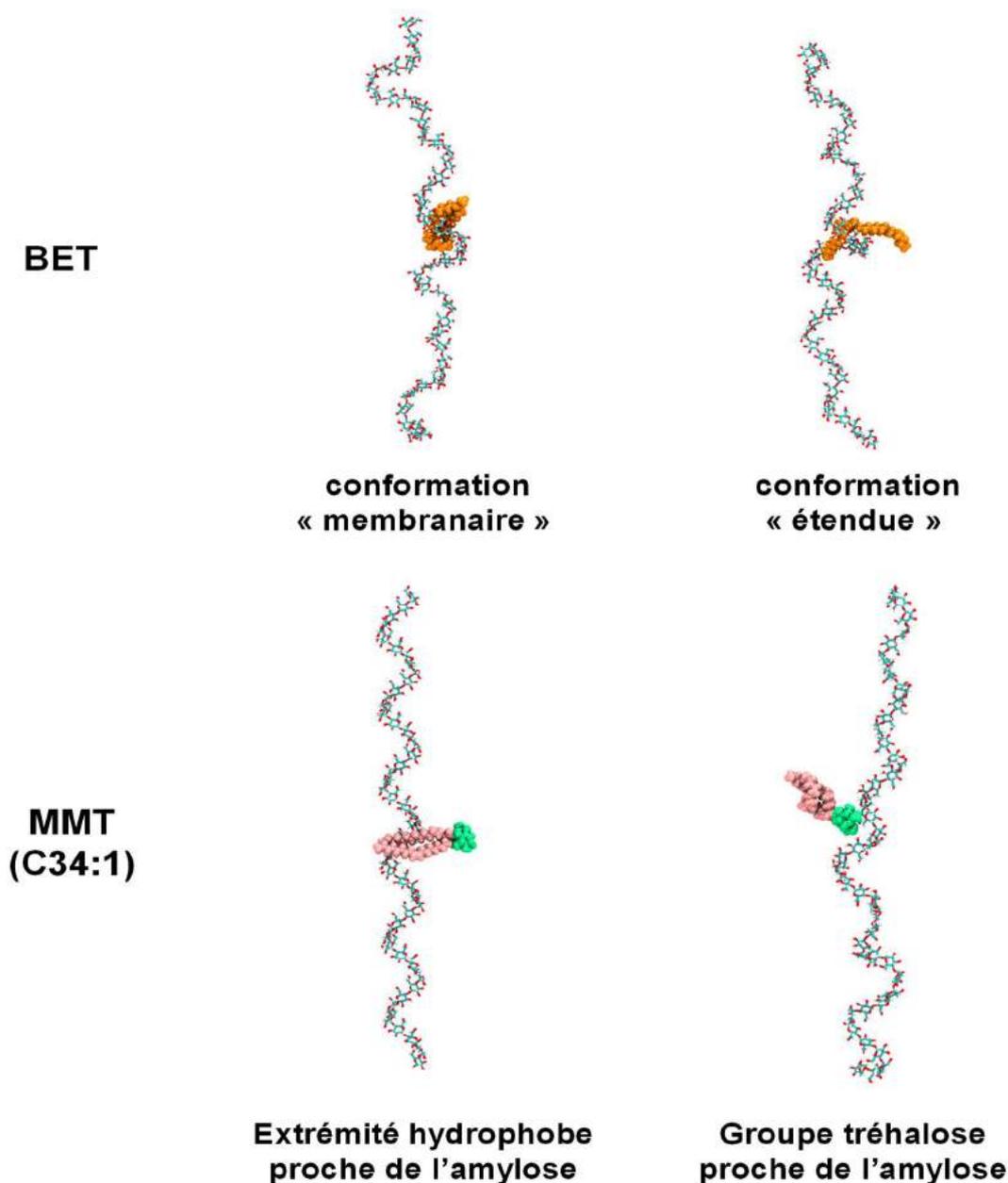


Figure 42: Conformations initiales des trajectoires de dynamique moléculaire réalisées en présence de BET (haut) et MMT (bas).

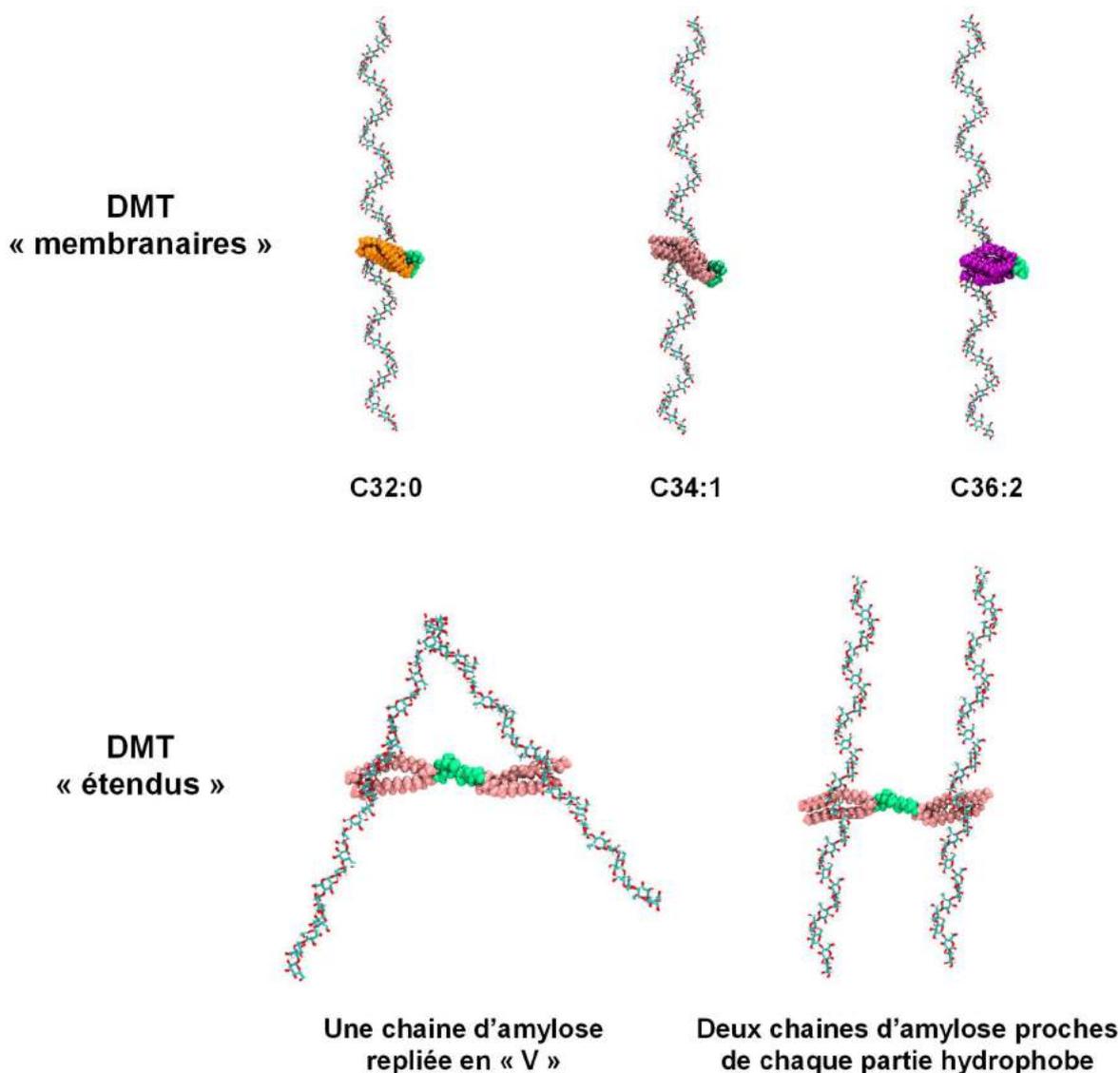


Figure 43 : Structures initiales des trajectoires de dynamique moléculaire réalisées en présence de DMT dans une conformation membranaire (haut) et étendue (bas)

Les acides gras présentait une conformation étendue, comme dans le chapitre 1 de cette thèse, et les phospholipides présentait une conformation similaire à celle observée dans des membranes avec les deux chaînes hydrophobes étendues l'une à proximité de l'autre. Pour le DPPC, deux calculs supplémentaires ont été réalisés dans lesquels le lipide était présent dans sa conformation « membranaire » au sein d'une chaîne d'amylose repliée en polymorphe V6 ou V8. Deux calculs ont été réalisés sur le BET, l'un dans une conformation de type « membranaire », l'autre sur une conformation complètement étendue du lipide, afin de tester une encapsulation complète de ce lipide. Pour le MMT (C34:1), deux calculs ont été réalisés afin de tester l'impact du groupe tréhalose : le premier présente une structure de départ avec le groupe tréhalose positionné à proximité de l'amylose, le deuxième avec les extrémités

hydrophobes positionnées proche de l'amylose. Les structures initiales correspondant aux trajectoires réalisées en présence de BET et de MMT sont présentées dans la Figure 42.

Enfin, 5 types de calculs ont été réalisés sur les DMT : 3 calculs utilisaient des conformations « membranaires » pour chaque type de DMT (C32:0, C34:1 et C36:2), un calcul présentait un DMT dans une conformation « étendue » à proximité d'une chaîne d'amylose de 45 résidus repliée en forme de « V » et le dernier présentait un DMT étendu positionné à proximité de deux chaînes d'amylose de 30 résidus, chacune positionnée à proximité d'une des deux parties hydrophobes du lipide. Les structures initiales correspondant aux conformations membranaires et étendues sont présentées dans la Figure 43.

## II. Protocole de complexation et analyses des résultats

Afin de réaliser les complexes de type polymorphe V, plusieurs types de protocoles ont été décrits dans la littérature, pouvant être séparés en trois parties : les méthodes dites classiques, les méthodes enzymatiques et les méthodes thermomécaniques<sup>11</sup>. Ceci ayant déjà été précisé dans l'introduction, nous allons maintenant décrire notre protocole, issu des méthodes classiques et qui a été légèrement adapté de sa version publiée par Lebail et al. en 2000<sup>2</sup>.

Dans un Falcon de 15 ml, sont dissouts 60 mg d'amylose dans (1.5 – x) mL de DMSO à 90°C. x mL de solution de lipides dissouts dans du DMSO à 90°C sont ajoutés à la solution d'amylose, de manière à obtenir le pourcentage molaire lipide/glucose désiré. Afin de pouvoir effectuer des comparaisons entre les différents lipides utilisés, nous avons défini le pourcentage molaire de la façon suivante :  $\%_{molaire} = \frac{n_{lip}}{n_{glucose}} \times 100$ . Ici  $n_{glucose}$  représente le nombre de mole de glucose (masse molaire = 180 mg/mol) correspondant à masse d'amylose donnée.

Les concentrations des solutions de lipides « x » ont été ajustées de manière à obtenir les ratios molaires décrits dans le Tableau 7 et pour que le volume ajouté à la solution d'amylose ne dépasse pas 0.5 ml. Le mélange est maintenu à 90°C pendant 4h afin de promouvoir une dissolution complète des lipides et de l'amylose. En suivant, 2.5 ml

d'eau MilliQ portée à 90°C sont ajoutés dans le Falcon, puis le mélange est refroidi jusqu'à 25°C par pas d'un degré toutes les dix minutes. En général, une précipitation est observée dans le Falcon.

Tableau 7: Pourcentages molaires ( $\%_{molaire} = \frac{n_{lip}}{n_{glucose}} \times 100$ ) et masses utilisées pour la réalisation des complexes amylose/lipide.

$\%_{molaire}$	0.78%	1.56%	3.13%	6.25%	12.5%
Masse amylose (mg)	60	60	60	60	60
Masse PAL (mg)	0.75	1.5	3.0	6.0	12.0
Masse OLE (mg)	0.87	1.65	3.3	6.6	13.2
Masse DPPC (mg)	2.15	4.3	8.6	17.2	-
Masse POPC (mg)	2.23	4.45	8.9	17.8	35.6
Masse DOPC (mg)	2.3	4.6	9.2	18.4	36.8
Masse BET (mg)	1.48	2.97	5.95	11.9	-
Masse Myco. A. (mg)	-	-	-	12.2	-
Masse MMT (mg)	-	-	-	15.3	-
Masse DMT (mg)	-	-	-	30.0	-
Masse DIM (mg)	-	9.7	17.4	-	-

L'échantillon est centrifugé 30 minutes à 10 000 G à température ambiante. Le surnageant est prélevé et conservé à 4 °C. Le culot est suspendu dans 2 ml d'un mélange 50/50 eau/éthanol afin de procéder au lavage de l'échantillon et d'éliminer les lipides non complexés. Trois lavages sont réalisés au total. A partir de là, deux types d'échantillons sont à dissocier : ceux dont le culot a été lyophilisé et ceux dont le culot a été laissé à sécher à l'air libre. La majorité des résultats présentés dans ce travail ont été obtenus sur des échantillons lyophilisés pendant 24 heures, puis réhydratés avec une masse d'eau équivalente à la masse d'échantillon déshydraté, nous précisons donc dans la suite de ce chapitre si l'échantillon n'a pas été lyophilisé. Le complexe précipité et hydraté d'amylose/lipide est transféré dans un rotor en zirconium de 4mm de diamètre pour enregistrement des spectres RMN.

Les expériences de RMN ont été réalisées sur un spectromètre Bruker DMX avec une fréquence de Larmor du proton de 500.13 MHz. Les spectres ont été acquis en utilisant une sonde Bruker de polarisation croisée en rotation à l'angle magique (CP-MAS) 4mm

double résonance H/X. Les spectres ont été enregistrés sur les échantillons en rotation à la vitesse de  $\omega r/2\pi = 10$  kHz et à température ambiante. Les spectres  $^{13}\text{C}$  ont été obtenus en utilisant une séquence d'impulsion CP-MAS classique et la longueur d'impulsion  $\pi/2$  pour le proton était de 3.1  $\mu\text{s}$ . L'étape de transfert de magnétisation  $^1\text{H}$ - $^{13}\text{C}$  a été réalisée en utilisant une rampe linéaire (70% à 100% de la force du champ) sur le canal proton avec un temps de contact de 2 ms et un champ radiofréquence  $^{13}\text{C}$  de 62.5 kHz. Le découplage des protons a été réalisé en utilisant une séquence de type TPPM15 et un champ de découplage de 80.6 kHz. Le temps de recyclage a été fixé à 3 s et 12256 scans ont été acquis pour chaque spectre. Les données ont été traitées en utilisant le logiciel Bruker Topspin (3.6.2) et l'élargissement de ligne a été fixé à 10 Hz. Après transformée de Fourier, les spectres ont subi une correction de phase et de ligne de base. La référence utilisée pour les déplacements chimiques du  $^{13}\text{C}$  est le TMS.

La capacité des différents lipides à engendrer la structuration en polymorphe V de l'amylose a été déterminée en calculant le pourcentage de polymorphe V présent dans nos échantillons. Pour ceci, nous avons réalisé un ajustement de 8 fonctions lorentziennes aux spectres expérimentaux obtenus. 3 de ces fonctions ont été utilisées pour reproduire le signal des C1 (99-103 ppm), une pour celui des C4 (82 ppm), 2 pour le signal des C6 (59-61 ppm) et les deux dernières ont été ajustées au massif compris entre 70 et 76 ppm. Nous avons ensuite défini le pourcentage de polymorphe V dans nos échantillons de la manière suivante :  $\%_{\text{polymorphe V}} = \%_{\text{PV}} = \frac{\text{Aire}_{\text{C4V}}}{\text{Aire}_{\text{C1T}}} \times 100$ . Ici  $\text{Aire}_{\text{C4V}}$  définit l'aire sous la fonction lorentzienne correspondant au pic présent à 82 ppm, signal correspondant aux C4 des glucoses inclus dans la structure du polymorphe V, et  $\text{Aire}_{\text{C1T}}$  définit l'aire sous les 3 fonctions lorentziennes correspondant aux signaux du C1 de tous les glucoses présents dans l'échantillon (signaux compris entre 99 et 103 ppm). L'erreur sur la détermination du  $\%_{\text{PV}}$  est estimée à +/- 3% en tenant compte de la reproductibilité des échantillons et de la précision de l'ajustement des lorentziennes.

## 2. Complexation de lipides « classiques »

La structure primaire du lipide utilisé est connue pour avoir une influence sur la formation de polymorphes V<sup>12</sup>. Avant d'utiliser des lipides issus de mycobactéries, nous nous sommes concentrés sur des lipides simples, dont les structures primaires sont analogues (une tête polaire liant une ou plusieurs chaînes aliphatiques) à celles des lipides mycobactériens que nous voulions étudier. Comme nous l'avons décrit au début de ce chapitre, le polymorphe V a été très peu caractérisé dans la littérature en présence de lipides contenant plus d'une chaîne hydrophobe. Il était donc nécessaire de réaliser une étude sur des lipides disponibles commercialement avant de commencer à étudier des lipides mycobactériens.

### I. Résultats expérimentaux

#### a. Choix de l'amylose et taux d'hydratation des échantillons

Le choix de l'amylose est un paramètre déterminant pour la complexation de ces lipides. Afin de choisir le type d'amylose que nous allons utiliser par la suite, nous avons utilisé de l'amylose de pomme de terre, de l'A400 et de l'A600 en présence d'acide palmitique (PAL). Ces trois échantillons présentaient un ratio massique de 10% (soit un ratio molaire de 6.26%).

Les spectres RMN <sup>13</sup>C des échantillons obtenus sur ces derniers sont présentés dans la Figure 44. Les proportions relatives entre le pic du C<sub>1</sub> à 102.6 ppm et les pics à 100 et 99 ppm laissent à penser que l'amylose de pomme de terre engendre un pourcentage relatif de polymorphe V supérieur à ceux engendrés par l'A400 et l'A600. Le calcul de %<sub>PV</sub> le confirme : 81.4 +/-3% des résidus glucose de l'échantillon sont dans un polymorphe V en utilisant de l'amylose de pomme de terre et 60.4 +/-3% et 60.0 +/-3 % le sont pour les échantillons A400 et A600 respectivement. Deux paramètres permettent d'expliquer ce résultat : la longueur moyenne des chaînes d'amylose utilisée, et la cristallinité de cette dernière. L'amylose générée par AsDg présentant un degré de polymérisation moyen aux alentours de 35-45, ce dernier est grandement inférieur à celui de l'amylose de pomme de terre commerciale qui a

récemment été caractérisé aux alentours de 2500<sup>13</sup>. Il a été montré que l'amylose produite par l'amylosucrase de *Neisseria polysaccharea* formait des polymorphes B présentant une cristallinité élevée<sup>14</sup>, ce qui suggère que les chaînes d'amylose produites présentent un DP assez homogène pour permettre l'établissement d'un réseau régulier. Ces deux caractéristiques engendrent donc une compétition non-négligeable entre la formation de polymorphes V et de polymorphes B que nous cherchons ici à éviter. De ce fait, les prochains échantillons décrits ici seront ceux réalisés en présence uniquement d'amylose de pomme de terre commerciale.

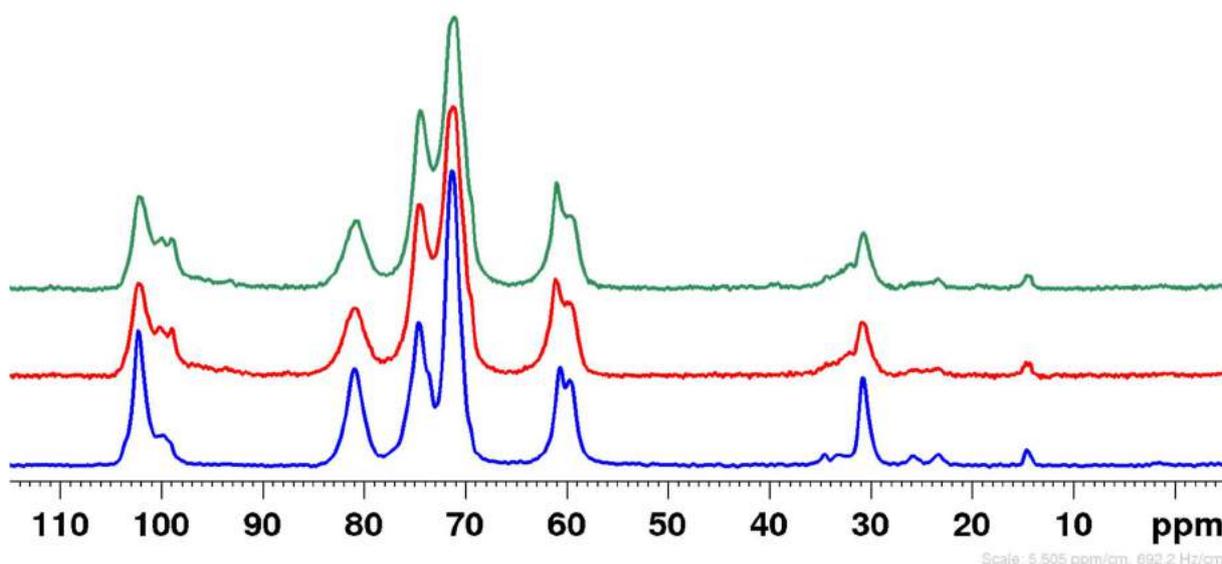


Figure 44: Spectres RMN <sup>13</sup>C d'échantillons contenant de l'acide palmitique en présence d'amylose de patate (Bleu), d'A400 (rouge) et d'A600 (Vert).

Il est établi dans la littérature que le taux d'hydratation de l'échantillon peut avoir un impact non négligeable sur l'échantillon<sup>2,15</sup>. Le protocole original utilisé par Le-bail et co-auteurs précise que les échantillons doivent être laissés à sécher à l'air libre après les lavages. Ce procédé ne permet pas un contrôle optimal du taux d'hydratation présent dans l'échantillon. Afin de mesurer l'impact de ce paramètre sur les signaux RMN obtenus, plusieurs spectres ont été enregistrés à différents taux d'hydratation, sur des échantillons contenant de l'amylose en présence de POPC. L'échantillon de POPC considéré ici comprend un ratio molaire de 6.26%. Le spectre RMN a été enregistré sur l'échantillon séché à l'air libre après lavage (Figure 45, Bleu). L'échantillon a ensuite été lyophilisé et un second spectre RMN a été enregistré.

La lyophilisation provoque un élargissement des pics observés sur l'échantillon (Figure 45, Rouge) qui peut être expliqué par le manque de dynamique du système. Plusieurs types de conformations différentes sont ainsi piégés et ne peuvent plus s'échanger. Nous avons réalisé des réhydratations progressives, de façon que la masse d'eau présente dans l'échantillon soit égale à 10%, 25% et 50% de la masse totale de l'échantillon. La réhydratation que nous appelons 50% correspond donc à une masse d'eau équivalente à la masse de l'échantillon lyophilisé. L'ajout progressif d'eau permet une diminution de la largeur à mi-hauteur des pics en introduisant une flexibilité et une diminution de l'hétérogénéité de l'échantillon. Par soucis de clarté la Figure 45 ne présente que la réhydratation à 50% (vert), qui représente le taux d'hydratation nécessaire et suffisant pour ne plus voir évoluer les largeurs de raie des signaux et retrouver un spectre carbone similaire à celui obtenu quand l'échantillon a été laissé 24h à l'air libre. Nous avons donc décidé pour la suite de l'étude de ce type d'échantillon de systématiquement procéder à la lyophilisation, puis réhydratation à 50% des échantillons afin d'être sûr que ce paramètre ne varie pas et donc n'influe pas sur l'aire sous la courbe des différents pics.

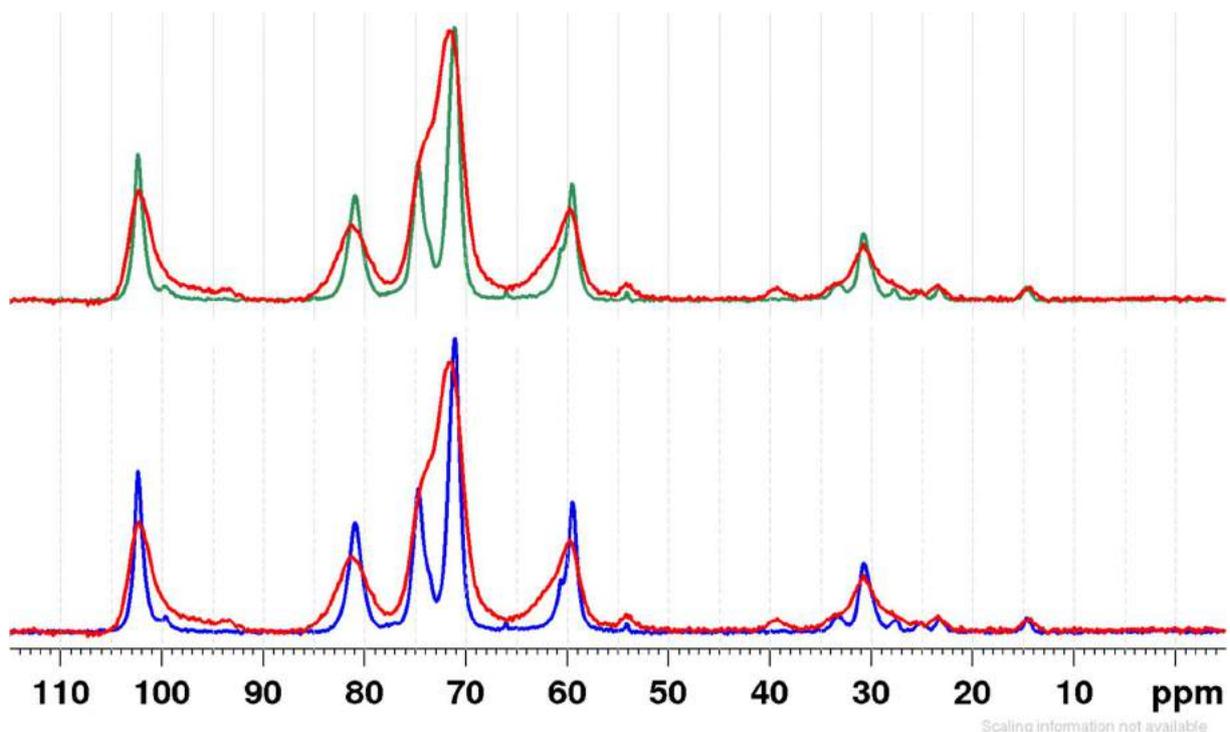


Figure 45 : Comparaison du spectre RMN  $^{13}\text{C}$  obtenu en fonction du taux d'hydratation de l'échantillon. Bleu : Echantillon laissé à l'air libre 24h. Rouge: Echantillon lyophilisé. Vert: Echantillon réhydraté avec une masse d'eau MilliQ équivalente à la masse d'échantillon lyophilisé.

## b. Etude comparative de la complexation sur les lipides classiques

Pour commencer, nous nous sommes basés sur les expériences réalisées par Le-Bail et coauteurs<sup>2</sup>. Au cours de ces dernières, des échantillons contenant différentes concentrations d'acide palmitique ont été fabriqués pour être étudiés par RMN <sup>13</sup>C à l'état solide. Ils ont montré que plus la concentration en lipide était importante plus les intensités des signaux du C<sub>1</sub> et C<sub>4</sub> associés au polymorphe V étaient élevées. Nous avons observé un comportement similaire sur des échantillons équivalents, réhydratés à 50% en masse et contenant différentes quantités d'acide palmitique (Tableau 7, Figure 46).

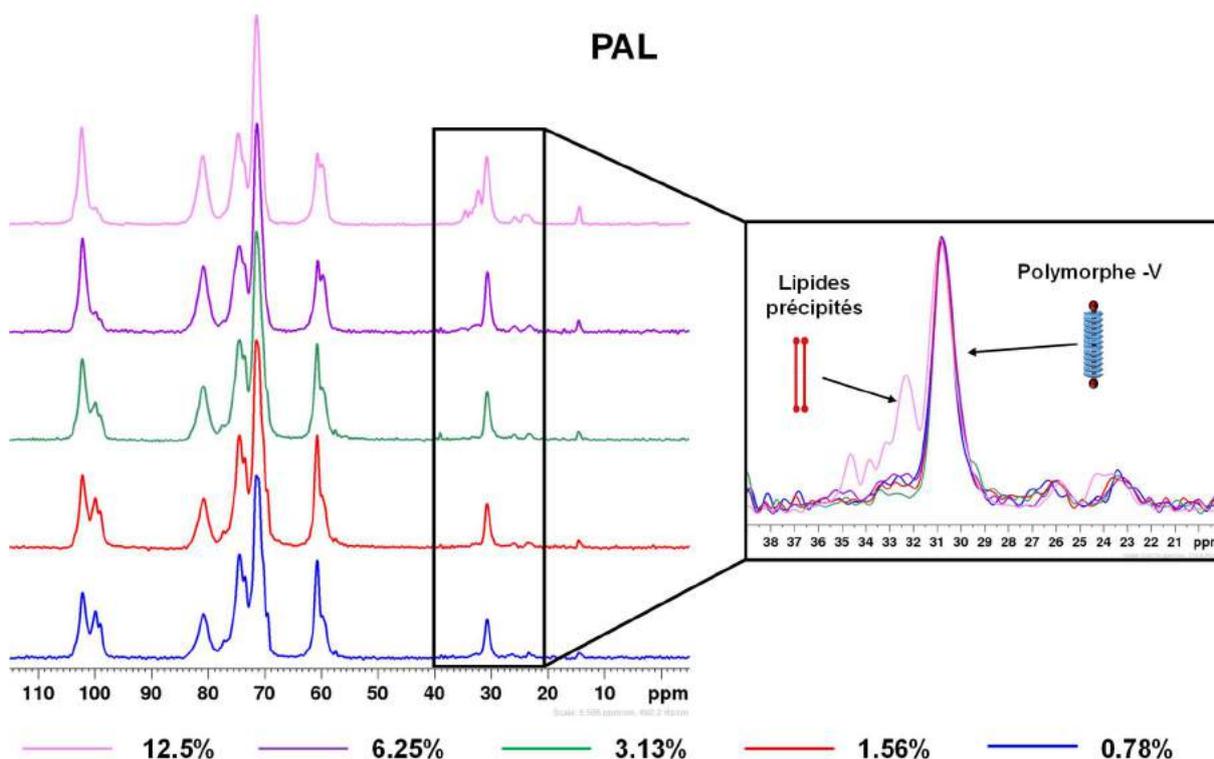


Figure 46: Spectres RMN <sup>13</sup>C d'échantillons d'amylose-acide palmitique dont les ratios molaires sont de 0.78% (bleu), 1.56% (rouge), 3.13% (vert), 6.25% (violet) et 12.5% (rose). Droite : agrandissement de la partie du spectre contenant le signal RMN des carbones CH<sub>2</sub> des lipides présents dans les échantillons.

En effet, plusieurs comportements peuvent être notés concernant l'évolution du signal en fonction de la concentration en lipide de l'échantillon. L'augmentation de cette concentration entraîne une diminution du signal des C<sub>1</sub> présent à 99-100 ppm en faveur de celui présent à 102.6 ppm, ce qui témoigne de l'augmentation de la quantité

de polymorphe V dans l'échantillon. Nous observons aussi la diminution du signal présent à 61 ppm en faveur du signal présent à 59 ppm. Au vu des résultats théoriques obtenus sur l'étude du complexe amylose-acide palmitique, nous pouvons affirmer que le signal présent à 61 ppm correspond au signal des C<sub>6</sub> qui ne sont pas impliqués dans la formation d'une structure en polymorphe V. Nous avons tracé l'évolution du pourcentage de polymorphe V défini plus haut (%<sub>PV</sub>) en fonction de la concentration en acide palmitique (Figure 48) et montré ainsi l'augmentation de la proportion de polymorphe V associée à l'augmentation de la quantité de lipide.

Le signal observé à 30.5 ppm correspond à celui des chaînes lipidiques incluses dans l'amylose, et dans l'échantillon présentant un ratio molaire de 12.5%, on observe un signal à 32.5 ppm correspondant au signal de lipides précipités<sup>2</sup>. Cette série d'échantillons nous servira de référence à des fins de comparaison avec les autres lipides classiques utilisés.

La première modification dans la structure des lipides qui nous intéresse réside dans l'ajout d'une insaturation de type Z en position 9 de la chaîne lipidique car il s'agit du type de chaînes acyles que l'on retrouve dans les lipides mycobactériens étudiés ci-

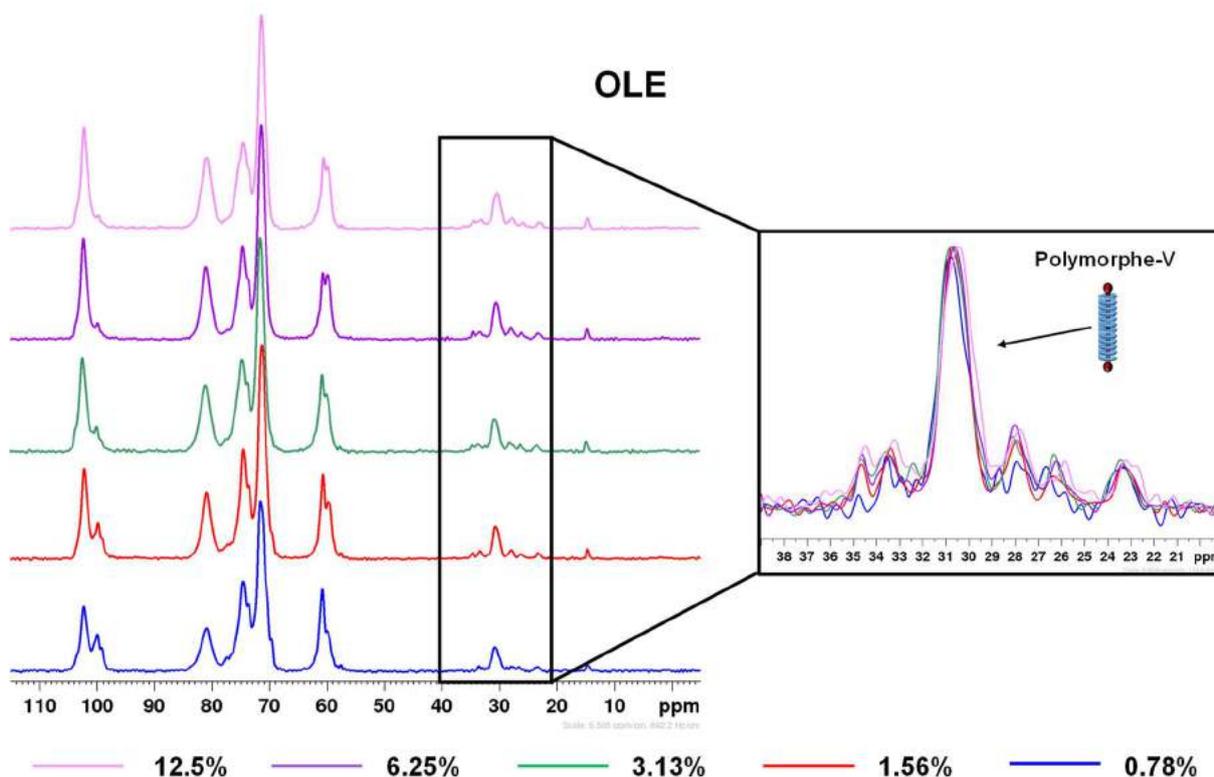


Figure 47: Spectres RMN <sup>13</sup>C d'échantillons d'amylose-acide oléique dont les ratios molaires sont de 0.78% (bleu), 1.56% (rouge), 3.13% (vert), 6.25% (violet) et 12.5% (rose). Droite : agrandissement de la partie du spectre contenant le signal RMN des carbones CH<sub>2</sub> des lipides présents dans les échantillons.

après. Nous avons donc réalisé des échantillons d'amylose-acide oléique aux mêmes ratios molaires que ceux utilisés pour l'acide palmitique (Figure 47).

Le comportement général des échantillons est similaire à celui observé dans le cadre de l'acide palmitique. Plus la concentration en lipides augmente dans l'échantillon, plus le signal RMN du  $^{13}\text{C}$  présent à 99-100 ppm diminue au profit du signal à 102.6 ppm et plus le signal à 61 ppm diminue au profit de celui présent à 59 ppm. Cependant, concernant le signal des lipides, quelle que soit la concentration, aucun signal ne peut être attribué à la présence de lipides précipités. Ceci peut être expliqué par la température de fusion de ce lipide qui est inférieure à la température ambiante. De ce fait, l'acide oléique libre est plus à même de rester en solution ou d'être solubilisé lors des lavages subséquents.

L'évolution du pourcentage de polymorphe V ( $\%_{\text{PV}}$ ) en fonction de la concentration en acide oléique est similaire à celle observé pour l'acide palmitique (Figure 48). On observe aussi que l'acide oléique induit systématiquement un pourcentage de polymorphe V supérieur à celui induit par l'acide palmitique, à concentration égale. L'ajout d'une insaturation semble donc faciliter la structuration de l'amylose en polymorphe V.

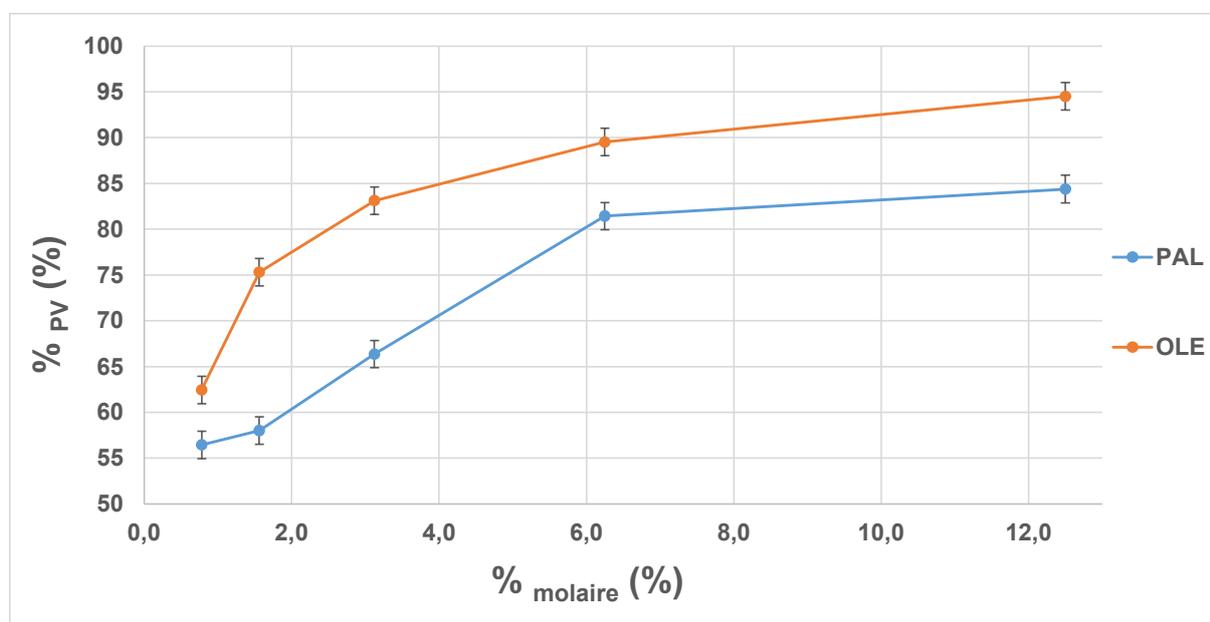


Figure 48: Pourcentage de polymorphe V induit par les acides gras en fonction du pourcentage molaire initial en lipide dans l'échantillon. Orange : Acide oléique. Bleu : Acide palmitique.

Par la suite, nous avons réalisé la complexation en présence de phospholipides : le DPPC, le POPC et le DOPC (Figure 40) présentant une même tête polaire

phosphatidyl choline et 2 chaînes d'acyle en C16:0 (sans insaturation), une chaîne en C16:0 et une en C18:1 (1 insaturation) et deux chaînes en C18:1 (2 insaturations).

Dans les échantillons contenant du DPPC, une différence de comportement est observée par rapport aux échantillons contenant des acides gras. En effet, l'échantillon contenant un ratio molaire de 3.13% présente une quantité de polymorphe V similaire à celui contenant un ratio molaire de 0.78%, comme l'atteste la comparaison des signaux des C<sub>1</sub> (Figure 49) et le %<sub>PV</sub> (Figure 13). De plus, le signal des CH<sub>2</sub> lipidiques semble montrer que, quelle que soit la concentration, on observe une présence de lipides précipités dans l'échantillon.

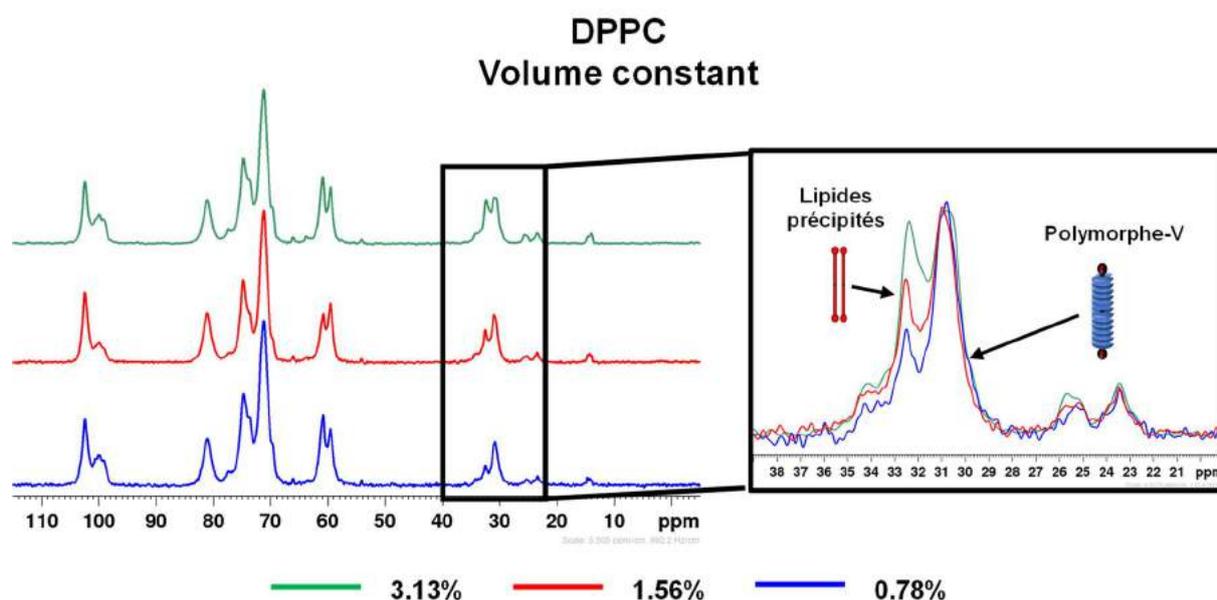


Figure 49 : Spectres RMN d'échantillons d'amylose-DPPC dont les ratios molaires sont de 0.78% (bleu), 1.56% (rouge) et 3.13% (vert). Droite : agrandissement de la partie du spectre contenant le signal RMN des carbones CH<sub>2</sub> des lipides présents dans les échantillons.

Au cours du protocole utilisé pour fabriquer ces échantillons, la dissolution des lipides dans du DMSO se fait à chaud (90°C). Le DPPC est le lipide qui a été le plus difficile à dissoudre dans ce solvant à concentration fixe, nécessitant régulièrement la sonication de la solution pour y parvenir. Cette information nous a amené à postuler que la concentration initiale en lipide en interaction avec l'amylose à 90°C dans le DMSO pouvait jouer un rôle dans la capacité de ces lipides à induire la formation du polymorphe V. Pour vérifier cette hypothèse, nous avons doublé et quadruplé le volume de solvant utilisé dans les échantillons présentant un ratio molaire de 3.13% et 6.25%, respectivement (Figure 50). De ce fait, nous avons gardé le ratio molaire constant, tout en diminuant la concentration du lipide et de l'amylose dans la solution.

Il en résulte une formation accrue de polymorphe V pour ces deux échantillons (Figure 51). La solubilité intrinsèque du lipide utilisé est donc un facteur déterminant dans la formation de polymorphe V.

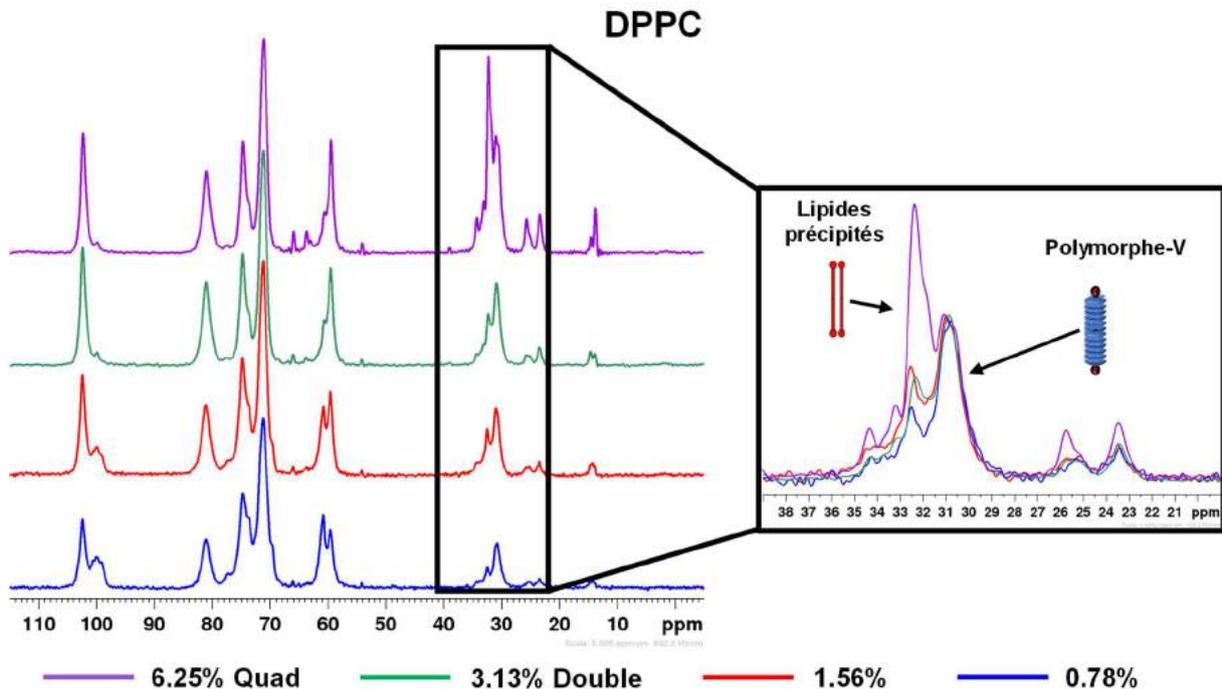


Figure 50 : Spectres RMN  $^{13}\text{C}$  d'échantillons d'amylose-DPPC dont les ratios molaires sont de 0.78% (bleu), 1.56% (rouge) et 3.13% (vert) et 6.25% (violet). Les volumes initiaux de solvant des échantillons à 3.13% et 6.25% ont été doublés et quadruplés respectivement. Droite : agrandissement de la partie du spectre contenant le signal RMN des carbones  $\text{CH}_2$  des lipides présents dans les échantillons.

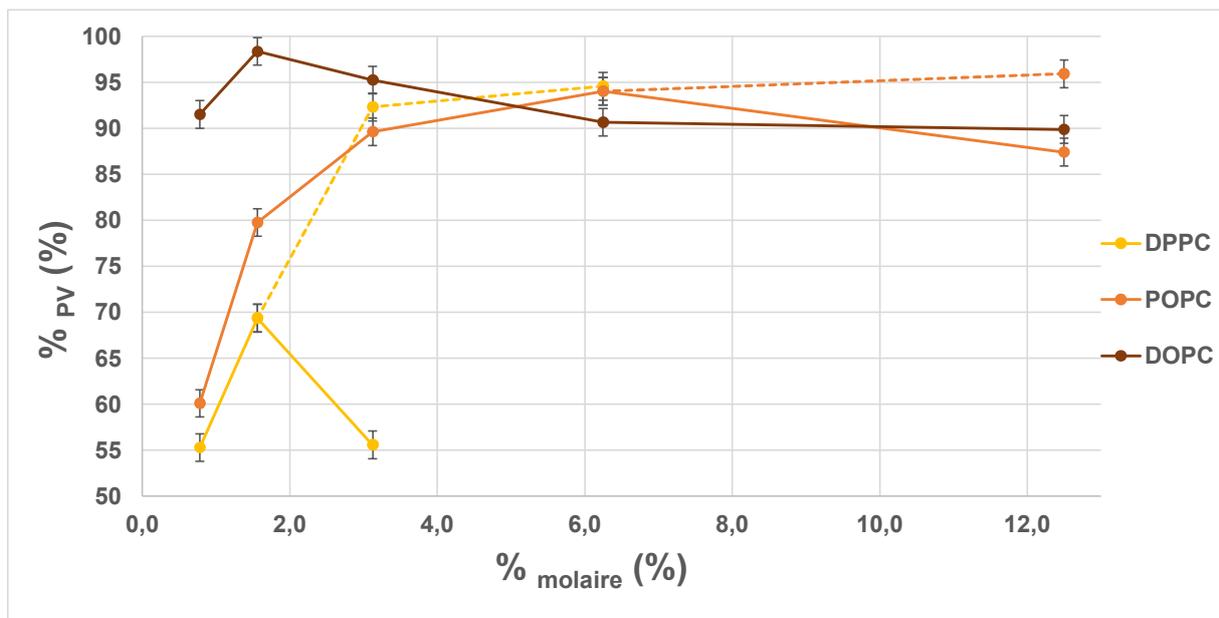


Figure 51 : Pourcentage de polymorphe V induit par les phospholipides utilisés en fonction du ratio molaire initial en lipide dans l'échantillon. Jaune : DPPC. Orange : POPC. Marron : DOPC. La courbe en pointillés jaunes représente ce pourcentage pour les échantillons de DPPC ayant un volume double (3.13%) et quadruple (6.25%). La courbe en pointillés orange représente l'échantillon de POPC ayant un volume quadruple (12.5%).

Concernant le POPC, ce lipide se dissout plus facilement dans le DMSO à 90°C à concentration molaire équivalente. Le protocole de fabrication des complexes avait de ce fait plus de chances de fonctionner, ce qui peut être observé sur les spectres RMN des complexes associés (Figure 52). De manière similaire aux acides gras, l'augmentation du ratio molaire initial dans la solution engendre une augmentation du pourcentage de polymorphe V observé, excepté pour l'échantillon présentant un ratio molaire de 12.5% (Figure 51). Pour ce dernier, afin de retrouver une augmentation de la quantité de polymorphe V, il a de nouveau été nécessaire de quadrupler le volume initial de solvant utilisé. Nous avons cependant observé des signaux RMN différents au niveau des CH<sub>2</sub> présents sur les chaînes lipidiques. Cette fois-ci, le signal qui leur est attribué quand les lipides sont précipités présente un déplacement chimique de 29.5 ppm. Il s'agit de POPC précipités entre eux et piégés dans l'amylose.

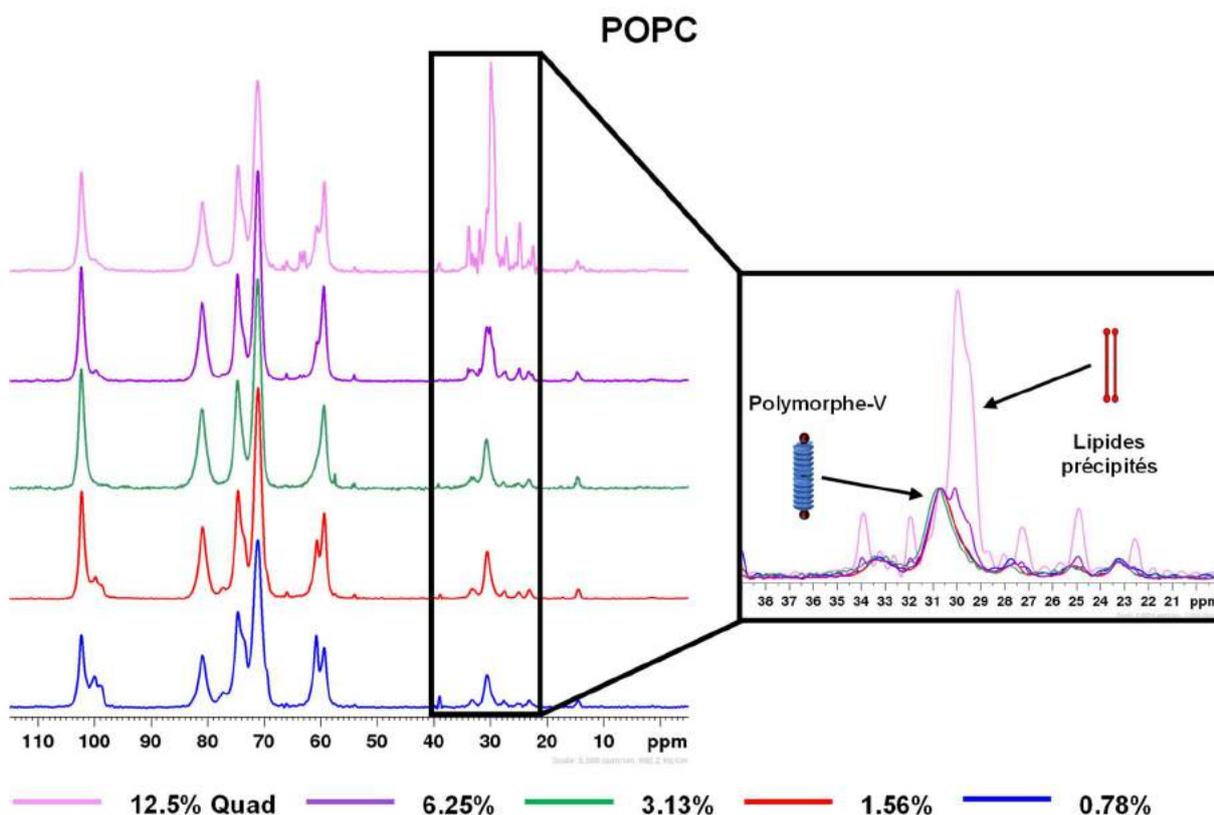


Figure 52: Spectres RMN d'échantillons d'amylose-POPC dont les ratios molaires sont de 0.78% (bleu), 1.56% (rouge), 3.13% (vert), 6.25% (violet) et 12.5% (rose). Droite : agrandissement de la partie du spectre contenant le signal RMN des carbones CH<sub>2</sub> des lipides présents dans les échantillons.

Ce comportement sera retrouvé dans les complexes suivant réalisés en présence de DOPC (Figure 53). Les échantillons contenant du DOPC présentent tous un pourcentage de polymorphe V supérieur à 90% dans le précipité obtenu à la fin de la complexation (Figure 51). L'augmentation de la quantité de lipide change relativement peu le %<sub>PV</sub> car l'amylose précipitée est déjà saturée en lipide. Les trois échantillons de plus haut pourcentage molaire présentent quant à eux de plus en plus de lipides agrégés, suggérant une compétition entre l'agrégation des lipides et la formation de polymorphe V.

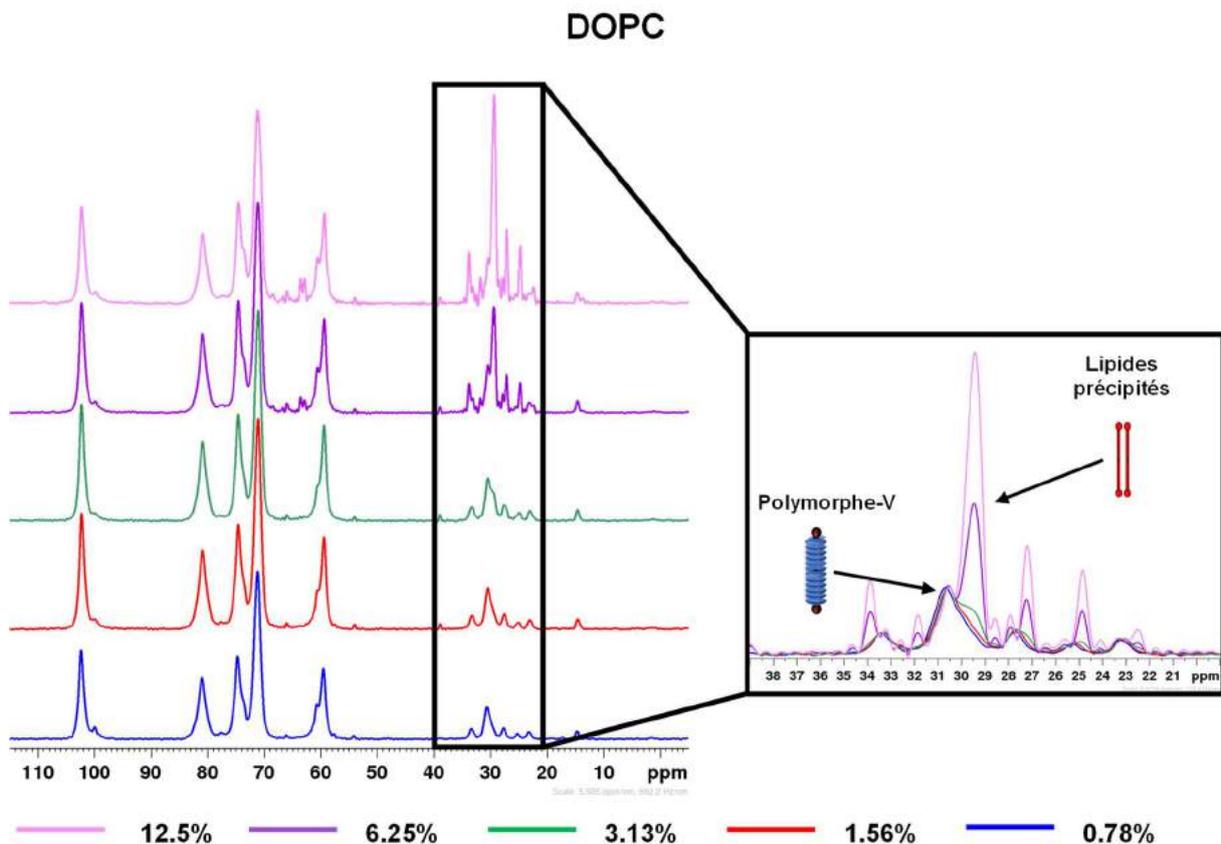


Figure 53: Spectres RMN <sup>13</sup>C d'échantillons d'amylose-DOPC dont les ratios molaires sont de 0.78% (bleu), 1.56% (rouge), 3.13% (vert), 6.25% (violet) et 12.5% (rose). Droite : agrandissement de la partie du spectre contenant le signal RMN des carbones CH<sub>2</sub> des lipides présents dans les échantillons.

La comparaison des pourcentages de polymorphe induit par le DPPC, le POPC et le DOPC semble indiquer que, à volume constant, plus on ajoute d'insaturations au sein des chaînes lipidiques, plus la quantité de polymorphe V formée est grande. La quantité maximum de polymorphe V %<sub>PVmax</sub> observable dans le précipité pour un volume constant dépend de la concentration du lipide utilisé, ainsi que de sa solubilité. Si la concentration en lipides est supérieure à celle nécessaire afin d'atteindre %<sub>PVmax</sub>, les lipides présents en surplus dans l'échantillon vont avoir tendance à s'agréger et

former un cluster prisonnier de l'amylose précipitée, comme en atteste le signal des CH<sub>2</sub> des chaînes acyles.

Ce mécanisme peut être formulé en d'autres termes : quel que soit le lipide utilisé, il existe une compétition entre son agrégation et la formation de polymorphe V au sein des échantillons (Figure 54).

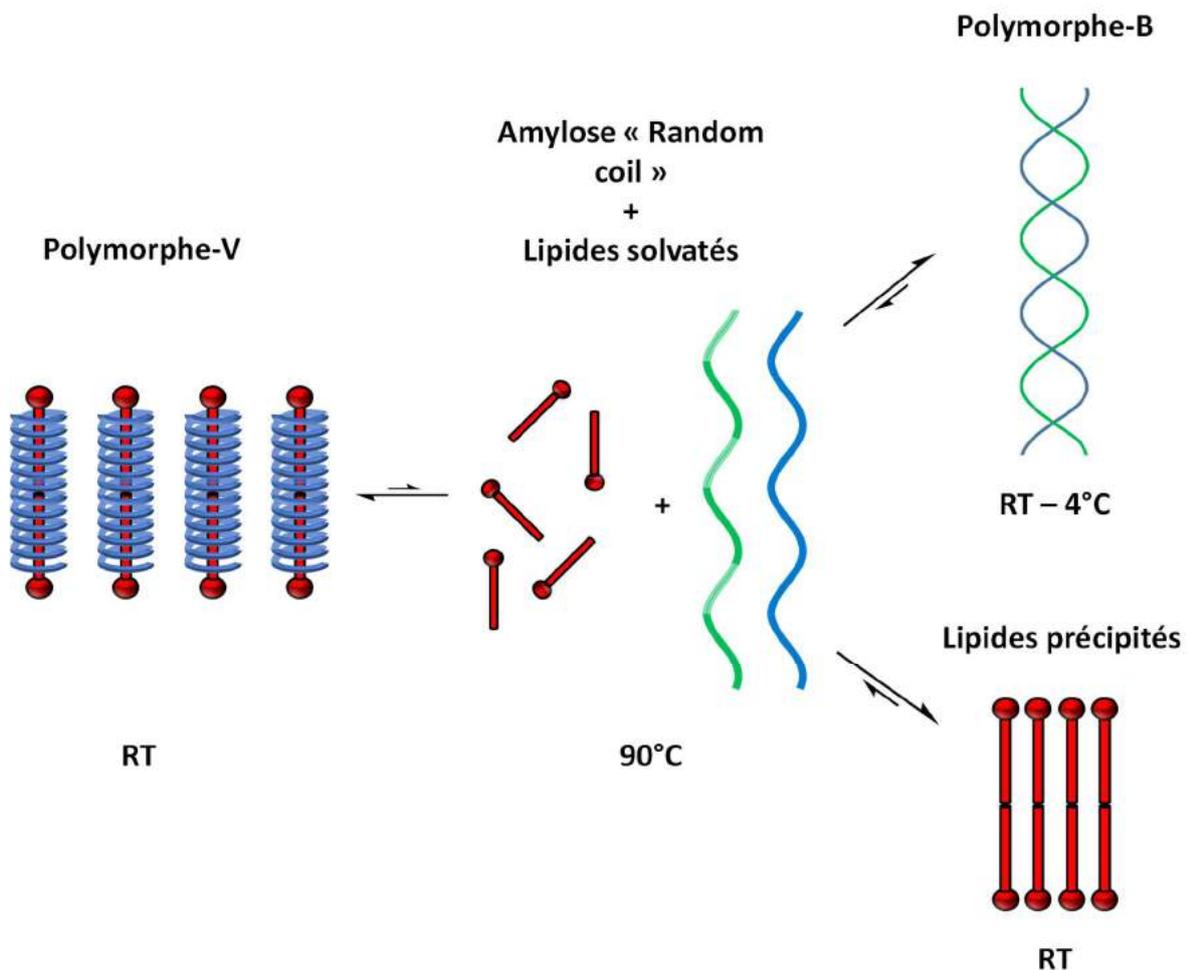


Figure 54: Mécanisme de compétition entre le mécanisme d'agrégation des lipides, la formation de polymorphes V et B.

Si le lipide utilisé présente une solubilité moindre, son agrégation va être privilégiée lors de la descente en température. De ce fait, moins de lipides seront disponibles pour engendrer la formation de polymorphe V dans l'échantillon. La rétrogradation de l'amylose engendrera la formation de polymorphe B, comme cela a déjà été démontré précédemment.

Dans le cadre des lipides que nous avons utilisés jusqu'ici, les acides gras présentent une très bonne solubilité dans le solvant utilisé. L'acide palmitique contient une chaîne

lipidique complètement saturée, rendant de ce fait les interactions inter-lipides plus fortes que dans le cas de l'acide oléique. Ceci se traduit aussi par la différence de température de fusion de ces lipides (Acide palmitique : 62°C ; Acide oléique : 13°C) qui rend l'acide oléique plus susceptible d'être retiré lors des lavages subséquents à la complexation.

Les phospholipides eux présentent une solubilité moindre qui engendre la présence de précipités ou d'agrégats dans l'échantillon en fonction de la concentration initiale en lipides lors du protocole de complexation. Le DPPC forme des précipités rapidement, et nécessite une concentration initiale plus faible afin de pouvoir être complexé par l'amylose. Les insaturations présentes dans le POPC et le DOPC induisent des interactions plus faibles entre lipides, et donc une capacité supérieure de ces lipides à former du polymorphe-V. L'étude de ces deux derniers phospholipides révèle qu'il existe une concentration initiale idéale, dépendante du lipide considéré, qui permet d'engendrer un maximum de polymorphe V au sein de l'amylose lors de sa rétrogradation.

## II. Résultats théoriques

Si le signal RMN  $^{13}\text{C}$  observé expérimentalement pour ces lipides est caractéristique de la formation du polymorphe V, leur diversité structurale ne permet pas d'affirmer qu'ils forment tous le même allomorphe. De plus, il était important de vérifier que les signaux RMN observés étaient bien dus à la formation de complexes d'inclusion. Afin d'obtenir de plus amples informations sur les caractéristiques structurales de ces complexes ainsi que sur leur processus de complexation, des trajectoires de dynamique moléculaire ont été réalisées sur des systèmes contenant un polymère d'amylose long de 45 résidus, déplié, en présence d'une molécule de PAL, OLE, DPPC, POPC ou DOPC.

Chaque simulation réalisée témoigne de la formation d'un complexe d'inclusion contenant le lipide (Figure 55).

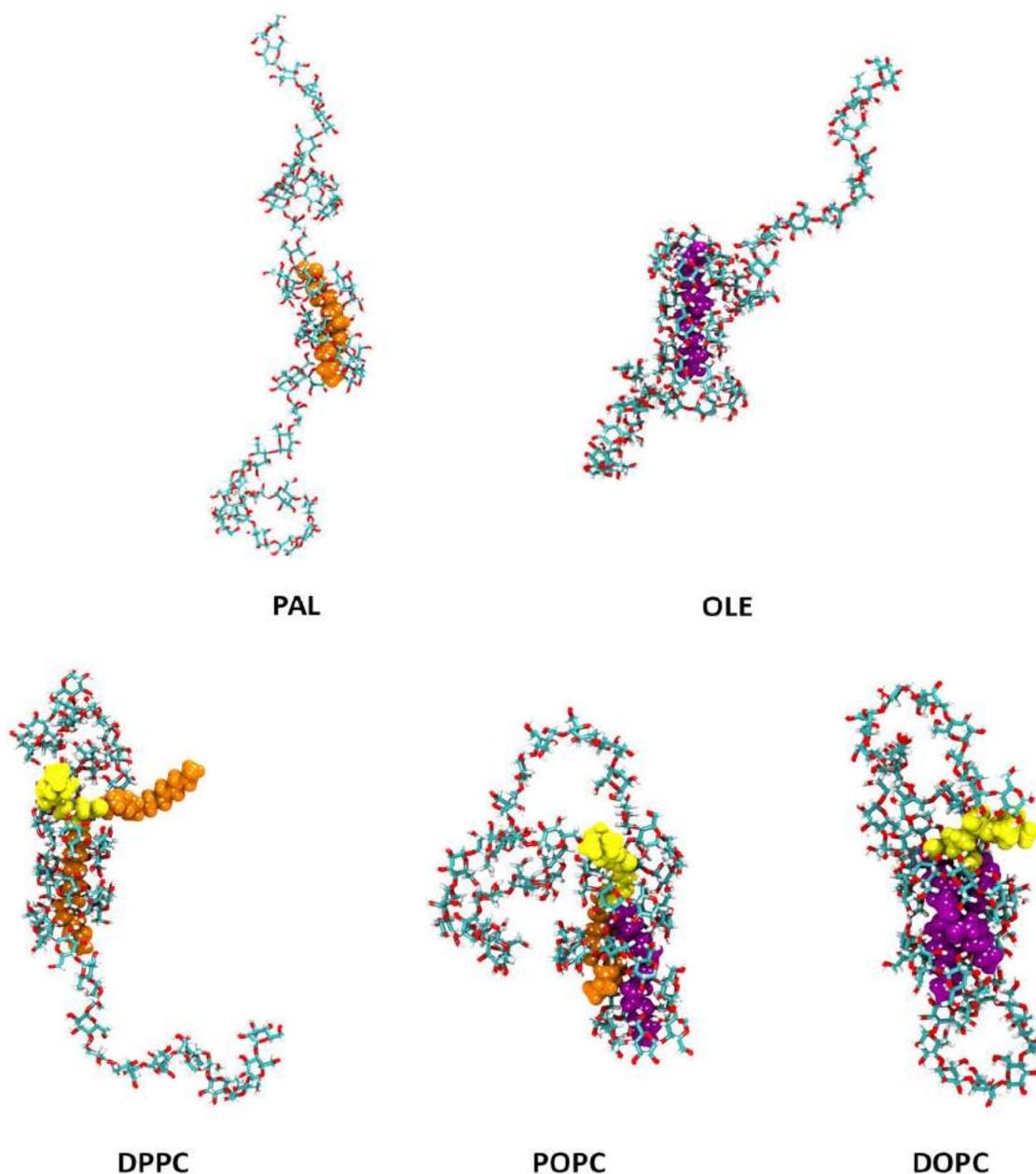


Figure 55: Structures sélectionnées après 100 ns de dynamique moléculaire contenant un polymère de 45 résidus en présence d'une molécule issue de la liste des lipides classiques. Jaune : Phosphatidylcholine. Orange : chaîne palmitoyl (C16 :0). Violet : chaîne oleoyl (C18 :1).

Les structures obtenues après complexation diffèrent cependant d'un lipide à l'autre. Les acides gras sont encapsulés dans une conformation étendue, ce qui est cohérent avec la littérature sur le sujet<sup>16-18</sup>. Concernant les phospholipides, les comportements diffèrent en fonction de la structure primaire des chaînes acyle du lipide. Pour le DPPC, les premières structures obtenues par dynamique moléculaire montrent un début d'encapsulation des deux chaînes hydrophobes (Figure 56,A – 20ns). Cependant, l'une des deux chaînes est finalement expulsée du polysaccharide qui forme

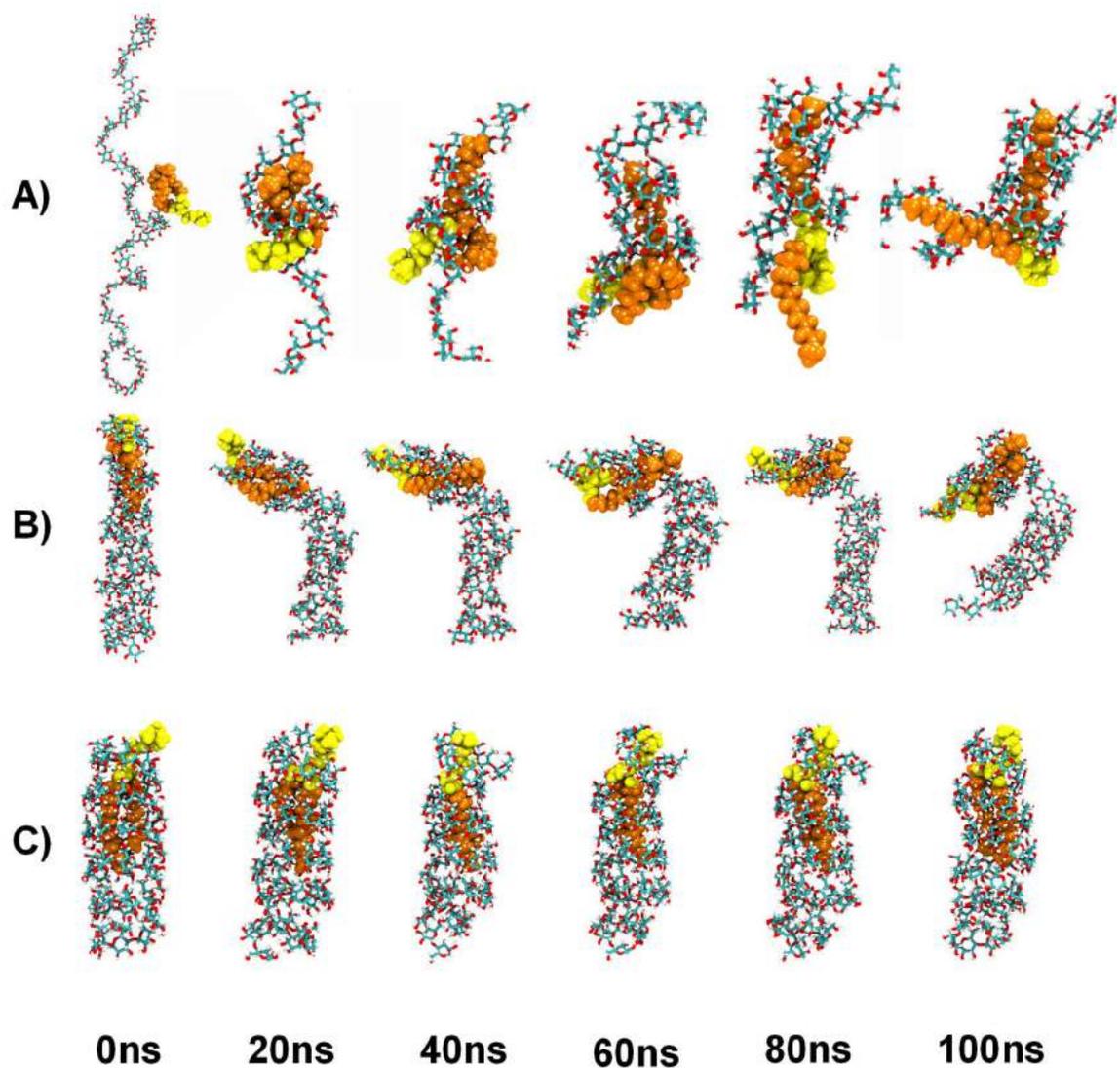


Figure 56: Structures extraites des trajectoires de calcul des systèmes contenant un polymère d'amylose de 45 résidus en présence d'une molécule de DPPC : Trajectoire de repliement (A), trajectoire avec comme point initial un polymorphe V6 (B) ou V8 (C).

alors un polymorphe V comprenant 6 résidus par tour autour d'une seule des deux chaînes acyles (Figure 56, A – 80-100ns). Afin de vérifier si l'encapsulation des deux chaînes était susceptible de former une conformation stable, nous avons réalisé deux calculs dont les structures initiales étaient formées d'une chaîne d'amylose repliée en polymorphe V6 (6 résidus glucose par tour d'hélice) et V8 (8 résidus glucose par tour d'hélice) autour d'une molécule de DPPC. On observe la déformation du polymorphe V6 autour du lipide, alors que le polymorphe V8 lui reste stable. Le polymorphe V6 ne présente pas une cavité hydrophobe suffisamment large pour pouvoir accueillir les deux chaînes lipidiques, ce qui entraîne une déformation du polymorphe (Figure 56, B). La chaîne d'amylose reste cependant repliée autour du lipide. Ce n'est pas le cas du polymorphe V8 qui lui reste stable au cours de la simulation (Figure 56, C). Il

semblerait de ce fait que l'encapsulation des deux chaînes soit envisageable, même si le processus de complexation que nous avons simulé engendre la formation d'un polymorphe replié autour d'une seule chaîne. A notre connaissance, Il s'agit du premier calcul de dynamique moléculaire caractérisant un tel complexe.

Les trajectoires de complexations réalisées en présence de DOPC et de POPC convergent vers des structures dans lesquels les deux chaînes lipidiques sont complexées. En effet, l'insaturation présente au milieu des chaînes oleoyl engendre une perte de degrés de libertés qui rend plus difficile l'expulsion de ces dernières de l'amylose. En contrepartie, le polymorphe formé ne peut plus être réellement assimilé à un allomorphe en particulier, mais plutôt à un mélange de différents allomorphes V7/V8 pour le POPC et V6/V7 pour le DOPC. Ceci est confirmé en observant la distribution des distances entre le centre de masse des résidus  $[n]$  et  $[n+X]$ , réalisée sur la totalité de la trajectoire (Figure 57). On observe la présence d'un pic de distribution proche de la valeur expérimentale pour les distances  $[n]$  et  $[n+7]$  et  $[n]$  et  $[n+8]$  dans le cadre du POPC et pour les distances  $[n]$  et  $[n+6]$  et  $[n]$  et  $[n+7]$  concernant le DOPC.

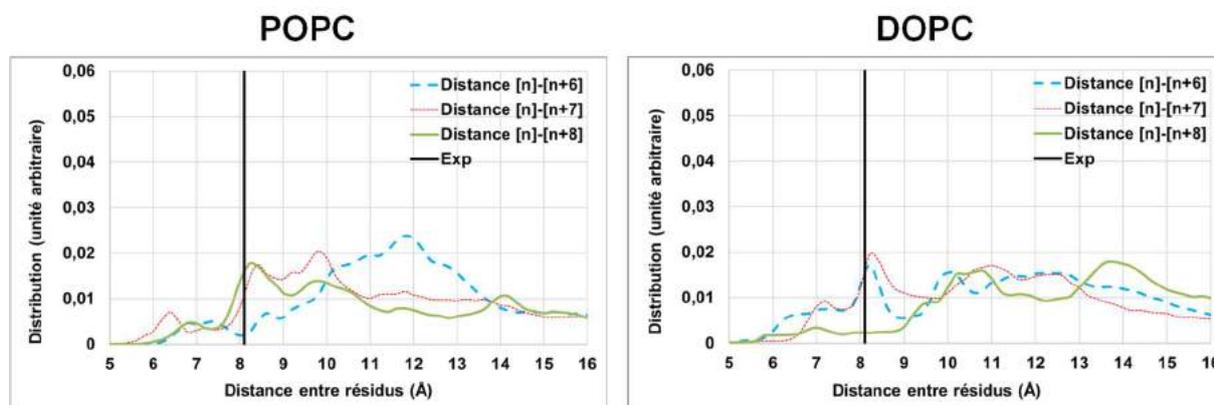


Figure 57: Distribution des distances  $[n]$ - $[n+6]$  (pointillés bleus),  $[n]$ - $[n+7]$  (tirets rouges),  $[n]$ - $[n+8]$  (trait vert) obtenus sur tous les résidus et échantillonnées sur les 100ns des simulations de repliement en présence d'une molécule de POPC et de DOPC.

Si les conformations observées à l'issue des trajectoires de repliement pour les acides gras et les phospholipides sont diverses, il était important de vérifier deux résultats de ces simulations : 1) les résidus situés autour des chaînes lipidiques forment bien une structure hélicoïdale et 2) cette structure est bien responsable du signal RMN observé.

Pour répondre à la première question, nous avons analysé les distributions des angles dièdres  $\phi$  (O5-C1-O4'-C4') et  $\psi$  (C1-O4'-C4'-C3') le long des chaînes d'amylose. Les

distributions calculées sur le long de la trajectoire des résidus situés dans les parties hélicoïdales présentent systématiquement un maximum global aux alentours de  $\phi=115^\circ$  et  $\psi=105^\circ$ . Nous avons ainsi pu rassembler toutes leurs distributions au sein de diagrammes ne concernant que les parties hélicoïdales (Figure 58).

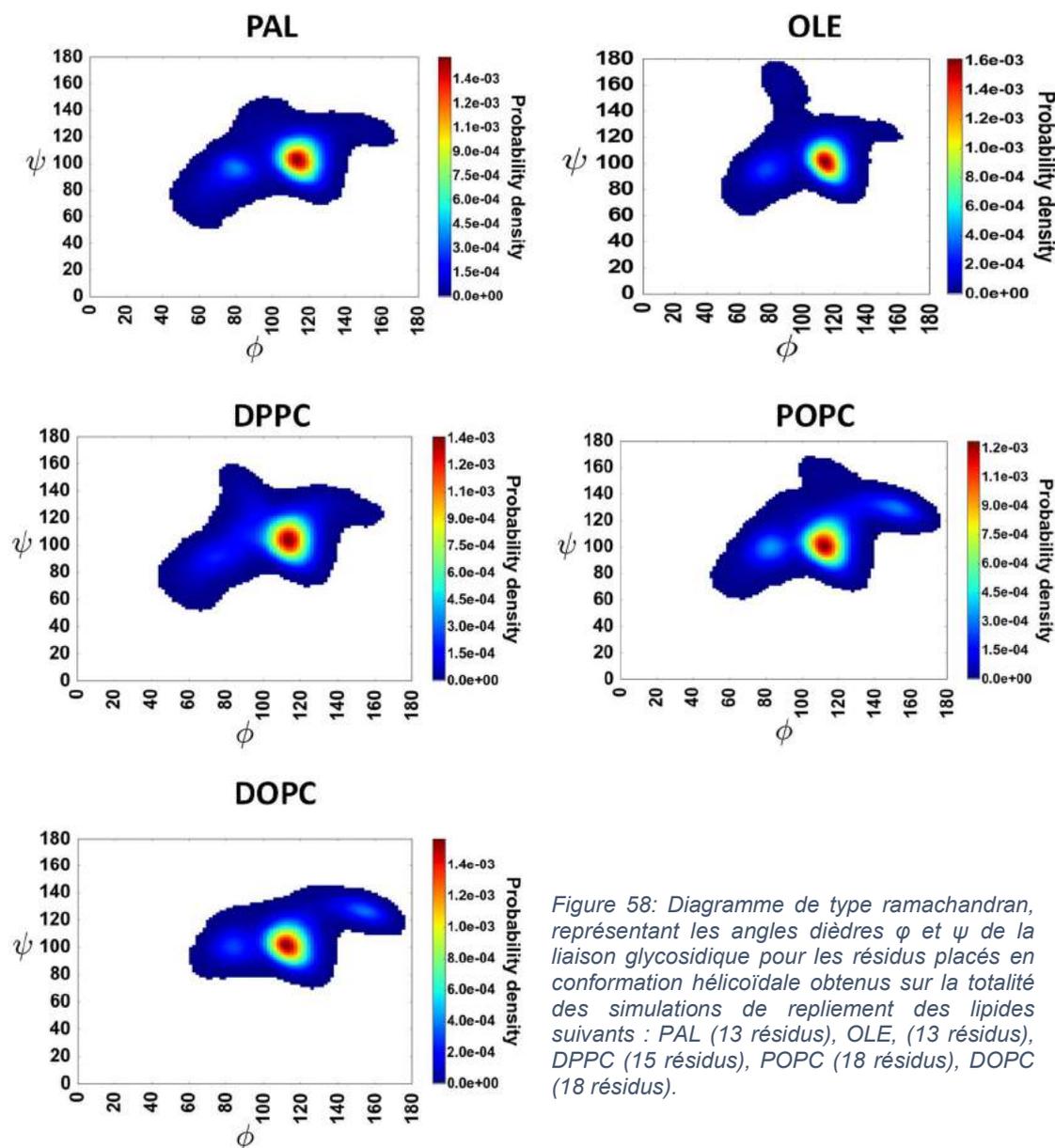


Figure 58: Diagramme de type ramachandran, représentant les angles dièdres  $\phi$  et  $\psi$  de la liaison glycosidique pour les résidus placés en conformation hélicoïdale obtenus sur la totalité des simulations de repliement des lipides suivants : PAL (13 résidus), OLE, (13 résidus), DPPC (15 résidus), POPC (18 résidus), DOPC (18 résidus).

De manière complémentaire, les autres résidus non complexés des chaînes d'amylose présentent un maximum global se situant aux alentours de  $\phi=75^\circ$  et  $\psi=95^\circ$ , et un second maximum d'intensité plus faible peut être aussi observé aux alentours de  $\phi=110^\circ$  et  $\psi=100^\circ$  (Figure 59). Le premier maximum est induit par des résidus présents dans une conformation étendue de l'amylose, et le second par des phases hélicoïdales transitoires en solution. Un comportement similaire a été récemment décrit dans une

étude comparant les différents champs de force utilisables pour la modélisation des polysaccharides<sup>19</sup>. La complexation autour d'une partie lipidique implique donc une forte diminution de la contribution présente à  $\phi=75^\circ$  et  $\psi=95^\circ$ .

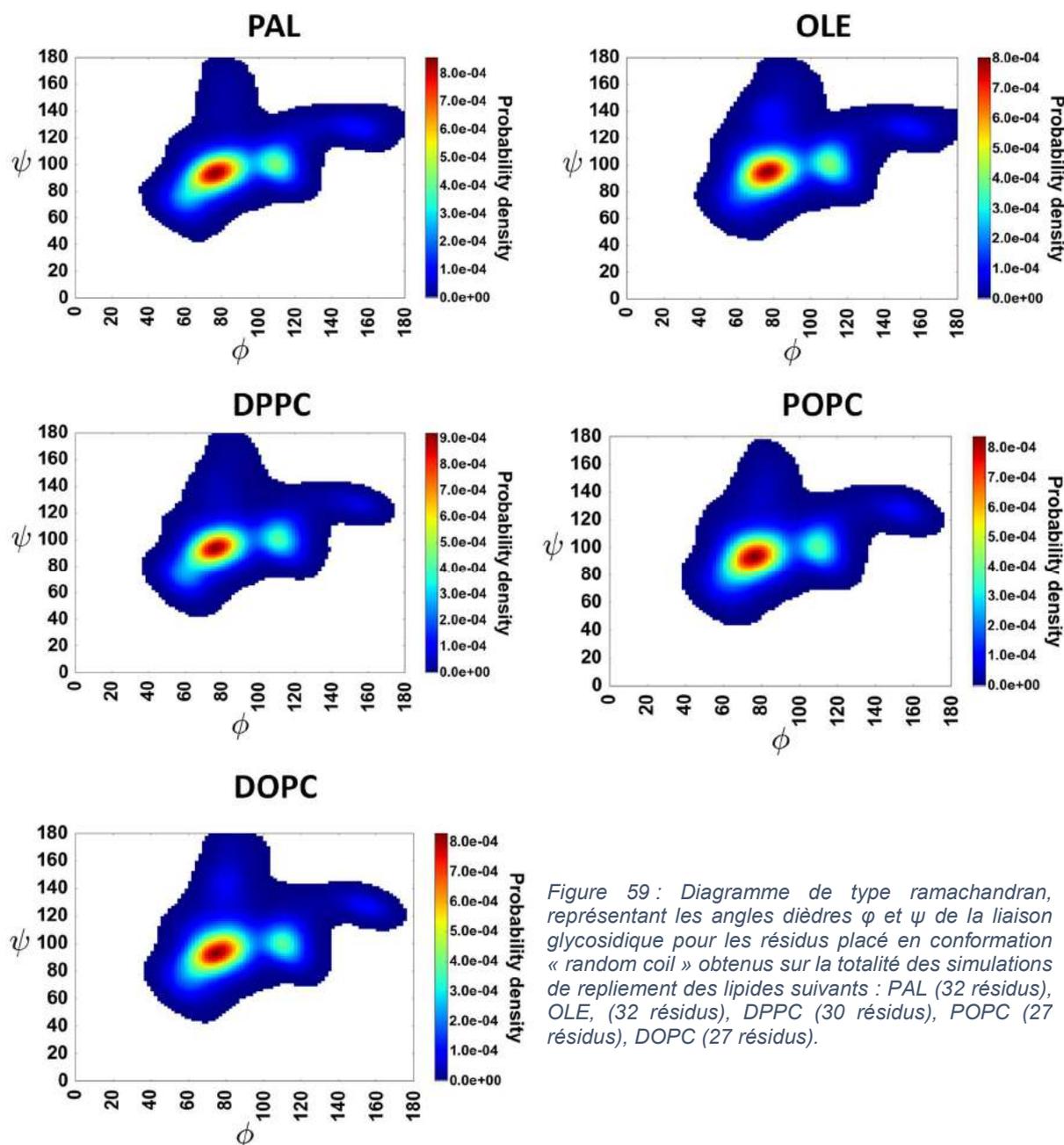


Figure 59 : Diagramme de type ramachandran, représentant les angles dièdres  $\phi$  et  $\psi$  de la liaison glycosidique pour les résidus placés en conformation « random coil » obtenus sur la totalité des simulations de repliement des lipides suivants : PAL (32 résidus), OLE, (32 résidus), DPPC (30 résidus), POPC (27 résidus), DOPC (27 résidus).

Il a été prouvé que la modification de ces angles dièdres présentait un impact non négligeable sur le calcul de paramètres RMN des  $C_1$  et  $C_4$  de disaccharides<sup>20</sup>. En partant de ce constat, nous avons extrait cinquante structures issues des dix dernières ns de chaque trajectoire afin de réaliser des calculs de RMN au niveau DFT. Par la suite, de manière similaire à la stratégie décrite dans le chapitre précédent, nous avons moyenné les cinquante échantillages chimiques calculés pour chaque carbone. Afin de

vérifier que ces résidus sont bien en conformation hélicoïdale, les diagrammes de type Ramachandran ont été réalisés pour chaque résidu sur les dix dernières ns de chaque simulation, plages correspondant à celles dont on extrait les structures utilisées pour le calcul de RMN. Ces diagrammes ne sont pas présentés ici car ils sont extrêmement similaires à ceux représentés sur la Figure 58 et la Figure 59.

Les sucres ont ensuite été séparés en deux groupes : 1) ceux dont le diagramme présente un maximum aux alentours  $\varphi=115^\circ$  et  $\psi=105^\circ$  qui forment la partie hélicoïdale et 2) ceux qui présentent un maximum caractérisé par d'autres angles, qui forment la partie « random coil ». Pour chaque type de carbone, C<sub>1</sub> à C<sub>6</sub>, nous avons calculé la moyenne d'écrantage chimique de tous les sucres compris dans chaque groupe. Nous avons donc obtenu un écrantage chimique moyen pour chaque type de carbone au sein de la partie hélicoïdale et au sein de la partie « random coil »

Afin de pouvoir comparer ces résultats aux résultats expérimentaux, nous avons simulé des spectres de déplacements chimiques <sup>13</sup>C à partir des écrantages chimiques calculés, tel que décrits dans le chapitre précédent (Figure 60). Pour chaque simulation et quel que soit le lipide considéré, les sucres présents dans la partie hélicoïdale présentent en moyenne un signal du C<sub>1</sub> supérieur de 2 ppm à ceux présents au sein du « random coil », et un signal du C<sub>4</sub> qui se situe aux alentours de 80 ppm. Ces résultats sont cohérents avec nos résultats expérimentaux, dans lesquels les signaux des C<sub>1</sub> et C<sub>4</sub> sont présents à 102.6 ppm et 82 ppm respectivement. Le déplacement moyen des carbones C<sub>2</sub> C<sub>3</sub> et C<sub>5</sub> se situe aux alentours de 72-73 ppm ce qui est cohérent avec les résultats obtenus dans le chapitre précédent. Concernant les résidus présents dans la partie « random coil », le pic présent à 76 ppm est dû au C<sub>4</sub>, et les C<sub>2</sub>, C<sub>3</sub> et C<sub>5</sub> sont présents entre 70 et 75 ppm. Nous pouvons observer des différences entre les déplacements chimiques de ces trois carbones d'une simulation à une autre ce qui laisse penser qu'un nombre plus important de structures devrait être considéré lors du calcul de paramètres RMN afin de pouvoir observer la convergence des déplacements chimiques calculés vers les valeurs moyennes similaires. Finalement les déplacements chimiques des carbones C<sub>6</sub> présentent un déplacement chimique aux alentours de 59 ppm dans les deux parties du polymère. Nous avons vu lors du premier chapitre que la présence de deux liaisons hydrogène sur le groupement

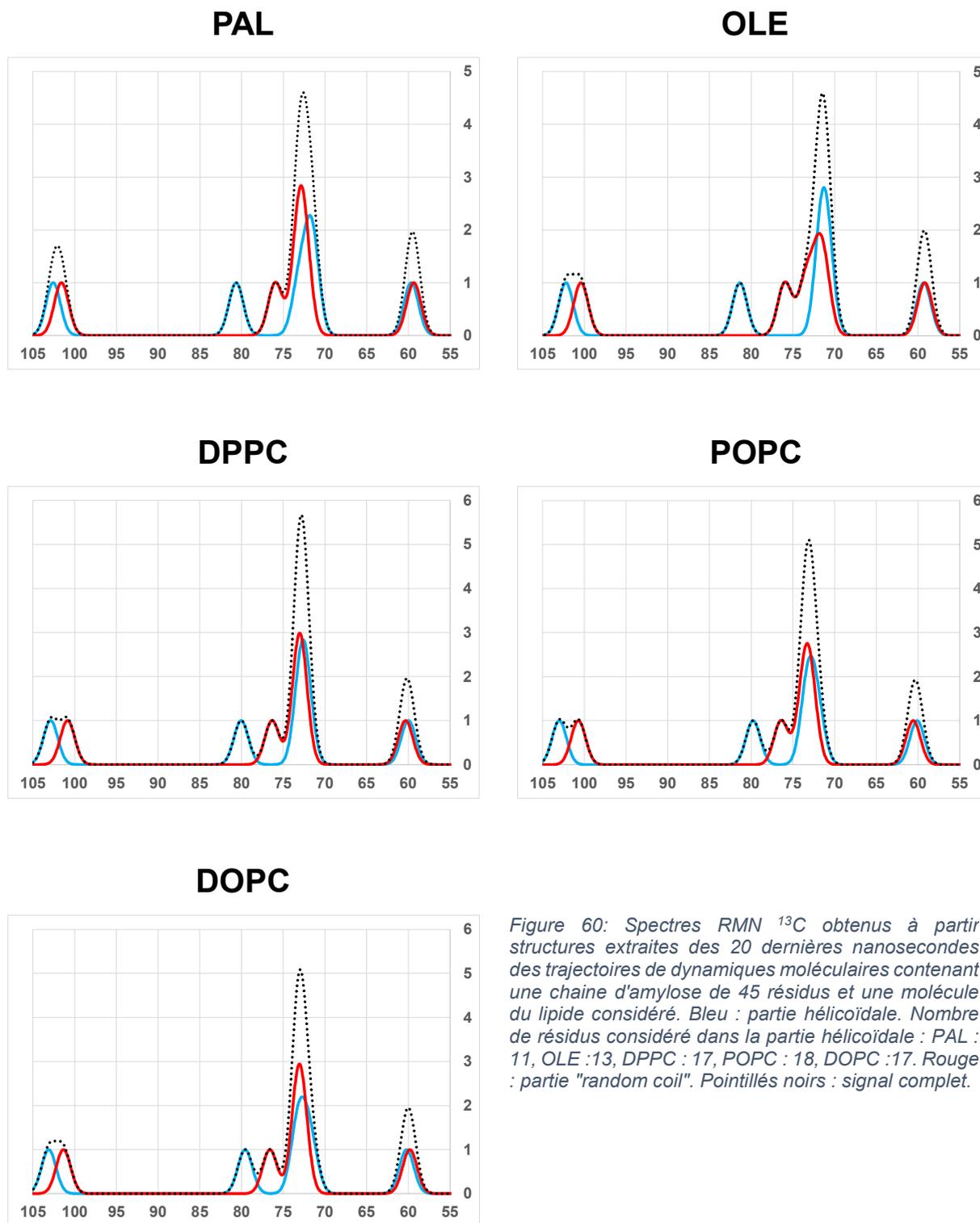


Figure 60: Spectres RMN <sup>13</sup>C obtenus à partir structures extraites des 20 dernières nanosecondes des trajectoires de dynamiques moléculaires contenant une chaîne d'amylose de 45 résidus et une molécule du lipide considéré. Bleu : partie hélicoïdale. Nombre de résidus considéré dans la partie hélicoïdale : PAL : 11, OLE :13, DPPC : 17, POPC : 18, DOPC :17. Rouge : partie "random coil". Pointillés noirs : signal complet.

hydroxyle du C<sub>6</sub> engendrait une diminution du déplacement chimique de ce dernier, de 61 à 59 ppm. Ces liaisons ne sont pas observées systématiquement dans nos structures, et de ce fait notre méthode ne permet pas de reproduire la différence de déplacement chimique des C<sub>6</sub> entre la partie hélicoïdale et la partie random-coil. Il s'agit donc ici d'une des limites de notre modèle chimique. Il est cependant raisonnable

de penser que la formation d'un ou plusieurs allomorphes présente une probabilité plus élevée de générer ces liaisons hydrogène au sein du polymorphe.

### III. Résumé des résultats obtenus

Nos résultats expérimentaux mettent en évidence l'importance de la concentration initiale des lipides utilisés, ainsi que le lien avec leur structure primaire. Il faut une quantité de lipides suffisante pour pouvoir observer la complexation par RMN. De plus, la concentration initiale en lipide doit être ajustée afin d'éviter la précipitation ou l'agrégation des lipides entre eux et ainsi les rendre « disponibles » à la complexation et promouvoir la formation du polymorphe V. La concentration optimale pour générer ces structures varie selon la structure primaire du lipide. Des lipides contenant des insaturations en milieu de chaîne seront plus à même de former des complexes d'inclusion, car les interactions lipides-lipides seront défavorisées. En contrepartie, les lipides contenant des chaînes complètement saturées seront plus difficiles à complexer. La présence de deux chaînes lipidiques augmente la surface hydrophobe accessible et devrait en soit engendrer plus de polymorphe V, comme on peut l'observer dans le cadre du POPC et du DOPC. Cependant, l'ajout d'une chaîne complètement saturée engendre plus d'interactions entre lipides et désavantage donc la formation de ces complexes (POPC vs DPPC).

Nos résultats théoriques mettent en évidence un effet particulier impliqué par l'ajout d'une chaîne lipidique : le polymorphe V formé en présence de tels lipides n'est pas une structure homogène. Cette information est d'importance car elle permet d'expliquer pourquoi les méthodes d'études thermiques et rhéologiques et les méthodes de diffractions présentent des difficultés à caractériser de tels complexes. Si les complexes d'inclusion formés ne sont pas assez homogènes, la cristallisation ne pourra pas avoir lieu, ou alors de manière incomplète. Les méthodes de diffraction ne seront donc pas adaptées à l'étude de ces échantillons. Les méthodes thermiques et rhéologiques nécessitent des échantillons présentant aussi une homogénéité structurale afin de pouvoir être utilisées. Nos résultats prouvent que le signal  $^{13}\text{C}$  RMN observé sera toujours similaire en cas de présence de complexes d'inclusion dans l'échantillon, quel que soit l'allomorphe considéré. De ce fait, la RMN du solide du  $^{13}\text{C}$

permettant de caractériser pleinement des échantillons hétérogènes, elle est donc une méthode de choix pour l'étude de ce type de complexes en présence de lipides présentant de multiples chaînes.

L'ensemble de nos résultats ont permis de mettre en évidence un mécanisme de formation des complexes, et de démontrer que les structures de nos échantillons peuvent être hétérogènes tout en générant un signal  $^{13}\text{C}$  RMN caractéristique de complexes d'inclusion.

### 3. Utilisation de lipides extraits de mycobactéries et corynébactéries.

Les acides mycoliques, MMTs, DMTs et DIMs sont des lipides qui sont retrouvés dans la paroi cellulaire et dans la capsule des mycobactéries. Nous souhaitons savoir si ces différents lipides sont susceptibles de provoquer la formation de polymorphe V au sein de l'amylose. Les DIMs étudiés ici sont ceux issus de MTB, qui ont déjà été étudiés expérimentalement et théoriquement au sein de l'institut<sup>9</sup>. Les acides mycoliques et leurs dérivés présentent une structure primaire différente selon la bactérie considérée et ceux issus de MTB sont composés de chaînes acyles contenant 60 à 90 carbones et pouvant contenir des groupes oxygénés ou des cyclopropanes<sup>21</sup>. La mise en culture et purification de ces acides présentant des contraintes expérimentales non-négligeables, nous avons décidé d'étudier leurs analogues issus de *Corynebacterium Glutamicum*. Ces derniers présentent aussi des chaînes hydrophobes similaires à celles présentes dans les acides gras et phospholipides que nous avons étudiés ci-dessus (contenant au total 32 à 36 carbones) et de ce fait ont permis de mettre à l'épreuve le mécanisme de compétition que nous avons décrit. Nous disposons aussi d'un analogue de synthèse des acides mycoliques, le BET, qui nous a permis de réaliser des complexes à différentes concentrations. Nous avons donc purifié des lipides issus de corynébactéries et réalisé la synthèse de complexes en utilisant le même protocole que décrit précédemment.

## I. Résultats expérimentaux

Le premier lipide auquel nous nous sommes intéressés est le Méthyl 3-oxo-2-tetradecyloctadecanoate. Ce lipide est un analogue de synthèse des acides mycoliques, présentant une fonction cétone en position  $\beta$  de la fonction ester, qu'on appelle aussi acide bêta-céto-mycolique (BET). Nous disposons donc de grandes quantités de ce lipide permettant de mettre à l'épreuve le modèle développé dans la partie précédente. De ce fait, des échantillons présentant des ratios molaires de 0.78%, 1.56%, 3.13% et 6.25% ont été préparés et nous avons enregistré leurs spectres RMN  $^{13}\text{C}$  (Figure 61).

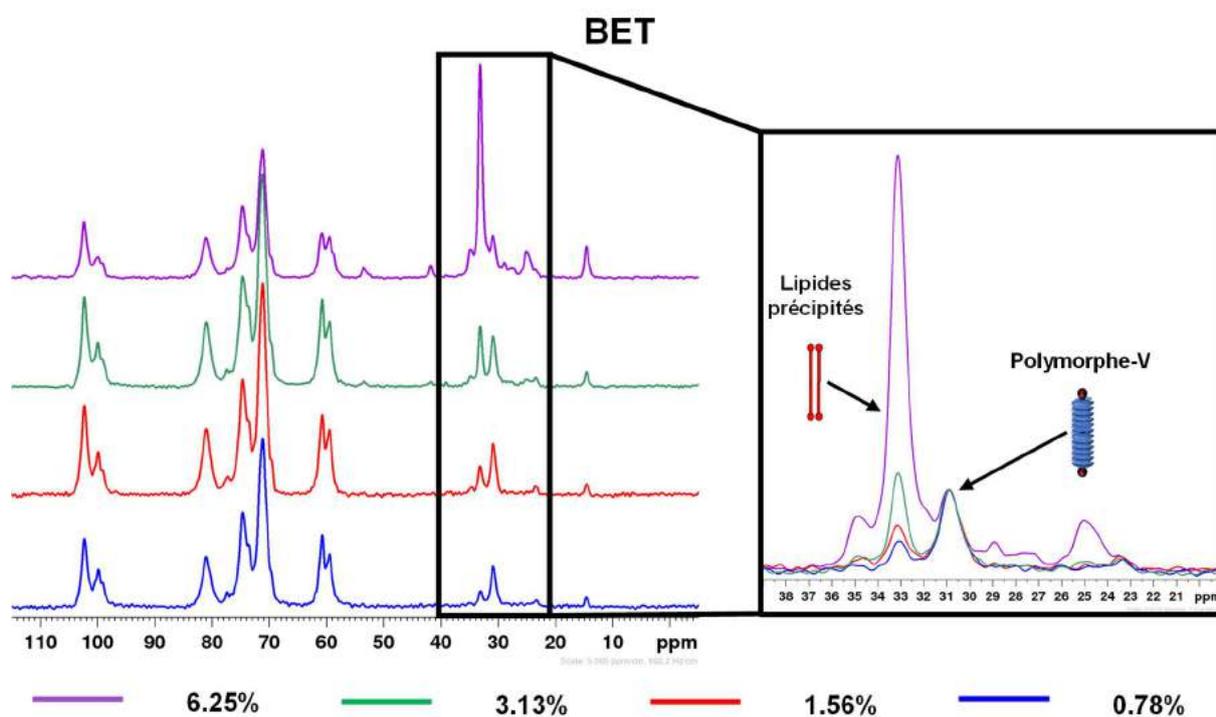


Figure 61: Spectres RMN  $^{13}\text{C}$  d'échantillons d'amylose-BET dont les ratios molaires sont de 0.78% (bleu), 1.56% (rouge) et 3.13% (vert) et 6.25% (violet). Droite : agrandissement de la partie du spectre contenant le signal RMN des carbones  $\text{CH}_2$  des lipides présents dans les échantillons.

Ces derniers témoignent de la présence de polymorphe V, comme en atteste la présence des pics localisés à 102.6 ppm et 81 ppm. De manière analogue au DPPC, l'utilisation du BET engendre des précipités lipidiques quel que soit le ratio molaire utilisé initialement. Le pourcentage de polymorphe V engendré par l'augmentation du ratio molaire du BET présente un comportement stationnaire, les valeurs obtenues étant, dans l'ordre croissant des ratios molaires utilisés, de 62%, 66%, 63% et 70% +/-

3%, respectivement. Ce lipide étant complètement apolaire et présentant en plus deux chaînes complètement saturées, le pourcentage maximum de polymorphe qu'il peut générer dans l'amylose est inférieur à celui des acides gras et des phospholipides, à l'exception du DPPC. Il nous servira de référence par rapport à laquelle nous comparerons les lipides suivants.

Nous avons par la suite réalisé un échantillon contenant des acides mycoliques (MYC), des monomycolates de tréhalose (MMT) ou des dimycolates de tréhalose (DMT) en utilisant un ratio molaire de 6.25% afin de comparer l'impact des modifications structurales d'un lipide à un autre. Les spectres RMN résultant montrent la présence de polymorphe V (cf. déplacements chimiques des carbones C<sub>1</sub> et C<sub>4</sub>), de lipides encapsulés et de lipides agrégés entre eux (Figure 62).

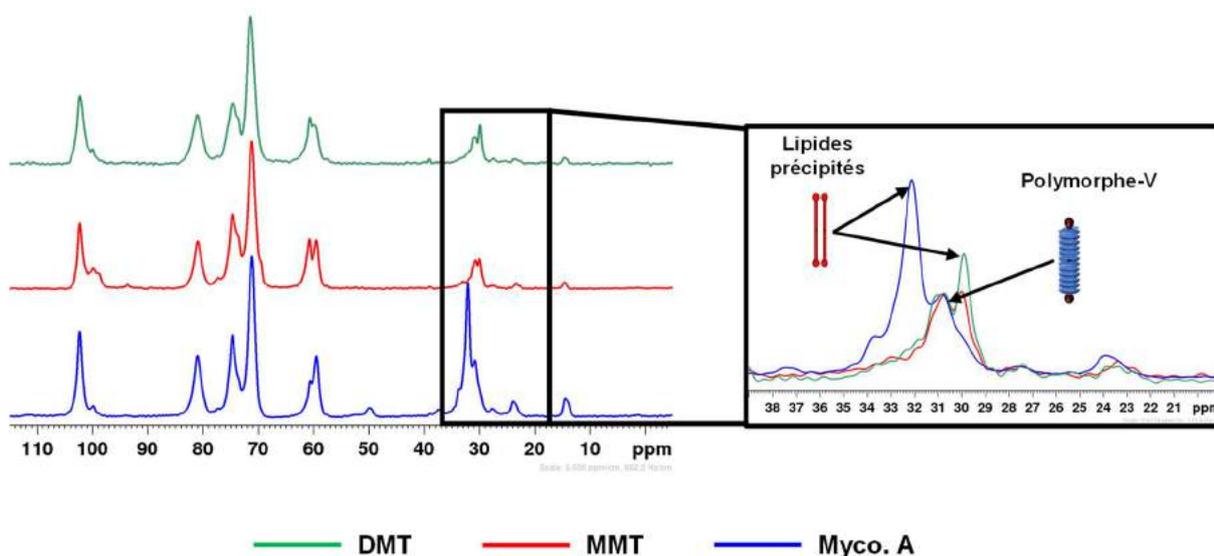


Figure 62: Spectres RMN <sup>13</sup>C d'échantillons d'amylose en présence d'acide mycoliques (bleu), de MMT (rouge) et de DMT (vert) dont les ratios molaires sont de 6.25%. Droite : agrandissement de la partie du spectre contenant le signal RMN des carbones CH<sub>2</sub> des lipides présents dans les échantillons.

Pour les acides mycoliques, le signal des lipides agrégés se situe à 32 ppm tandis que dans le cas des MMT et des DMT, il se situe à 30 ppm. Ceci laisse penser que la plupart des acides mycoliques non complexés présentent une majorité de chaînes saturées, là où les MMTs et DMTs agrégés contiennent aussi des chaînes insaturées.

Le pourcentage de polymorphe V induit par chaque type de lipide diffère en fonction de leur structure primaire (Figure 63).

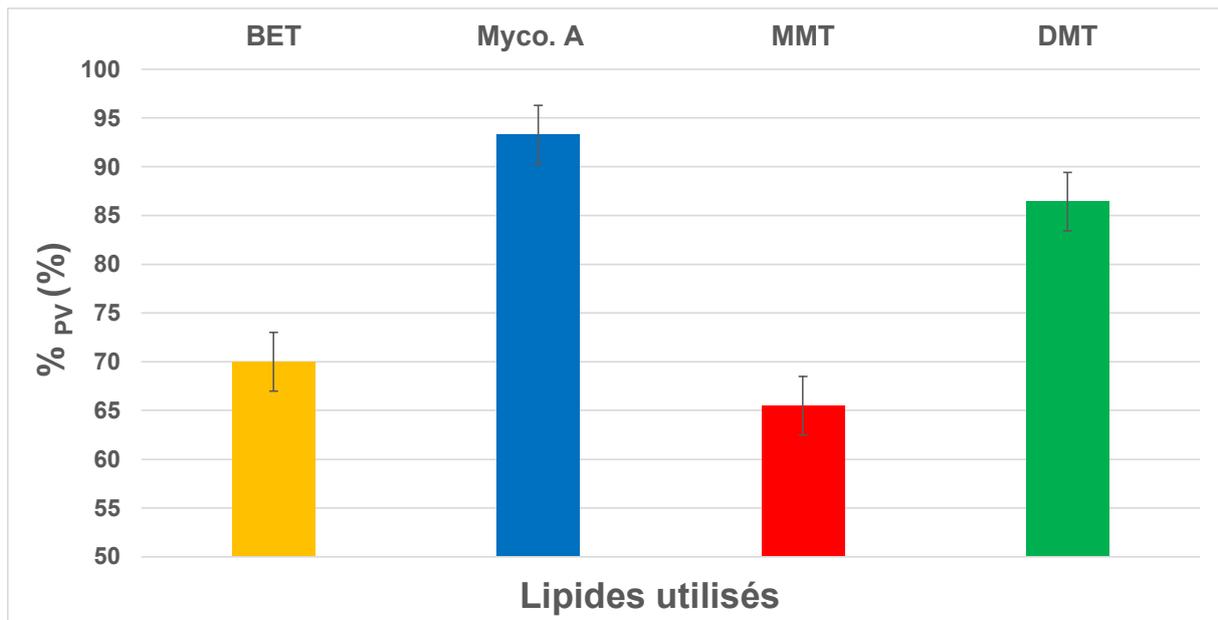


Figure 63: Pourcentage de polymorphe V induit par les acides bêta-céto-mycoliques, les acides mycoliques, les monomycolates de tréhalose et les dimycolates de tréhalose en utilisant un ratio molaire initial de 6.25% dans chaque échantillon.

Pour les acides mycoliques, il existe deux modifications structurales particulières, en comparaison avec le BET (Figure 41) : 1) la tête polaire du lipide présente une fonction acide et un groupe hydroxyle en bêta de cette dernière, et 2) l'échantillon étant purifié à partir d'une culture de corynébactéries (souche *C. Glutamicum* ATC13032), il correspond à un mélange d'acides mycoliques avec, majoritairement, des chaînes acyles en C32:0 (32 carbones et 0 insaturation) et C34:1 (34 carbones et 1 insaturation) ainsi qu'une faible proportion de chaînes en C36:2 (34 carbones et 2 insaturations) telles que décrites dans la Figure 41. Comme pour l'acide oléique, la présence de chaînes lipidiques contenant des insaturations promeut la complexation et la formation de polymorphe V, qui représente 93.3% +/-3% de la structure de l'amylose dans l'échantillon au ratio molaire Lipide/Amylose de 6.25%.

Les MMTs, purifiés à partir de culture de la même souche de corynébactéries, présentent aussi une distribution de chaînes aliphatiques similaires à celle des acides mycoliques. La tête polaire étant cette fois-ci un groupe tréhalose branché en position 6, il s'agit de la seule partie du lipide pouvant expliquer la baisse du pourcentage de polymorphe V dans l'échantillon après précipitation. Cette tête étant un disaccharide, des interactions spécifiques sont envisageables avec le polysaccharide, déstabilisant de ce fait le polymorphe V. L'impact de l'ajout du groupe tréhalose sera discuté dans la partie théorique concernant les lipides bactériens.

Enfin les DMT présentent une structure identique aux MMT, avec l'ajout d'une chaîne mycolique supplémentaire en position 6' du tréhalose. Cette chaîne mycolique peut de nouveau être de la forme C32:0, C34:1 ou C36:2. Si différentes combinaisons de chaînes mycoliques peuvent alors être envisagées, les trois structures présentées dans la Figure 41 sont celles qui sont le plus couramment rencontrées<sup>8</sup>. En comparaison avec les MMT, on observe une augmentation du pourcentage de polymorphe V dans l'échantillon contenant des DMT. Le doublement du nombre de chaînes, ainsi que le fait qu'elles contiennent des insaturations, peuvent expliquer cette augmentation de manière intuitive. Si on compare cet échantillon et l'échantillon Myco. A., la présence de la tête tréhalose tend tout de même à déstabiliser le polymorphe V.

Les résultats obtenus sur les lipides bactériens sont cohérents avec ceux observés sur les lipides classiques. En passant d'un type de lipide ne contenant aucune insaturation (BET) à un mélange de divers lipides contenant zéro à deux insaturations (Myco. A.), la formation du polymorphe V est favorisée. L'ajout d'un groupement tréhalose branché en 6 (MMT) implique la formation de liaisons hydrogènes avec l'amylose qui sont délétères pour la formation de complexes d'inclusions. Nous reparlerons de cet effet dans la partie théorique discutant de ce lipide. Enfin l'ajout de deux chaînes supplémentaires en position 6' du groupe tréhalose (DMT) permet de favoriser de nouveau la formation du polymorphe, sans pour autant arriver au taux d'inclusion présent dans l'échantillon MYC.

Nous avons étendu l'étude de complexation avec l'amylose aux dimycocérosate de phtiocérol (Figure 41 ,DIM), lipides que l'on ne retrouve que chez les souches pathogènes de *Mycobacterium tuberculosis* et qui sont étudiés à l'IPBS<sup>9</sup>. Ce type de lipides présente une structure relativement différente des lipides étudiés jusqu'alors : il contient des chaînes hydrophobes plus longues (20 à 30 carbones) qui, pour deux entre elles portent 3 à 5 groupements méthyles. Nous avons réalisé deux échantillons à des ratios molaires de 1.8% et 3.7% visant à étudier la possible formation de complexes d'inclusion. Les spectres RMN <sup>13</sup>C de ces échantillons démontrent la possibilité de formation de polymorphe V en présence de ces lipides (Figure 64). En effet, on y observe la présence des pics caractéristiques du C<sub>1</sub> et du C<sub>4</sub> localisés respectivement à 102.6 et 82 ppm. Dans les deux échantillons, aucun lipide précipité, piégé dans l'amylose, n'est observé, le seul signal des CH<sub>2</sub> des chaînes lipidiques

étant localisé à 32 ppm. Ces deux échantillons génèrent respectivement 41 et 77% +/- 3% de polymorphe V.

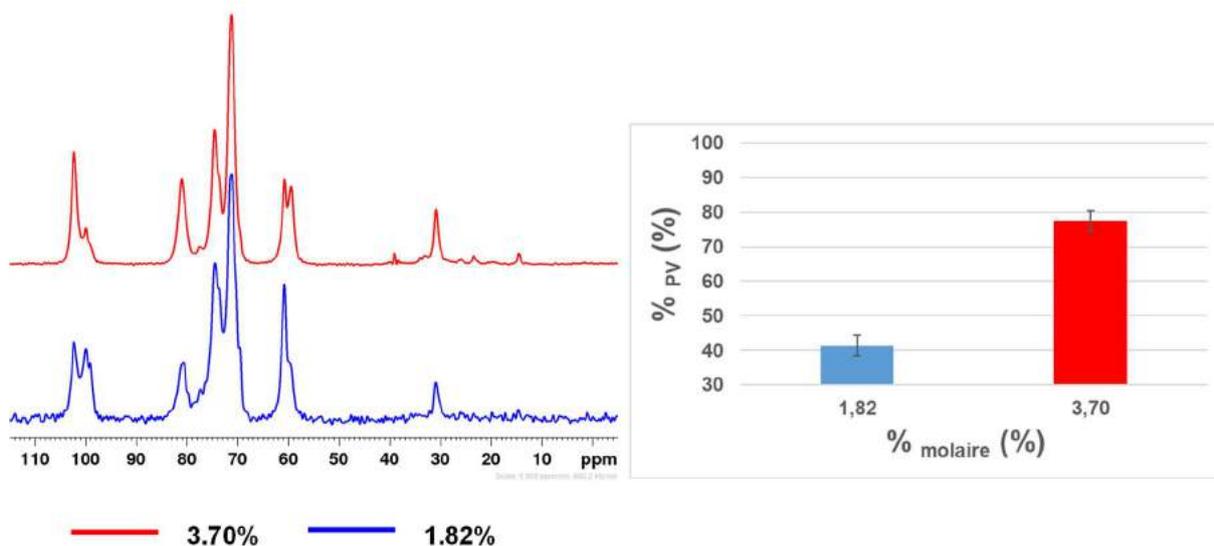


Figure 64: Spectres RMN <sup>13</sup>C d'échantillons d'amylose-DIM dont les ratios molaires sont de 1.82% (bleu) et 1.56% (rouge). Droite : Pourcentage de polymorphe V induit par les DIM en fonction du ratio molaire initial en lipide dans l'échantillon.

L'étude des échantillons ci-dessus montrent qu'il est possible de former des complexes d'inclusion de type polymorphe V en utilisant des lipides issus de systèmes bactériens. La comparaison entre l'échantillon de BET et d'acide mycoliques peut être expliquée par des effets similaires à ceux observés dans le cadre des phospholipides, mais dans le cas des MMTs et DMTs, d'autres paramètres doivent être pris en compte. La présence d'un groupe tréhalose semble être délétère pour la formation de ces structures, sans empêcher complètement la complexation de ces lipides. Afin d'obtenir de plus amples informations structurales sur ces lipides, nous avons donc réalisé des calculs de dynamique moléculaire qui vont maintenant être décrits.

## II. Résultats théoriques

De manière analogue aux lipides classiques, des calculs de dynamique moléculaire ont été entrepris sur des systèmes contenant un lipide proche d'une chaîne d'amylose de 45 résidus.

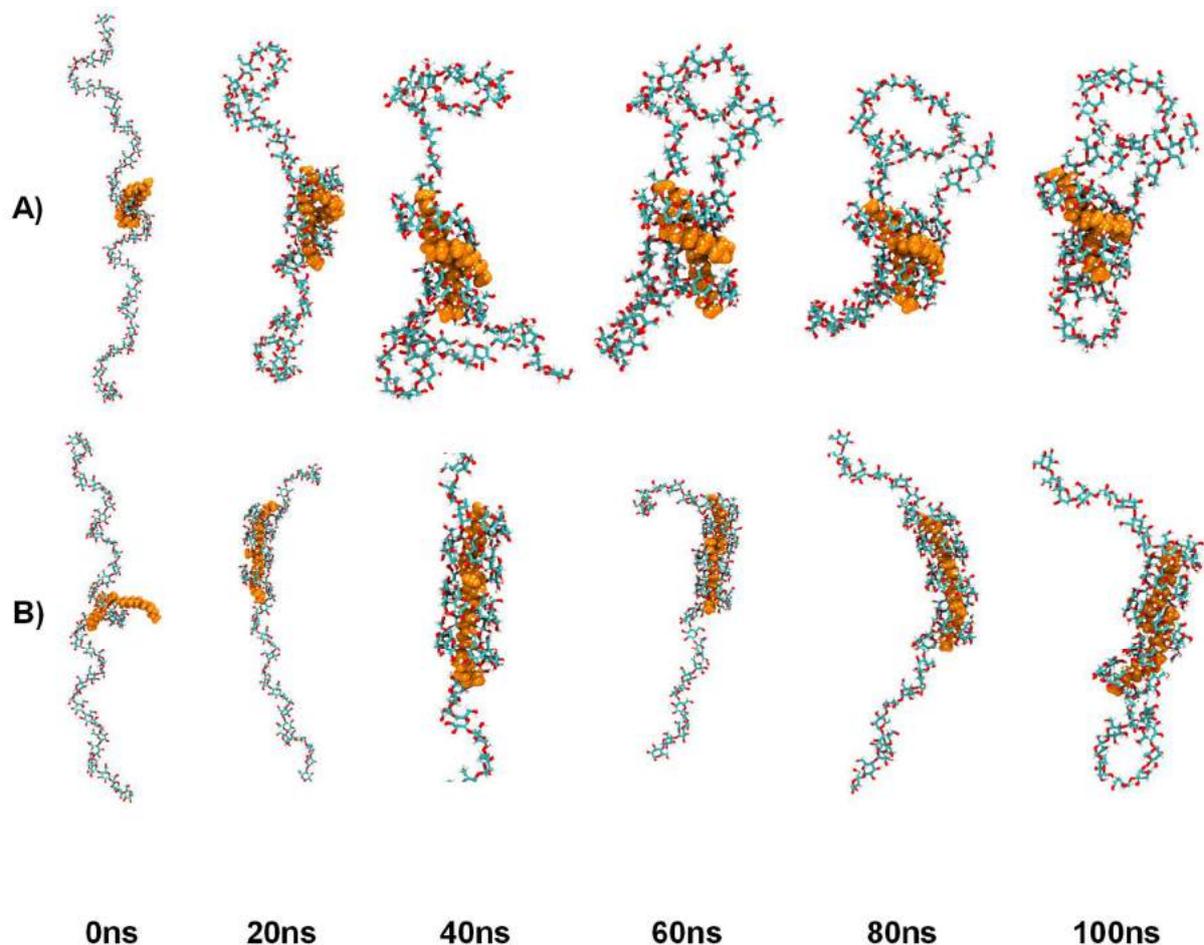


Figure 65: Structures extraites des trajectoires de calcul des systèmes contenant un polymère d'amylose de 45 résidus en présence d'une molécule de BET : Trajectoire de repliement avec comme structure initiale une conformation « membranaire » (A) ou une conformation étendue (B).

Les calculs réalisés sur le BET montrent la formation d'une structure hélicoïdale de l'amylose autour du lipide Figure 65, A). Cependant, une des deux chaînes lipidiques du BET se replie autour de l'autre ce qui empêche la formation d'un allomorphe particulier ni même d'un mélange d'allomorphes comme en atteste la distribution des distances  $[n]-[n+X]$  (Figure 66). Bien que ce lipide présente deux chaînes hydrophobes identiques à celles du DPPC, nous n'observons pas l'exclusion de l'une d'entre elle au cours de la simulation, comme nous l'avons observé pour le DPPC. Ceci peut être expliqué par l'encapsulation entière du lipide, groupe ester inclus, au sein du polysaccharide, là où les groupes phosphatidylcholine des phospholipides restaient à l'extérieur de l'amylose. L'absence d'allomorphe n'empêche cependant pas d'observer un maximum de distribution des angles dièdres aux alentours de  $\varphi=115^\circ$  et  $\psi=105^\circ$  (Figure 67), ce qui montre bien qu'une structure hélicoïdale (définie sur la

base des angles dièdres de la liaison glycosidique) peut être observable sans que l'empilement des résidus entre deux tours d'hélice ne soit observé.

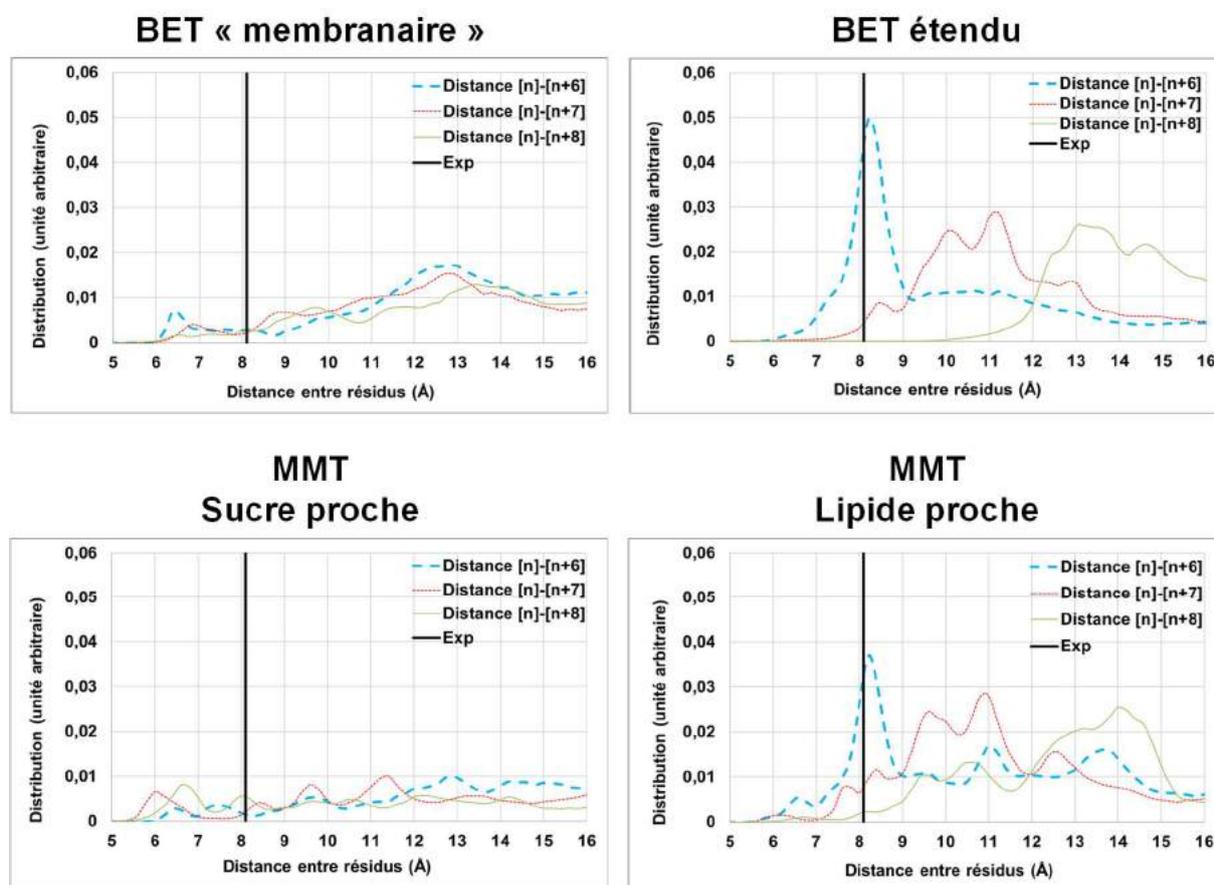


Figure 66: Distribution des distances  $[n]-[n+6]$  (pointillés bleus),  $[n]-[n+7]$  (tirets rouges),  $[n]-[n+8]$  (trait vert) obtenus sur tous les résidus et échantillonnées sur les 100ns des simulations de repliement en présence d'une molécule de BET et de MMT.

Le groupement méthyl présent sur la fonction ester implique donc la possibilité que le BET puisse être entièrement encapsulé de manière complètement étendue. Afin de tester cette possibilité, nous avons réalisé une simulation de complexation dans laquelle la conformation initiale du BET est étendue (Figure 65, B). Au cours du calcul, le BET se retrouve rapidement encapsulé par l'amylose, groupe ester compris et la distribution des distances  $[n]-[n+X]$  atteste de la formation d'un allomorphe V6 (Figure 66). De plus, la distribution des angles dièdres se localise fortement aux alentours de  $\varphi=115^\circ$  et  $\psi=105^\circ$  (Figure 67). La différence de distribution entre le BET et BET étendu obtenue semble bien indiquer que dans sa conformation étendue, le BET implique la structuration en hélice d'une proportion relative plus grande des 45 résidus de la chaîne d'amylose.

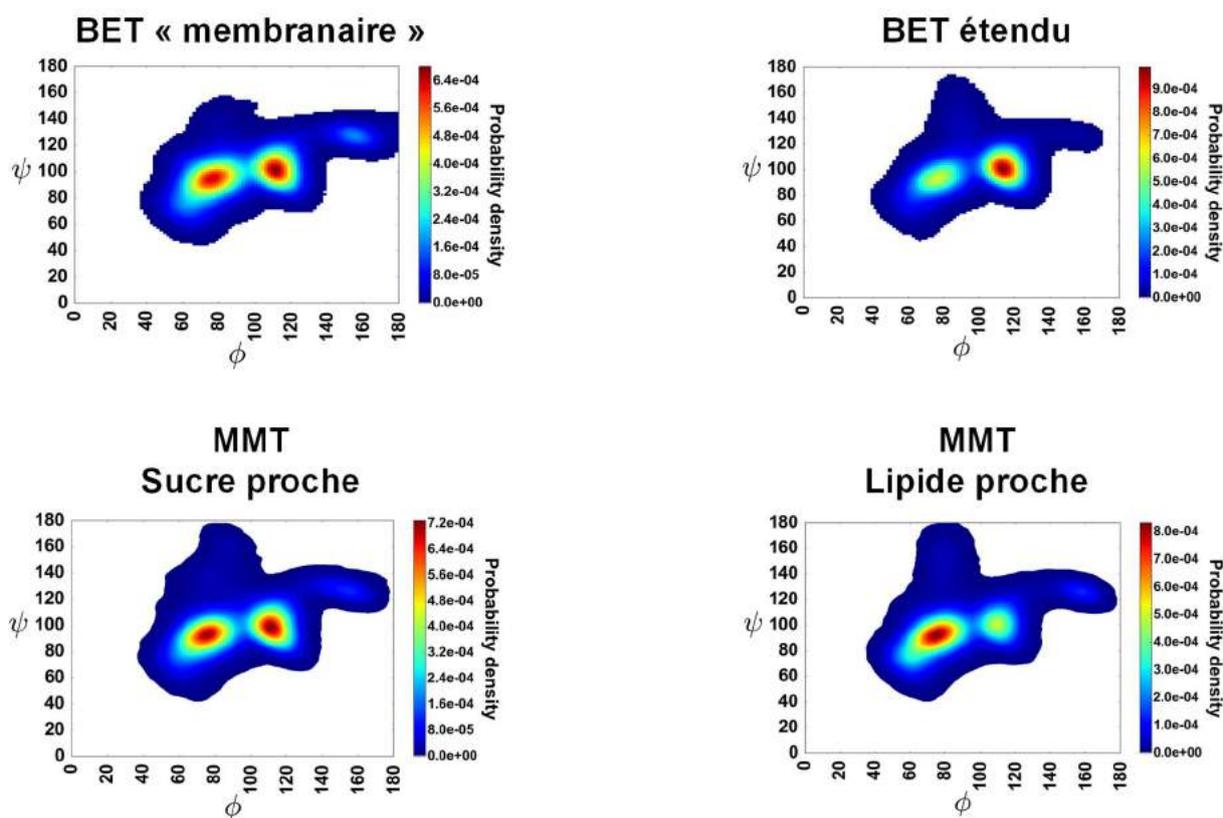


Figure 67: Diagramme de type Ramachandran représentant les angles dièdres  $\phi$  et  $\psi$  de la liaison glycosidique pour tous les résidus sur l'ensemble des trajectoires obtenues sur les systèmes contenant du BET et du MMT.

Deux structures sont donc a priori possibles concernant ces complexes dont l'une présente un empilement plus régulier et stable des résidus glucose entre deux tours d'hélice. Dans le premier chapitre nous avons démontré qu'un tel empilement devait engendrer plus de liaisons hydrogène sur les groupes hydroxyles des carbones C<sub>6</sub> et donc devait impliquer la diminution du signal RMN de ces derniers de 61 à 59 ppm. Nos données expérimentales révèlent la présence de deux populations de C<sub>6</sub> à 59 et 61 ppm. Il est donc probable que les deux structures que nous avons caractérisées ci-dessus soient présentes dans nos échantillons expérimentaux et contribuent à la présence de ces deux pics.

Par la suite, nous nous sommes intéressés à la différence observée entre les acides mycoliques et les MMTs, qui réside principalement dans l'ajout du groupe tréhalose branché en position 6. Pour ceci, deux calculs ont été réalisés sur des systèmes contenant une molécule de MMT placée à proximité d'une chaîne d'amylose de 45 résidus, tels que décrits dans la partie méthode du chapitre.

Lorsque la partie lipidique est proche de l'amylose dans la structure initiale, le polymère se replie rapidement autour des chaînes hydrophobes (Figure 68). La chaîne complètement saturée est isolée et encapsulée seule, et de ce fait le pas de l'hélice présente une distribution caractéristique un allomorphe V6 (Figure 66) ce qui est cohérent avec les résultats précédents d'encapsulation d'une seule chaîne grasse. La chaîne hydrophobe présentant une insaturation est quant à elle encapsulée dans une forme repliée. En contrepartie, lorsque le groupement tréhalose est positionné proche de l'amylose, aucun allomorphe défini ne peut être observé (Figure 66).

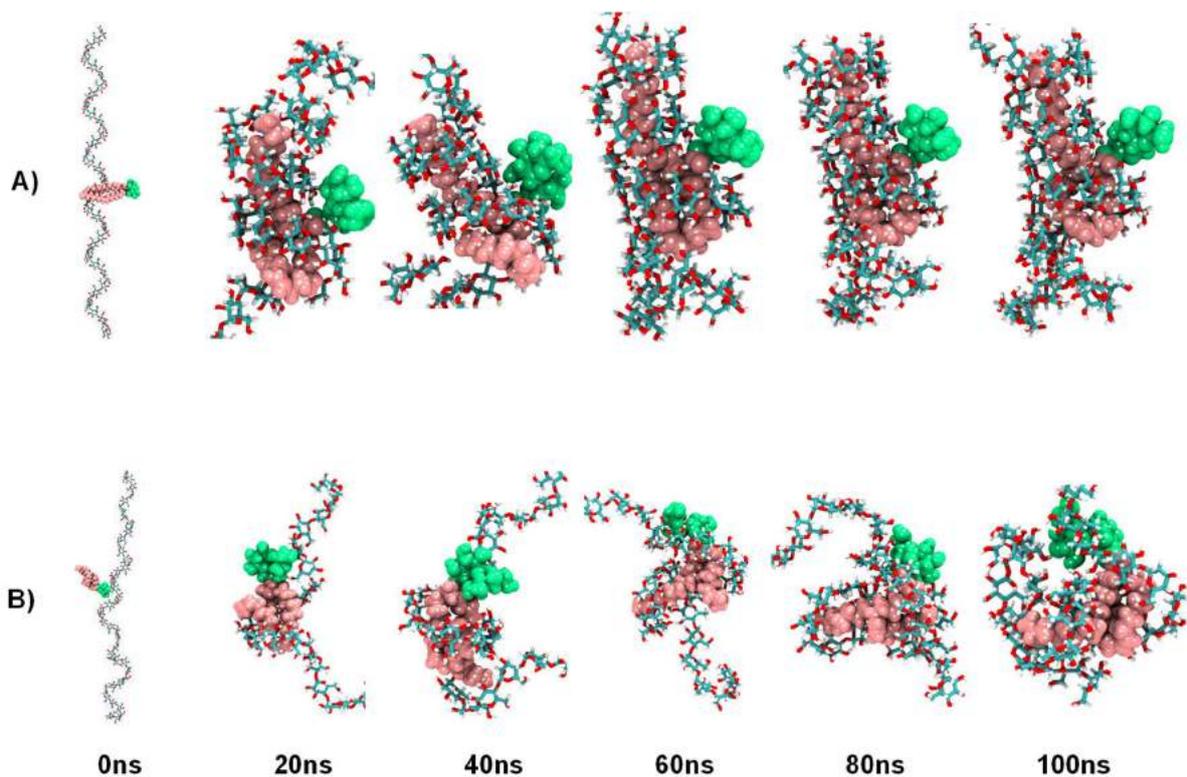


Figure 68: Structures extraites des trajectoires de calcul des systèmes contenant un polymère d'amylose de 45 résidus en présence d'une molécule de MMT : Trajectoire avec comme structures initiales la partie lipidique proche de l'amylose(A) et le groupement tréhalose proche de l'amylose(B). Rose : chaîne mycolique 341. Vert : groupe tréhalose.

Ce comportement peut être expliqué par la présence de liaisons hydrogène formées entre les groupes hydroxyles du tréhalose et ceux des résidus glucose de l'amylose. Lorsque la partie hydrophobe du lipide est au contact de l'amylose, le polymorphe V6 se structure rapidement. Ce n'est aussi qu'à partir de 10 ns que des liaisons hydrogène peuvent être constatées entre le groupe tréhalose et les résidus de glucose de l'amylose qui en sont les plus proches (Figure 69). Par la suite, une à deux liaisons hydrogène resteront établies jusqu'à la fin de la simulation. Lorsque le groupe tréhalose est placé à proximité de l'amylose, des liaisons hydrogène sont

instantanément établies entre ce groupe et les résidus glucoses du polymère et resteront présentes jusqu'à la fin de la simulation. Ces liaisons stabilisent les interactions entre les groupements saccharidiques des deux molécules au détriment des interactions hydrophobes présentes dans un polymorphe V (Figure 69). La formation d'un des allomorphes est alors compromise.

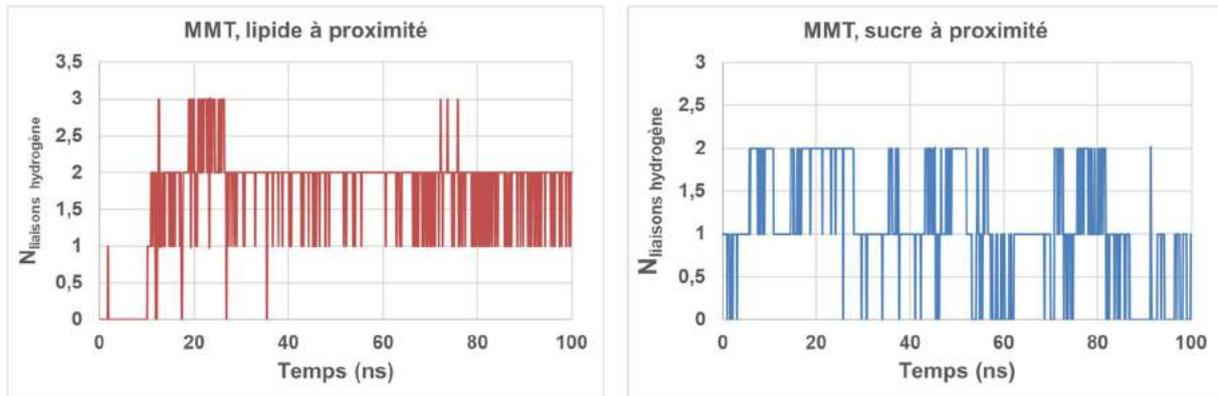


Figure 69: Nombre de liaison hydrogène établies entre le groupe tréhalose et l'amylose pour les trajectoires avec la chaîne lipidique (haut) ou le groupe tréhalose (bas) initialement placée à proximité de la chaîne d'amylose.

La distribution des angles dièdres présente cependant dans les deux cas un maximum aux alentours de  $\varphi=115^\circ$  et  $\psi=105^\circ$  (Figure 67). Ce maximum est plus intense dans la simulation dont la tête tréhalose est proche de l'amylose en raison d'un plus grand nombre de résidus présents à proximité du lipide. Bien qu'il soit impossible d'observer la formation d'un allomorphe, le lipide tend tout de même à structurer l'amylose autour des parties hydrophobes.

La différence de structure entre les MMT et les DMT réside dans l'ajout d'une chaîne mycolique en position 6' du tréhalose. Cette dernière permet de diminuer les interactions tréhalose-amylose au profit d'interactions chaîne mycolique-amylose. Concernant ce lipide, la présence des deux chaînes mycoliques séparées par un groupement tréhalose permet de générer différents types de structures en présence d'amylose. Les calculs réalisés sur ces systèmes sont partis de trois types de conformations initiales différentes (décrites plus haut) qui ne se veulent pas exhaustives des conformations possibles de ce type de système.

Les calculs dans lequel le DMT est dans sa conformation membranaire montrent un repliement de l'amylose autour des chaînes. Plusieurs conformations différentes sont obtenues en fonction de la nature des chaînes mycoliques (C32 :0, C34 :1 et C36 :2),

et la formation d'une hélice autour d'une à trois chaînes lipidiques est observée (Figure 70).

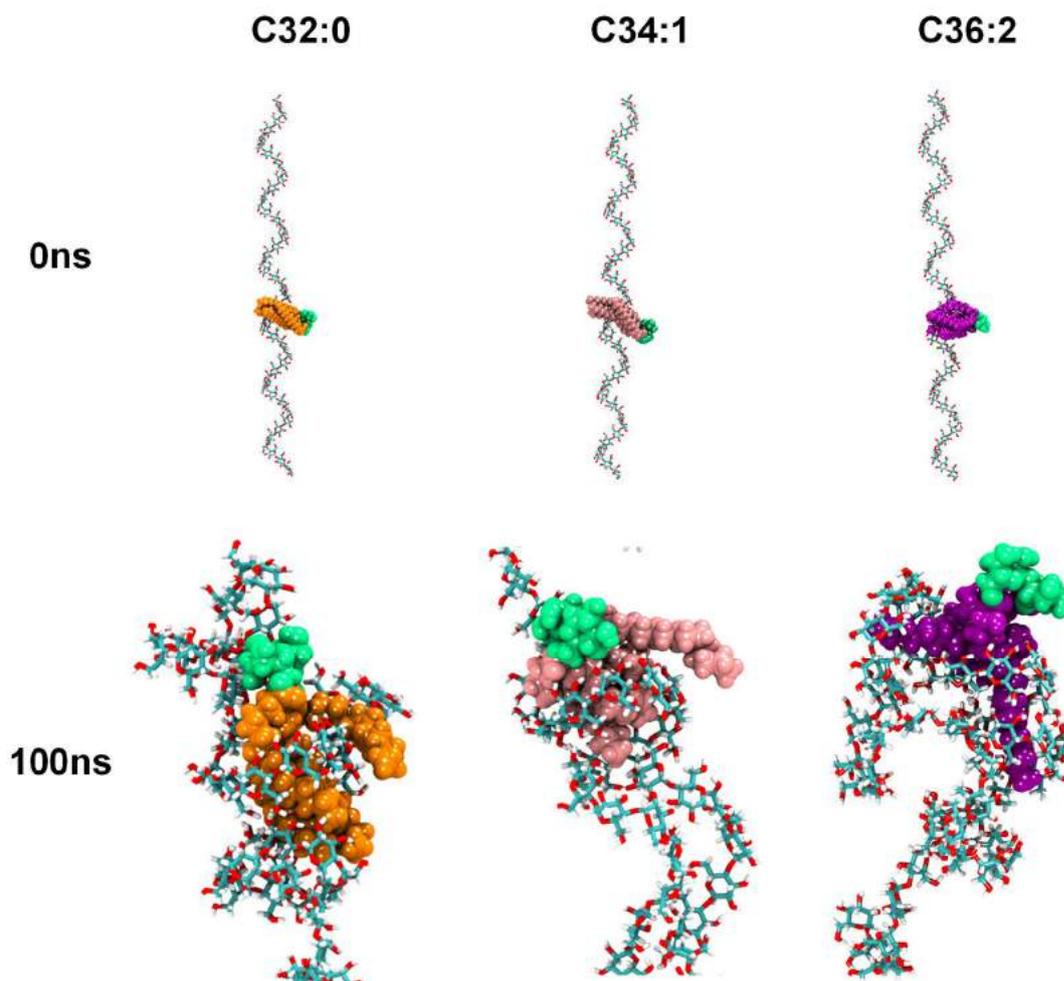


Figure 70: Structures obtenues après 100ns de simulation en présence d'une molécule de DMT dans sa conformation "membranaire". Orange : chaîne mycolique 320. Rose : chaîne mycolique 341. Violet : chaîne mycolique 362. Vert : groupe tréhalose.

Concernant ces trois calculs, la distribution des distances  $[n]-[n+X]$  n'atteste du début de formation d'un polymorphe que dans le cas du DMT C36:2 (Figure 71) ce qui est en accord avec les structures observées à 100ns (Figure 72). Cependant, les distributions d'angles dièdres présentent toutes un maximum situé aux alentours de  $\varphi=115^\circ$  et  $\psi=105^\circ$ . Il semblerait donc que ces lipides soient susceptibles d'engendrer de la structuration au sein de l'amylose, sans pour autant qu'un allomorphe ne puisse être observé. Cependant ceci entre en contradiction avec nos résultats expérimentaux qui présentent un signal des  $C_6$  à 59 ppm (caractéristique de polymorphes V stabilisés par des liaisons H intra-amylose).

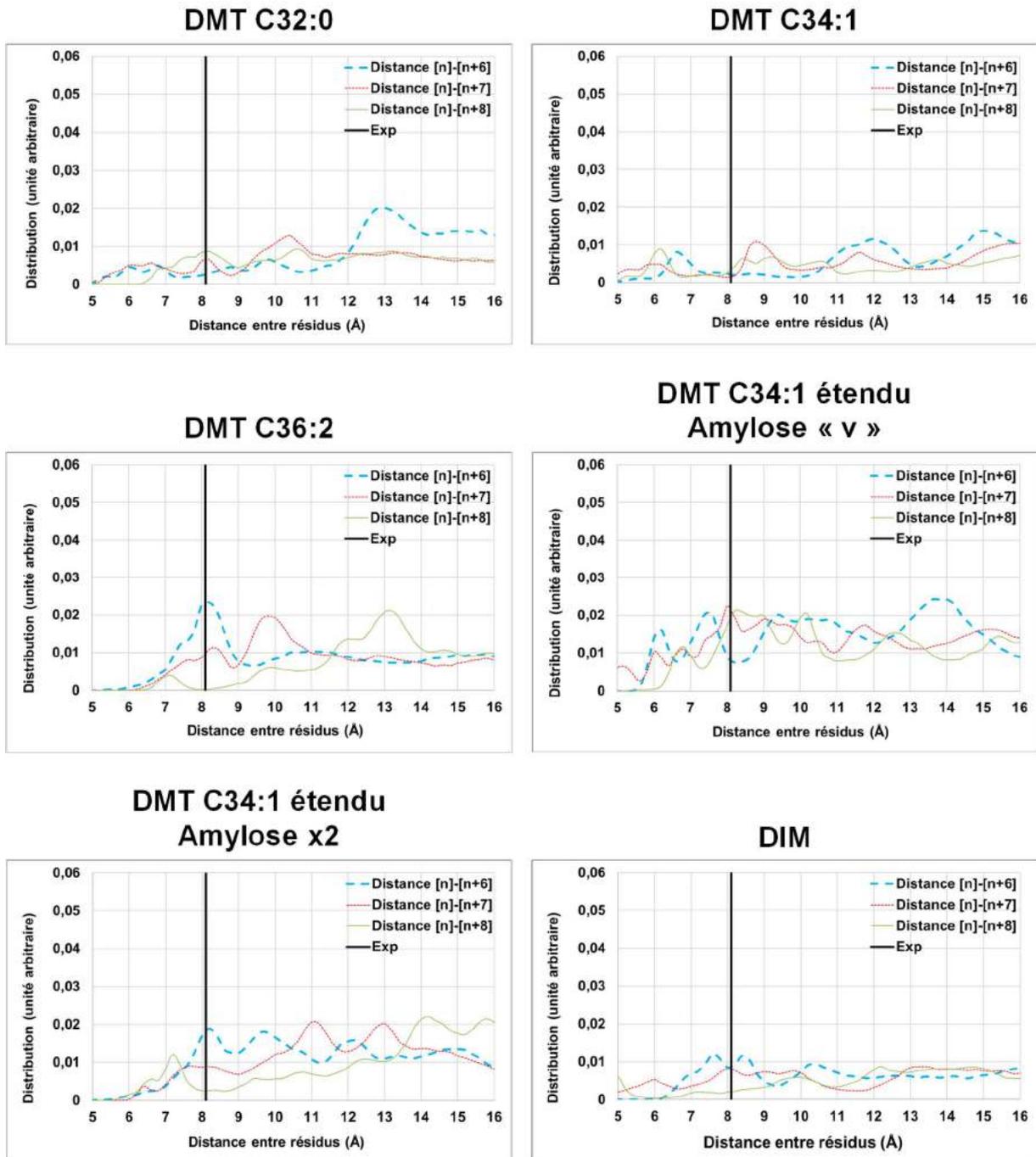


Figure 71: Distribution des distances  $[n]$ - $[n+6]$  (pointillés bleus),  $[n]$ - $[n+7]$  (tirets rouges),  $[n]$ - $[n+8]$  (trait vert) obtenus sur tous les résidus et échantillonnées sur les 100ns des simulations de repliement en présence d'une molécule de DMT et de DIM.

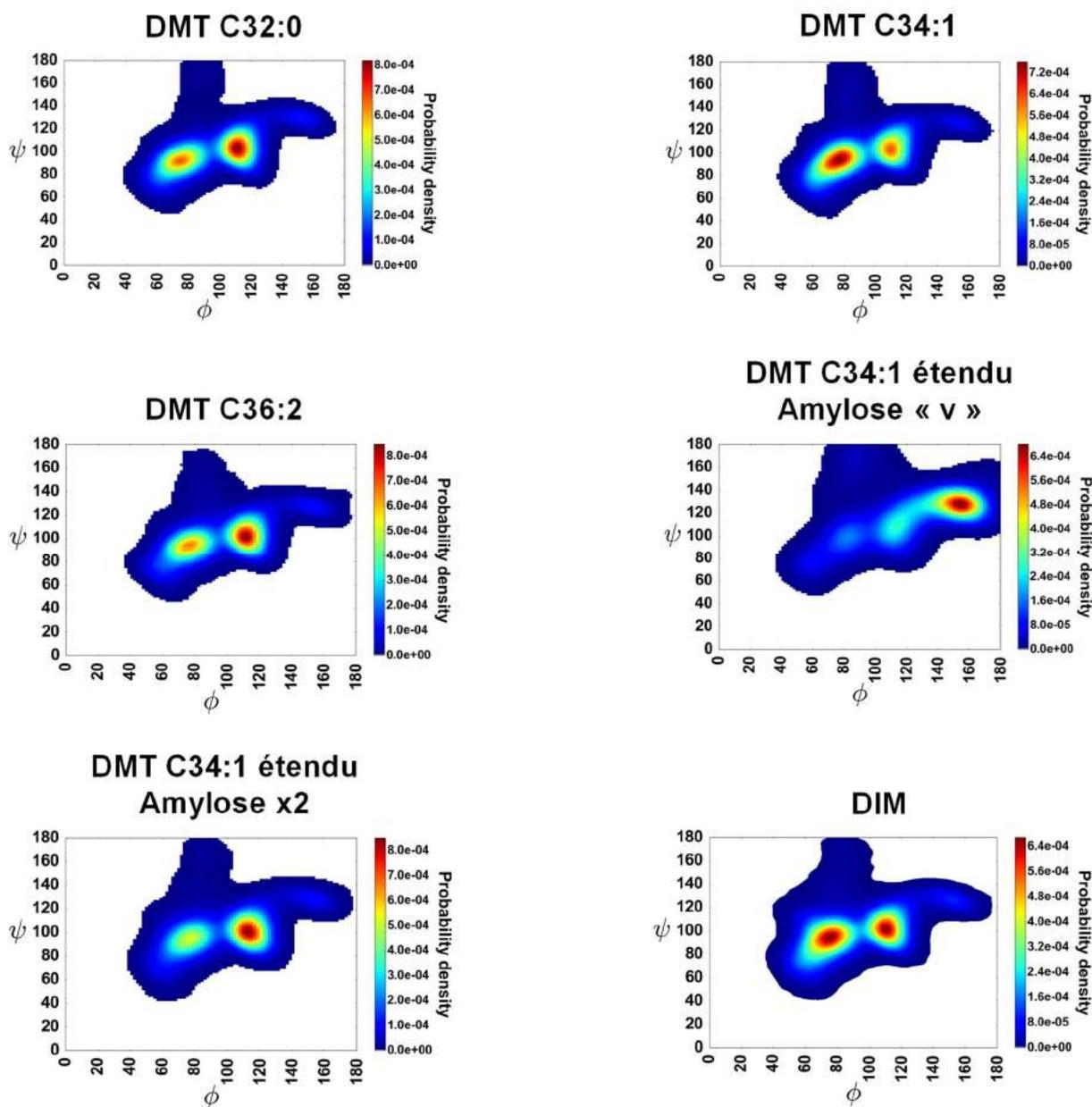


Figure 72: Diagramme de type Ramachandran représentant les angles dièdres  $\phi$  et  $\psi$  de la liaison glycosidique pour tous les résidus sur l'ensemble des 100ns de simulations réalisées sur les systèmes contenant du DMT et du DIM.

Nous avons postulé que notre modèle chimique pouvait atteindre sa limite et avons de ce fait envisagé d'autres structures possibles de ce lipide en présence d'amylose. Du fait de la séparation des deux chaînes mycoliques par le groupement tréhalose, une conformation complètement étendue pouvait notamment être envisagée au sein de l'amylose. Un calcul a donc été réalisé en utilisant une conformation étendue du DMT en présence d'amylose initialement courbée en forme de « V » afin de savoir si cette conformation du lipide pouvait être stabilisée. Les calculs effectués sur ce type de

structure montrent que l'amylose initialement placée courbée en V se replie effectivement autour de chacune des chaînes mycoliques (Figure 73, A).

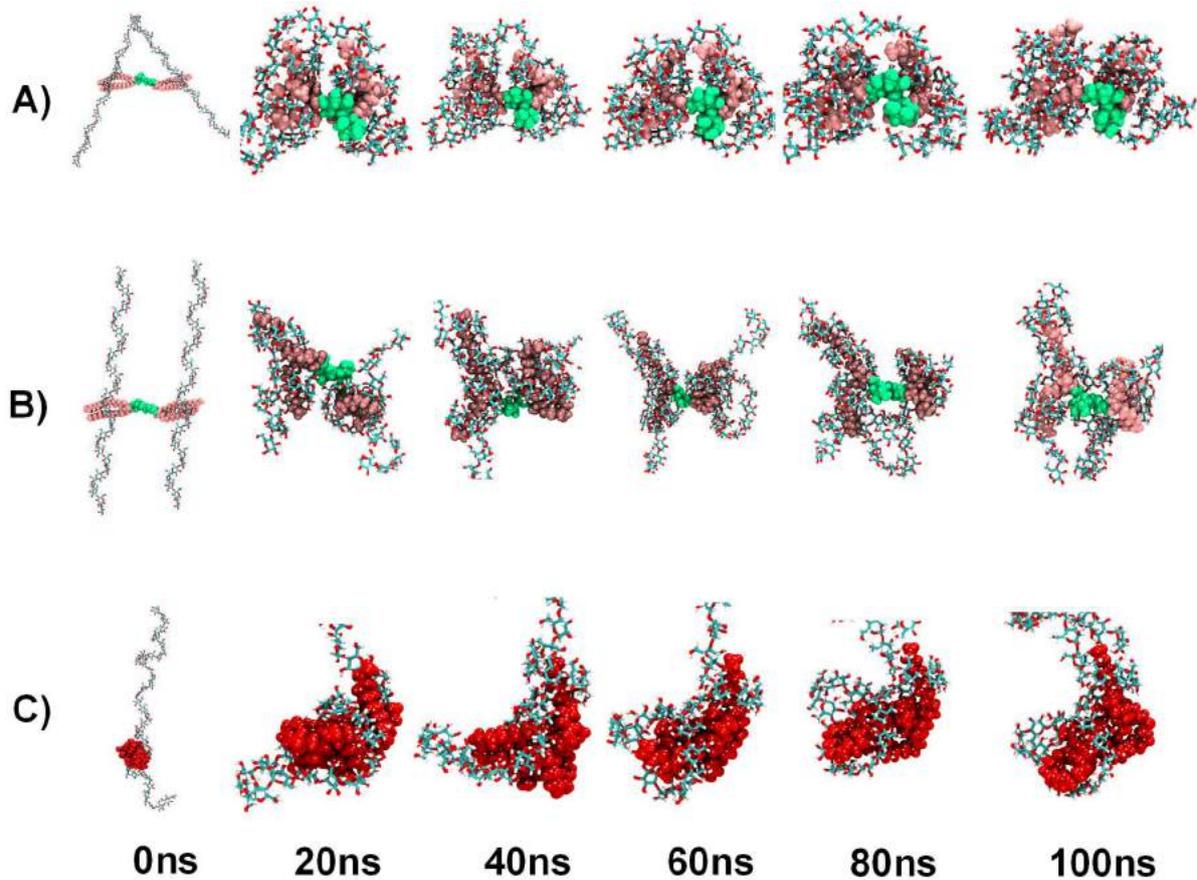


Figure 73: Structures extraites des trajectoires de calcul des systèmes contenant une molécule de DMT en conformation étendue à proximité d'une chaîne d'amylose de 45 résidus repliée en « V » (A) et de deux chaînes d'amylose de 30 résidus (B) ou contenant une molécule de DIM à proximité d'une chaîne d'amylose de 45 résidus (C).

Cependant une fois que le repliement est entamé, les degrés de libertés du polymère et du lipide deviennent limités. De plus, on observe peu à peu des interactions entre les résidus glucose de l'amylose repliés ainsi qu'avec le groupement tréhalose, ce qui empêche la formation complète du polymorphe. Ceci peut être caractérisé plus facilement en regardant les diagrammes de type Ramachandran établis pour cette trajectoire (Figure 72, étendu amylose V). On y observe un maximum de distribution des angles dièdres aux alentours de  $\varphi=155^\circ$  et  $\psi=130^\circ$  qui correspondent à des résidus qui sont situés entre la partie hélicoïdale et la partie « random-coil » de la chaîne d'amylose. Une contribution non négligeable est aussi observée aux alentours de  $\varphi=115^\circ$  et  $\psi=105^\circ$  ce qui atteste aussi de la présence de résidus dans une structure hélicoïdale. Cette contribution associée à la distribution des distances  $[n]-[n+X]$  nous

informe de la présence notable d'un mélange d'allomorphes V7/V8. Notre modèle ici semble se stabiliser dans une structure transitoire, qui peut être dû au faible nombre de résidus présent au sein de l'amylose. En effet, dans la partie de ce chapitre sur les phospholipides nos diagrammes de type Ramachandran montrent que 18 résidus sont stabilisés en hélice autour de deux chaînes hydrophobes (Figure 58). Notre modèle d'amylose ne contient que 45 résidus ce qui laisse penser qu'il n'y en a pas suffisamment pour générer une boucle entre les deux parties se repliant autour du DMT.

De ce fait, afin de minimiser les interactions intra amylose et de permettre une plus grande flexibilité au polymère lors de la formation du polymorphe V autour de chacune des chaînes mycoliques, la dernière trajectoire réalisée contenait deux chaînes d'amylose de 30 résidus (Figure 73, B). Cette fois-ci, la distribution des angles dièdres se localise en majorité aux valeurs de  $\varphi=115^\circ$  et  $\psi=105^\circ$ , ce qui atteste bien de la formation d'une structure hélicoïdale de l'amylose (Figure 72, étendu Amylose x2) De plus, la distribution des distances  $[n]-[n+X]$  atteste du début de formation d'un allomorphe V6 (Figure 71) ce qui est cohérent avec les structures observées (Figure 73). Nos données expérimentales montrent deux populations de  $C_6$  à 59 et 61 ppm, il est donc plausible que des structures étendues et membranaire du DMT soient présentes dans nos échantillons.

Suite à l'étude des acides mycoliques et de leurs dérivés, nous avons finalement réalisé une trajectoire de dynamique moléculaire sur un système contenant une molécule de DIM afin d'en observer le comportement à proximité de l'amylose. Ce lipide est extrêmement hydrophobe et de ce fait se replie très rapidement sur lui-même au cours de la simulation. Cependant, nous observons un repliement de l'amylose autour de ce lipide, notamment autour d'une des chaînes comme en atteste la structure observée à 100 ns (Figure 73). La distribution des angles dièdres atteste aussi d'une structure hélicoïdale du polysaccharide ce qui peut être confirmé par le maximum observé à  $\varphi=115^\circ$  et  $\psi=105^\circ$ . Nous n'observons cependant pas la présence d'un allomorphe particulier (Figure 71). Nos résultats expérimentaux montrent la présence d'une population non négligeable de  $C_6$  à 61 ppm, ce qui est en accord avec la formation de parties hélicoïdales au sein de l'amylose. En revanche, le signal présent à 59 ppm laisse supposer qu'il est envisageable que des allomorphes particuliers soient formés, bien que nous n'ayons pas pu les observer lors de nos calculs.

### III. Résumé des résultats obtenus

La formation de polymorphe V en présence de lipides mycobactériens n'a jusqu'à aujourd'hui jamais été caractérisée. Le protocole décrit ci-dessus nous a permis de mettre en évidence la formation ce type de complexes en présence de BET, de MYC, de MMT, de DMT et de DIM.

Comme dans le cas des lipides classiques, la présence d'insaturations au sein du lipide facilite la formation de polymorphe V, comme le montrent la différence entre les échantillons BET et MYC. Comme l'ont montré les différents calculs réalisés sur le BET, l'absence d'un groupement de haute polarité (phosphatidylcholine, acide carboxylique ...) permet, en contrepartie d'envisager des structures dans lesquelles le lipide est complètement encapsulé dans l'amylose. Ces structures attestent bien de la formation du polymorphe V, mais entre en contradiction avec les modèles jusqu'ici établis concernant ces complexes d'inclusion. En effet, la plupart des modèles récents postulent la présence du groupe hydrophile des lipides à l'extérieur du polymorphe V, et établissant la limite entre les phases cristallines et amorphes au sein des échantillons<sup>22-24</sup>. Nos structures obtenues en présence de BET étendu suggèrent que ces phases ordonnées pourraient ne pas être observées. Des phases dans lesquels les lipides sont encapsulés dans l'amylose sans empilement de ces complexes pourraient ainsi être envisagées dans nos échantillons. Le modèle proposé récemment par Kong et co-auteurs afin de caractériser des complexes de type « I.a » forme une bonne illustration de ce que nous décrivons ici<sup>12</sup> (Figure 74).

L'ajout du groupement tréhalose tend à empêcher la formation de ces complexes, comme l'atteste la différence de quantité de polymorphe V entre les échantillons MYC et MMT. Nos simulations de dynamique moléculaire nous indiquent que les interactions tréhalose-amylose pourraient être responsables de cela. Dans le cas où des liaisons hydrogène sont établies avant la formation de la structure hélicoïdale, les degrés de liberté du lipide seront contraints par rapport à l'amylose, ce qui rendra la formation du polymorphe moins favorable. Dans le cas où la formation du polymorphe est déjà entamée, le lipide va se retrouver stabilisé (chaines mycoliques encapsulées) permettant ensuite la formation de liaisons hydrogène entre le groupement tréhalose et les résidus glucose à proximité.

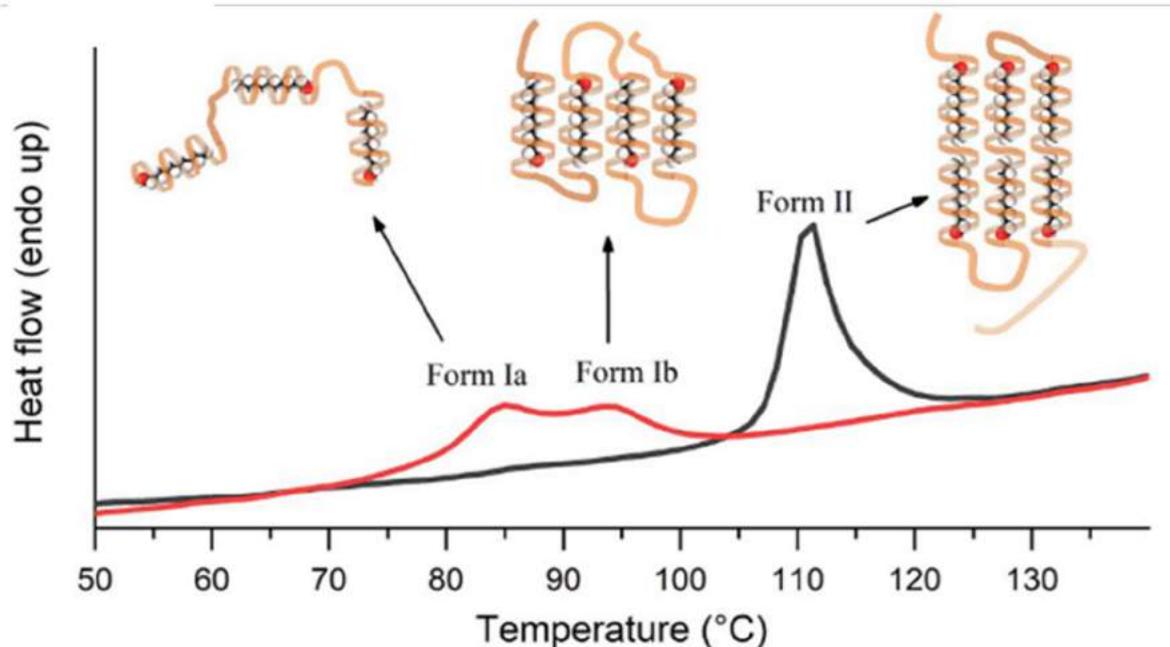


Figure 74: Pics Endomorphes récemment observés par Kong et co-auteurs, et modèles d'organisation leur correspondant. Extrait de Kong (2019)

Dans le cas du DMT, la chaîne mycolique ajoutée en position 6' du groupe tréhalose génère une surface hydrophobe plus grande, pour une tête polaire identique à celle du MMT. Dans sa conformation membranaire, ce lipide peut engendrer la formation de parties hélicoïdales au sein de l'amylose, bien qu'aucun allomorphe ne puisse être observé. Cependant, ce ne sont pas les seules structures pouvant être envisagées. En effet, l'augmentation de la surface hydrophobe peut aller jusqu'à être doublée si on envisage une conformation étendue du DMT, ce qui augmente de ce fait les chances de formation du polymorphe V. Cette conformation étendue permettrait de générer des structures en polymorphe V au sein d'un même brin d'amylose comme au sein de deux brins différents.

Finalement, nous avons démontré qu'il était possible d'observer aussi la formation de polymorphe V en présence de DIM. Nos calculs de dynamique moléculaire ont démontré que ce lipide avait tendance à se replier rapidement sur lui-même en milieu aqueux à température ambiante. Même dans cette conformation, il nous a été possible d'observer le début de formation d'un polymorphe V autour de ce lipide, comme en attestent les structures issues de dynamique moléculaire ainsi que les diagrammes de type Ramachandran.

Si nos résultats obtenus par dynamique moléculaire permettent d'apporter des informations complémentaires aux résultats expérimentaux, il nous faut cependant

rappeler certaines limites de notre méthode. La dynamique moléculaire en tant que méthode de calcul est intrinsèquement dépendante de la durée de la simulation ainsi que de la structure de départ utilisée, et ceci particulièrement si le nombre de degrés de liberté du système étudié est grand. Nous pouvons donc imaginer que nos simulations ne sont pas suffisamment longues, particulièrement dans le cadre des lipides les plus complexes, tels que les DMT et le DIM, pour échantillonner de manière satisfaisante l'espace des configurations de ces derniers. De plus, s'il a été relativement simple de séparer les résidus de glucose en deux groupes selon la distribution des angles dièdres  $\varphi$  et  $\psi$  pour l'étude des acides gras et des phospholipides, cette séparation n'est plus aussi évidente sur les systèmes contenant des lipides mycobactériens. Ces lipides engendrent des conformations de l'amylose qui ne présentent aucun allomorphe particulier, les unités glucose de l'amylose sont susceptibles d'alterner entre des conformations hélicoïdales et de type random-coil rendant ainsi l'interprétation des spectres RMN théoriques plus fastidieuse. Ces spectres RMN théoriques n'étant pas tous disponibles au moment de la rédaction de ce chapitre, nous avons préféré ne pas en discuter concernant ces lipides en particulier.

## 4. Conclusion

Par l'utilisation de RMN du solide du  $^{13}\text{C}$ , de calculs de dynamique moléculaire et de calculs de chimie quantique, nous avons pu caractériser la formation de polymorphe V sur deux acides gras, des phospholipides et les lipides mycobactériens associés.

Dans un premier temps l'utilisation de la RMN  $^{13}\text{C}$  à l'état solide nous a permis de démontrer que la formation de polymorphe V était possible en présence de phospholipides contenant plusieurs chaînes hydrophobes. La présence d'insaturations au sein des chaînes lipidiques promeut la formation de polymorphe V et la concentration en lipides de l'échantillon doit être également ajustée de manière à éviter les agrégats de lipides piégés au sein de l'amylose. De ce fait nous avons proposé un mécanisme de compétition entre la formation de lipides agrégés entre eux et la formation de polymorphe V en solution. Afin de comprendre plus en profondeur le type de structures que pouvait engendrer ces lipides nous avons analysé les structures

théoriques issues de calculs de dynamique moléculaire. Nous avons pu montrer que les structures générées pouvaient être hétérogènes, avec notamment la présence de différents allomorphes (V6/V7/V8) au sein d'une même hélice. Finalement les calculs de chimie quantique nous ont permis de confirmer que quel que soit l'hétérogénéité présente au sein de l'hélice, celle-ci présenterait tout de même un signal RMN  $^{13}\text{C}$  caractéristique de ce polymorphe.

Par la suite, nous avons appliqué notre stratégie couplant expérience et théorie à l'étude de complexes obtenus en présence de lipides mycobactériens. Nous avons prouvé qu'il était possible d'effectuer des complexations avec des acides mycoliques ainsi qu'avec leur analogue de synthèse : le BET. Nous avons pu démontrer que les interactions entre le groupe tréhalose des MMT et l'amylose déstabilisait la formation du polymorphe V et que l'ajout d'une deuxième chaîne mycolique permettait de retrouver un haut taux de formation du complexe. Enfin, nous avons pu aussi caractériser la présence de polymorphe V au sein d'échantillon contenant du DIM, le lipide le plus apolaire de ceux testés dans ce chapitre.

Les complexes d'inclusion amylose-lipide ont la réputation de ne pas pouvoir être observés en présence de lipides contenant plusieurs chaînes hydrophobes. Nos résultats obtenus sur les lipides classiques devraient ouvrir de nouvelles opportunités d'études qui seront utiles pour la communauté agro-alimentaire et nous espérons que la preuve de l'existence de complexes amylose-lipides mycobactérien permettra une meilleure compréhension des interactions polysaccharides-lipides au sein de systèmes biologiques divers.

## Références

- (1) Gidley, M. J.; Bociak, S. M.  $^{13}\text{C}$  CP/MAS NMR Studies of Amylose Inclusion Complexes, Cyclodextrins, and the Amorphous Phase of Starch Granules: Relationships Between Glycosidic Linkage Conformation and Solid-State  $^{13}\text{C}$  Chemical Shifts. *J. Am. Chem. Soc.* **1988**, *110* (12), 3820–3829. <https://doi.org/10.1021/ja00220a016>.
- (2) Lebail, P.; Buleon, A.; Shifan, D.; Marchessault, R. H. Mobility of Lipid in Complexes of Amylose-Fatty Acids by Deuterium And  $^{13}\text{C}$  Solid State NMR. *Carbohydr. Polym.* **2000**. [https://doi.org/10.1016/S0144-8617\(00\)00180-6](https://doi.org/10.1016/S0144-8617(00)00180-6).
- (3) Sang, S.; Chen, Y.; Zhu, X.; Narsimhan, G.; Hu, Q.; Jin, Z.; Xu, X. Effect of Egg Yolk Lipids on Structure and Properties of Wheat Starch in Steamed Bread. *J. Cereal Sci.* **2019**, *86* (October 2018), 77–85. <https://doi.org/10.1016/j.jcs.2019.01.009>.
- (4) Emond, S.; Mondeil, S.; Jaziri, K.; André, I.; Monsan, P.; Remaud-Siméon, M.; Potocki-Véronèse, G. Cloning, Purification and Characterization of a Thermostable Amylosucrase from *Deinococcus Geothermalis*. *FEMS Microbiol. Lett.* **2008**, *285* (1), 25–32. <https://doi.org/10.1111/j.1574-6968.2008.01204.x>.
- (5) Potocki-Veronese, G.; Putaux, J. L.; Dupeyre, D.; Albenne, C.; Remaud-Siméon, M.; Monsan, P.; Buleon, A. Amylose Synthesized in Vitro by Amylosucrase: Morphology, Structure, and Properties. *Biomacromolecules* **2005**. <https://doi.org/10.1021/bm049326g>.
- (6) Dickson, C. J.; Madej, B. D.; Skjevik, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014**. <https://doi.org/10.1021/ct4010307>.
- (7) Case, D. A. Amber 18. *Univ. California, San Fr.* **2018**.
- (8) Rath, P.; Saurel, O.; Czaplicki, G.; Tropis, M.; Daffé, M.; Ghazi, A.; Demange, P.; Milon, A. Cord Factor (Trehalose 6,6'-Dimycolate) Forms Fully Stable and Non-Permeable Lipid Bilayers Required for a Functional Outer Membrane. *Biochim. Biophys. Acta - Biomembr.* **2013**. <https://doi.org/10.1016/j.bbamem.2013.04.021>.
- (9) Augenstreich, J.; Haanappel, E.; Ferré, G.; Czaplicki, G.; Jolibois, F.; Destainville, N.; Guilhot, C.; Milon, A.; Astarie-Dequeker, C.; Chavent, M. The Conical Shape of DIM Lipids Promotes Mycobacterium Tuberculosis Infection of Macrophages. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (51), 25649–25658. <https://doi.org/10.1073/pnas.1910368116>.
- (10) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23* (16), 1623–1641. <https://doi.org/10.1002/jcc.10128>.
- (11) Wang, S.; Chao, C.; Cai, J.; Niu, B.; Copeland, L.; Wang, S. Starch–Lipid and Starch–Lipid–Protein Complexes: A Comprehensive Review. *Compr. Rev. Food Sci. Food Saf.* **2020**, *19* (3), 1056–1079. <https://doi.org/10.1111/1541-4337.12550>.
- (12) Kong, L.; Perez-Santos, D. M.; Ziegler, G. R. Effect of Guest Structure on

- Amylose-Guest Inclusion Complexation. *Food Hydrocoll.* **2019**, *97* (June), 105188. <https://doi.org/10.1016/j.foodhyd.2019.105188>.
- (13) Le, C. A. K. Inclusion Complexes of Amylose: Morphogenesis, Crystal Structure and Release of Bioactive Molecules, 2018.
- (14) Rondeau-Mouro, C.; Veronese, G.; Buleon, A. High-Resolution Solid-State NMR of B-Type Amylose. *Biomacromolecules* **2006**. <https://doi.org/10.1021/bm060330x>.
- (15) Le, C. A. K.; Choisnard, L.; Wouessidjewe, D.; Putaux, J. L. Polymorphism of Crystalline Complexes of V-Amylose with Fatty Acids. *Int. J. Biol. Macromol.* **2018**, *119*, 555–564. <https://doi.org/10.1016/j.ijbiomac.2018.07.163>.
- (16) Godet, M. C.; Buléon, A.; Tran, V.; Colonna, P. Structural Features of Fatty Acid-Amylose Complexes. *Carbohydr. Polym.* **1993**. [https://doi.org/10.1016/0144-8617\(93\)90003-M](https://doi.org/10.1016/0144-8617(93)90003-M).
- (17) López, C. A.; De Vries, A. H.; Marrink, S. J. Amylose Folding under the Influence of Lipids. *Carbohydr. Res.* **2012**. <https://doi.org/10.1016/j.carres.2012.10.007>.
- (18) Cheng, L.; Feng, T.; Zhang, B.; Zhu, X.; Hamaker, B.; Zhang, H.; Campanella, O. A Molecular Dynamics Simulation Study on the Conformational Stability of Amylose-Linoleic Acid Complex in Water. *Carbohydr. Polym.* **2018**, *196*, 56–65. <https://doi.org/10.1016/j.carbpol.2018.04.102>.
- (19) Koneru, J. K.; Zhu, X.; Mondal, J. A Quantitative Assessment of the Conformational Heterogeneity in Amylose across Force Fields. *J. Chem. Theory Comput.* **2019**. <https://doi.org/10.1021/acs.jctc.9b00630>.
- (20) Suzuki, S.; Horii, F.; Kurosu, H. Theoretical Investigations of <sup>13</sup>C Chemical Shifts in Glucose, Cellobiose, and Native Cellulose by Quantum Chemistry Calculations. *J. Mol. Struct.* **2009**, *921* (1–3), 219–226. <https://doi.org/10.1016/j.molstruc.2009.01.002>.
- (21) Marrakchi, H.; Lanéelle, M. A.; Daffé, M. Mycolic Acids: Structures, Biosynthesis, and Beyond. *Chemistry and Biology*. 2014. <https://doi.org/10.1016/j.chembiol.2013.11.011>.
- (22) Gelders, G. G.; Goesart, H.; Delcour, J. A. Potato Phosphorylase Catalyzed Synthesis of Amylose-Lipid Complexes. *Biomacromolecules* **2005**, *6* (5), 2622–2629. <https://doi.org/10.1021/bm0502011>.
- (23) Putseys, J. A.; Lamberts, L.; Delcour, J. A. Amylose-Inclusion Complexes: Formation, Identity and Physico-Chemical Properties. *J. Cereal Sci.* **2010**, *51* (3), 238–247. <https://doi.org/10.1016/j.jcs.2010.01.011>.
- (24) Goderis, B.; Putseys, J. A.; Gommès, C. J.; Bosmans, G. M.; Delcour, J. A. The Structure and Thermal Stability of Amylose-Lipid Complexes: A Case Study on Amylose-Glycerol Monostearate. *Cryst. Growth Des.* **2014**. <https://doi.org/10.1021/cg4016355>.

### Chapitre 3. Etude théorique et expérimentale préliminaire des interactions entre polysaccharides branchés en présence de lipides divers.

Afin de nous rapprocher le plus possible de notre thématique biologique initiale, nous avons décidé d'étudier la formation de complexes entre les lipides employés précédemment et des polysaccharides branchés.

Notre objectif initial résidait dans l'étude de l'interaction entre l' $\alpha$ -glucane issu de la capsule de *Mycobacterium tuberculosis*, et des lipides retrouvés en son sein. Nous avons pu démontrer dans le chapitre précédent que ces lipides étaient susceptibles de générer des complexes d'inclusion au sein de l'amylose, et nous voulions savoir si de telles structures sont envisageables au sein de l' $\alpha$ -glucane. Concernant ce polysaccharide, nous ne disposons que d'informations moyennes quant à sa structure : c'est une molécule de haut poids moléculaire ( $13 \times 10^6$  Da) composée de résidus glucoses branchés en  $\alpha$ -(1 $\rightarrow$ 4) et  $\alpha$ -(1 $\rightarrow$ 6). Il reste donc un grand nombre d'inconnues sur ce polysaccharide dont les plus limitantes pour nous sont les suivantes : la longueur des parties ne contenant aucune ramification en  $\alpha$ -(1 $\rightarrow$ 6), le taux et la position des ramifications ainsi que l'homogénéité de distribution de ces dernières.

Au vu de sa structure primaire, l' $\alpha$ -glucane est donc un analogue de l'amylopectine. L'utilisation de cette dernière molécule pourrait donc permettre d'établir un modèle d'étude viable afin de vérifier si le protocole expérimental que nous avons utilisé pour étudier l'amylose est transférable à l'étude de polysaccharides branchés. De plus, de récentes études menées par Beeren et co-auteurs ont démontré, par RMN à l'état liquide, une interaction entre un lipide amphiphile et l'amylopectine<sup>1,2</sup>. En utilisant des séquences d'impulsion de type NOESY, les auteurs ont postulé que ces interactions attestaient de la présence de complexes d'inclusion similaires au polymorphe V. Nous avons, de fait, décidé de réaliser aussi des expériences de complexations sur des échantillons contenant de l'amylopectine. Ces expériences étaient bien plus faciles à établir du fait de la disponibilité commerciale de ce polysaccharide.

Ce chapitre présentera les résultats préliminaires que nous avons obtenus sur des systèmes contenant des polysaccharides ramifiés. En raison de la crise de la COVID-19, toutes les expériences prévues n'ont pas pu être réalisées, ce chapitre ne permettra donc pas d'aboutir à une conclusion quant à la capacité de ces systèmes à former des complexes d'inclusion. Cependant il m'est apparu important de discuter des résultats théoriques et expérimentaux obtenus, car il s'agit d'informations qui pourront être utiles pour de futures études sur ce type de systèmes.

## 1. Résultats expérimentaux

L' $\alpha$ -glucane n'étant pas disponible commercialement, nous avons dû en purifier à partir de cultures de *Mycobacterium tuberculosis*. La limitation technique principale rencontrée lors de la mise en place des expériences réalisées sur les polysaccharides branchés réside dans la solubilité de ces derniers. Nous allons de ce fait décrire la purification du polysaccharide ainsi que la modification du protocole de synthèse de ces complexes.

### I. Purification du glucane et modifications du protocole de synthèse

L'amylopectine de maïs est disponible commercialement et a été fournie par Sigma-Aldrich. Le protocole de complexation entre amylopectine et lipides est identique à celui utilisé pour l'amylose et présenté dans le chapitre précédent.

L' $\alpha$ -glucane a été purifié en suivant un protocole publié précédemment<sup>3</sup>. Le milieu extracellulaire d'une culture de *Mycobactérium tuberculosis* H37Rv a été filtré deux fois sur filtre Nalgene 0.22 $\mu$ m, puis concentré environ 10 fois à l'évaporateur rotatif. Six volumes d'éthanol froid ont été ajoutés afin de provoquer la précipitation des macromolécules, c'est-à-dire des protéines et des polysaccharides dont l' $\alpha$ -glucane. Après centrifugation à 14000 g pendant 2 h, le culot a été récupéré, solubilisé dans un minimum d'eau et dialysé contre de l'eau MilliRo (membrane Spectra/Por MWCO 6-8000 SPECTRA) pendant 24 h afin d'éliminer toute trace de sel et de glycérol. Après

24 h de lyophilisation, les macromolécules ont été solubilisées dans 3 mL d'une solution de  $\text{NH}_4\text{Cl}$  0.01 M à pH 8,3 puis chromatographiées sur une colonne de DEAE-Trisacryl afin d'éliminer les protéines. Les polysaccharides neutres ( $\alpha$ -glucane, arabinomannane et mannane) ont été élués par 100mL de  $\text{NH}_4\text{Cl}$  0.01 M à pH 8,3, et les protéines par 100mL de  $\text{NaCl}$  1M. La fraction neutre a été récoltée en aveugle, puis concentrée jusqu'à approximativement 10 mL. Afin de séparer l' $\alpha$ -glucane de haute masse moléculaire (estimée à  $12 \cdot 10^6$  Da), de l'arabinomannane et du mannane (15 kDa et 4 kDa respectivement), le concentrat a ensuite été filtré sur centricon (30 kDa). La pureté en l' $\alpha$ -glucane de la fraction retenue a été vérifiée par RMN  $^1\text{H}$  liquide à haute température.

Si le protocole expérimental est resté le même que celui de l'amylose concernant les expériences réalisées sur les échantillons contenant de l'amylopectine, il a été modifié dans les expériences réalisées sur des échantillons contenant le polysaccharide mycobactérien. De manière opposée à l'amylose, l' $\alpha$ -glucane est soluble en milieu aqueux et de ce fait notre mélange de solvants (DMSO :  $\text{H}_2\text{O}$ ) ne pouvait pas être utilisé. Lors de la purification dont le protocole est décrit ci-dessus, nous réalisons une précipitation de l' $\alpha$ -glucane en milieu aqueux grâce à l'ajout d'éthanol. Après plusieurs essais, nous avons défini les conditions suivantes : 12.2 mg de polysaccharide purifié sont dissous dans 520  $\mu\text{L}$  d'un mélange Ethanol/eau (28:24, [V/V]) à 70°C. Ce mélange de solvant permet d'observer une solubilisation complète de l' $\alpha$ -glucane et sa précipitation à température ambiante, ce qui est similaire au comportement de l'amylose dans le mélange DMSO-Eau utilisé dans le chapitre précédent.

Le changement de solvant ainsi que la réduction de la température initial de solubilisation de 90°C à 70°C, forment les seules modifications du protocole utilisé dans le chapitre 2 de ce travail. La concentration molaire en polysaccharide (rapportée en concentration molaire de glucose) est identique à celle utilisée dans les échantillons contenant de l'amylose.

## II. Résultats

Nous avons réalisé deux expériences de complexation en utilisant de l'amylopectine seule et en présence d'acide palmitique. Nous avons observé une différence de texture

entre les culots obtenus sur des échantillons d'amylose et des échantillons d'amylopectine. En effet les culots d'amylopectine ne présentent pas un précipité blanc, mais plutôt un culot gélatineux. Cette phase pouvant tout de même être séparée du surnageant par centrifugation, nous avons décidé de poursuivre les étapes de lavage, de lyophilisation et de réhydratation du culot avant de procéder à la réalisation des spectres RMN  $^{13}\text{C}$  (Figure 75).

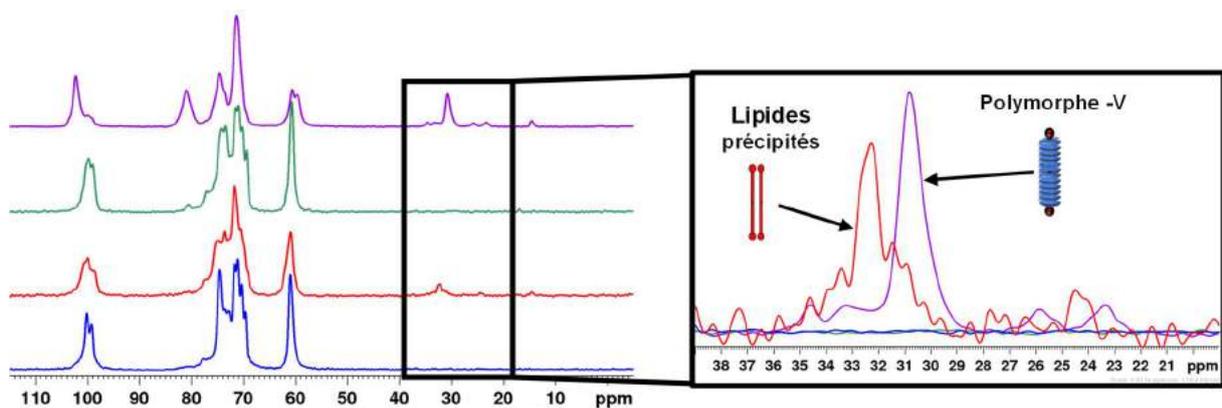


Figure 75: Spectres RMN  $^{13}\text{C}$  obtenus sur des échantillons d'amylopectine seule (bleu) et en présence d'acide palmitique à un ratio molaire de 6.25% (rouge). Pour comparaison, les spectres d'amylose seule (vert) et en présence d'acide palmitique au même ratio molaire (violet) sont présentés. A droite : un agrandissement de la partie concernant les  $\text{CH}_2$  des parties lipidiques.

L'amylopectine seule présente le spectre typique attestant de la formation de polymorphe B, avec notamment la présence du double pic caractéristique des  $\text{C}_1$  situé à 99-100 ppm. Celui-ci présente une homogénéité supérieure à celle observée dans l'échantillon contenant de l'amylose seule. Ceci peut être expliqué par les structures de l'amylopectine, qui présente des parties non-ramifiées contenant 11-15 résidus, et de l'amylose qui elle possède un DP beaucoup plus élevé. Ces chaînes courtes sont plus susceptibles de former des structures de type polymorphe B de haute cristallinité que les chaînes longues, impliquant une meilleure homogénéité de l'échantillon<sup>4,5</sup>. L'ajout d'acide palmitique dans l'échantillon ne permet pas d'observer la présence de polymorphe V comme en atteste l'absence de signal à 102 ppm pour les  $\text{C}_1$  et 82 ppm pour les  $\text{C}_4$  des molécules de glucose. De plus, seule une petite quantité de lipides (caractérisée par un signal de faible intensité) est présente dans l'échantillon, et le signal des  $\text{CH}_2$  de ces derniers est situé à 32.5 ppm ce qui indique qu'il s'agit de lipides agrégés entre eux. Par comparaison avec les changements d'allure du spectre de l'amylose en présence du lipide, nous pouvons penser que l'acide palmitique bien que non complexé de façon stable sous forme de polymorphe V a interagi avec l'amylopectine perturbant sa rétrogradation en un polymorphe B homogène. Les

lavages à l'éthanol peuvent avoir éliminé de façon efficace le lipide non complexé et non précipité du gel d'amylopectine.

Nous avons postulé que la solubilité particulière de l'amylopectine dans le solvant DMSO/Eau utilisé pouvait être la raison de l'absence de formation/stabilisation de polymorphe. Notre objectif étant l'étude de l' $\alpha$ -glucane, qui présente une solubilité différente, nous avons décidé de ne pas pousser l'étude de l'amylopectine plus en avant et d'explorer les interactions entre l' $\alpha$ -glucane et des lipides. Nous avons donc changé de solvant et de température comme indiqué ci-dessus et choisi un lipide qui permettait de générer une quantité plus importante de polymorphe V dans l'amylose que l'acide palmitique, à savoir le POPC.

Les spectres obtenus après réhydratation de l'échantillon à hauteur de 40% présentent très peu de signal du polysaccharide (Figure 76). Ceci peut être expliqué par la solubilité de ce dernier en milieu aqueux bien supérieure à celle de l'amylose et de l'amylopectine.

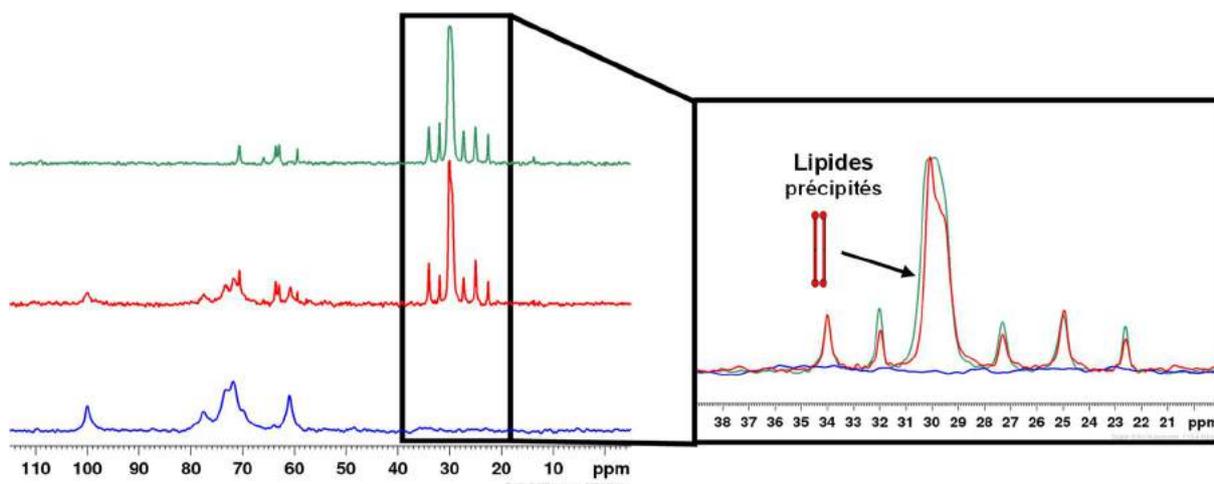


Figure 76 : Spectres RMN <sup>13</sup>C obtenus sur des échantillons d' $\alpha$ -glucane seul (bleu) et en présence de POPC à un ratio molaire de 6.25%(rouge). Pour comparaison, le spectre de POPC seul ayant subi le même protocole (vert) est aussi représenté. A droite est présenté un agrandissement de la partie concernant les CH<sub>2</sub> lipidiques.

En effet, la séquence d'impulsion que nous utilisons pour réaliser ces spectres est basée sur un transfert d'aimantation du proton vers le carbone via une polarisation croisée adaptée à l'étude d'échantillon solide présentant une dynamique restreinte. L' $\alpha$ -glucane étant plus soluble, son comportement est beaucoup plus dynamique après réhydratation, rendant ainsi le transfert de polarisation moins efficace.

Cela étant, le signal observé sur ces deux échantillons ne permet pas de caractériser la présence de polymorphe V ou B en leur sein. Le signal des C<sub>1</sub> une fois l'échantillon

réhydraté se présente sous la forme d'un pic unique situé à 100 ppm, et on observe un pic présent à 78 ppm que nous attribuons aux C<sub>4</sub>, par comparaison avec le signal observé dans des solutions aqueuses d'amylose<sup>6</sup>. Il semblerait que contrairement à l'amylopectine, l' $\alpha$ -glucane ne présente pas une distribution de longueur de chaîne ainsi que de ramifications susceptible d'engendrer des phases organisées sans ajout d'une autre molécule. L'ajout de POPC ne permet pas de générer d'autres signaux pour les C<sub>1</sub> et les C<sub>4</sub> de l'échantillon que ceux déjà observés dans l'échantillon ne comprenant que de l' $\alpha$ -glucane. Le signal des CH<sub>2</sub> des chaînes lipidiques du POPC se situe aux alentours de 29-30 ppm. Afin de vérifier qu'il s'agissait bien de lipides agrégés entre eux, nous avons réalisé un spectre RMN sur un échantillon ne contenant que du POPC et ayant subi le même protocole de « complexation ». Les deux spectres présentent la même distribution de pics, prouvant ainsi les signaux RMN étaient bien dus à de l'auto-précipitation du lipide et non pas à du POPC encapsulé dans l' $\alpha$ -glucane.

Pour le moment nos résultats expérimentaux ne permettent donc pas d'attester de la présence de complexes d'inclusions au sein de l' $\alpha$ -glucane. Des expériences sont prévues pour la suite de ce travail en utilisant de l'acide oléique à différentes concentrations. Il s'agit d'un lipide ne précipitant pas ce qui nous permettra de diminuer les interactions entre lipides mises en jeu dans le mécanisme de compétition que nous avons postulé dans le chapitre précédent.

## 2. Résultats théoriques

Malgré des résultats expérimentaux décevants, nous avons souhaité étudier, de manière théorique, les interactions entre polysaccharides branchés et divers lipides afin d'obtenir des informations complémentaires sur leur comportement en milieu aqueux. Pour mettre en œuvre ces calculs, nous avons besoin d'une structure afin de pouvoir réaliser des simulations de dynamique moléculaire. Or il n'existe à ce jour aucune structure tridimensionnelle expérimentale précise de l' $\alpha$ -glucane qui pourrait être utilisée à des fins d'études théoriques.

Toutefois, des expériences récentes ont été menées visant à reproduire un analogue de l' $\alpha$ -glucane en utilisant les enzymes GlgE et GlgB impliquées dans sa biosynthèse<sup>7</sup>.

Les auteurs ont établi un modèle dans lequel est proposé un mécanisme de biosynthèse de ce polysaccharide à partir de maltose (Figure 77) et basé sur l'activité in vitro des deux enzymes : une chaîne d'amylose « C » (chaîne contenant l'extrémité réductrice) est tout d'abord allongée (activité in vitro de l'enzyme GlgE) par pas de deux résidus glucose jusqu'à atteindre un DP moyen de 16 résidus. Ensuite, un oligosaccharide contenant en moyenne 8 résidus est extrait (activité in vitro de l'enzyme GlgB) de cette chaîne « C » à l'extrémité non-réductrice afin d'être branché en  $\alpha$ -(1→6) sur le résidu 3,4 ou 5 de cette dernière, formant ainsi une première chaîne « A » (chaîne terminale non-ramifiée). Cette chaîne « A » est ensuite allongée préférentiellement à la chaîne « C » jusqu'à atteindre un DP moyen de 16 résidus. Le processus se répète donc, la chaîne « A » devenant une chaîne « B » (chaîne ramifiée, ne présentant pas l'extrémité réductrice) jusqu'à l'obtention d'un polysaccharide présentant un très grand nombre de branchements et des chaînes de résidus glucose branchés en  $\alpha$ -(1→4) d'un DP faible (~7-8 résidus).

Il est impossible de savoir si la distribution de ces chaînes, le taux et la position des branchements correspondent à ceux présent au sein de la mycobactérie. Il s'agit cependant des informations les plus précises à ce jour publiées sur la structure primaire de ce polysaccharide et de ce fait ceux sont celles sur lesquelles seront basés les modèles de glucane que nous utiliserons dans la partie théorique de ce chapitre.

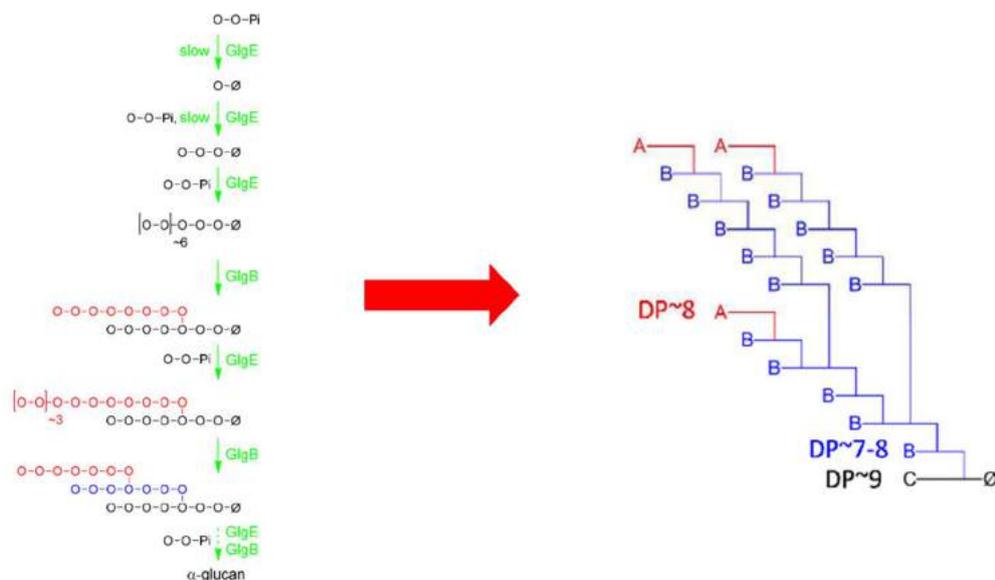


Figure 77: Modèle d' $\alpha$ -glucane (droite) et de sa biosynthèse par les enzymes GlgB et GlgE (gauche). Extrait de Rashid et al. (2016).

## I. Développement de modèle théorique de polysaccharides branchés.

Dans un premier temps, nous voulions vérifier l'impact de la présence de ramifications sur la capacité du polysaccharide à se structurer en polymorphe V. Pour ceci, nous avons modélisé un polysaccharide de 100 résidus avec des chaînes d'une longueur supérieure (20-32 résidus) à celles décrites par Rashid et co-auteurs et peu de ramifications que nous assimilons à un morceau d'amylopectine (Figure 78). Les chaînes devaient être suffisamment longues pour pouvoir impliquer la formation d'un polymorphe V en présence d'une molécule d'acide palmitique ( $DP > 15$ ) et nous voulions un polysaccharide plus grand taille que nos calculs précédents. Deux calculs ont été réalisés : l'un contenant deux molécules d'acide palmitique positionnées à proximité du milieu de chaînes internes « B » et « C » et l'autre contenant ces deux molécules à proximité des chaînes externes « A ».

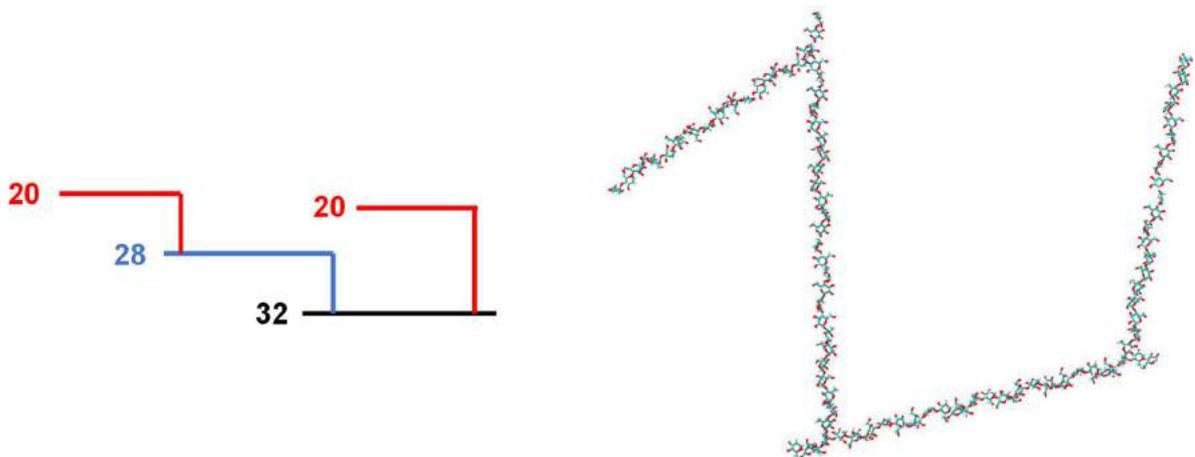


Figure 78: Modèle de polysaccharide branché : topologie (gauche) et structure de départ (droite). Les chiffres correspondent au nombre de résidu de chaque chaîne A (rouge), B(bleu) et C (noir).

Pour obtenir ces structures tridimensionnelles hypothétiques, nous avons développé un logiciel permettant de générer des structures aléatoires similaires à celles présentées ci-dessus. Ce logiciel qui s'inspire de la biosynthèse de polysaccharide in vitro par les enzymes GlgE et GlgB, fonctionne comme ceci : une première chaîne de glucoses branchés en  $\alpha$ -(1 $\rightarrow$ 4) est allongée par ajout de paires de résidus glucose. Quand la chaîne atteint 16 résidus, une probabilité élevée (70%) d'insertion d'une ramification est introduite. Cette ramification se fait en prélevant 7, 8, 9 ou 10 résidus sur la chaîne précédemment allongée, et en la greffant sur les résidus n°3, 4 ou 5 de

cette dernière. La ramification devient la chaîne qui est allongée, et à partir de 16 résidus, le protocole recommence jusqu'à atteindre le nombre total de résidus désirés. Deux modèles de glucane ont ainsi été générés (Figure 79), un contenant 100 résidus (glucane100) et le deuxième en contenant 200 (glucane200).

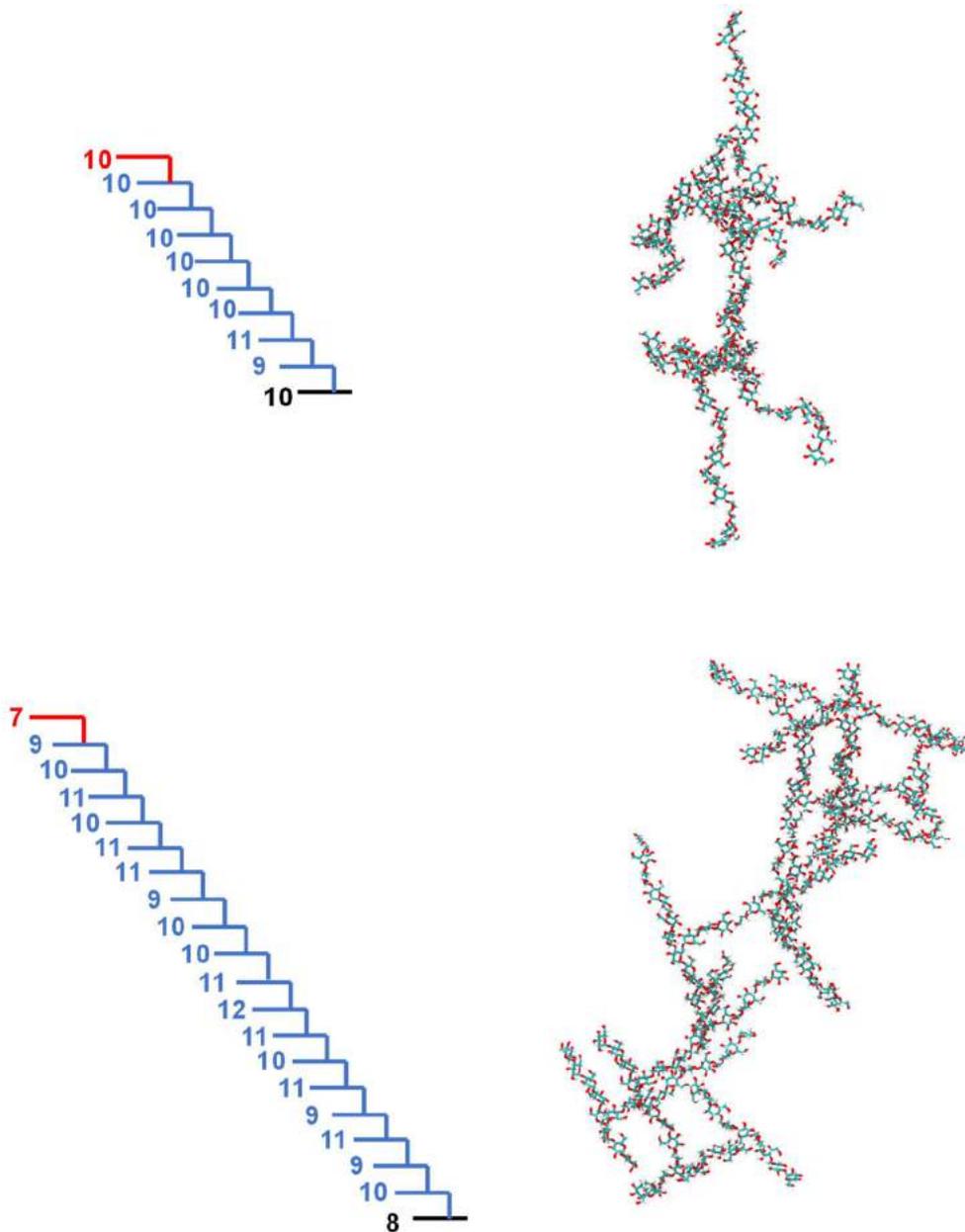


Figure 79: Modèle de glucane contenant 100 (haut) et 200 résidus (bas) : topologie (gauche) et structure de départ (droite). Les chiffres correspondent au nombre de résidus de chaque chaîne A (rouge), B (bleu) et C (noir).

Afin de nous rapprocher du système biologique, nous avons réalisé trois calculs sur le glucane100, contenant le polysaccharide seul, en présence d'une molécule de DMT ou d'une molécule de DIM. Afin de savoir si un effet de taille pouvait être envisagé, nous avons réalisé trois calculs supplémentaires sur des systèmes contenant notre

modèle de glucane 200 : un contenant le polysaccharide seul, un en présence d'une molécule de DMT (évaluation de l'effet de taille) et un contenant deux molécules de DMT (évaluation de l'effet de taille, même stœchiométrie que le calcul glucane100 en présence d'une molécule de DMT). Finalement, nous planifions une expérience de complexation en présence d' $\alpha$ -glucane et de différentes concentrations d'acide oléique. Nous voulions en effet savoir si l'utilisation d'un lipide présent à l'état liquide à température ambiante pouvait éviter son auto-précipitation et donc promouvoir la formation de complexes d'inclusion au sein de polysaccharides branchés. En accord avec ces expériences, nous avons donc réalisé deux calculs contenant une molécule de glucane200 en présence de 5 molécules d'acide oléique ou de 10 molécules d'acide oléique.

Les calculs de dynamique moléculaires ont été réalisés comme décrits dans le chapitre précédent, sur une durée de 50ns pour l'amylopectine, de 150ns pour ceux contenant le glucane100 et de 300ns pour ceux contenant le glucane200.

## II. Résultats

Concernant le modèle d'amylopectine, lorsque les molécules d'acide palmitique sont positionnées à proximité des chaînes internes, la formation de complexes d'inclusion n'est pas observée (Figure 80, A). Le polysaccharide tend à se replier sur lui-même sans former de structure particulière, et les lipides sont amenés à diffuser dans le solvant. En contrepartie, lorsque ces molécules sont placées à proximité des chaînes externes (Figure 80, B), la formation d'un complexe d'inclusion est observée pour l'une d'entre elles. Il en résulte une distribution des angles dièdres  $\varphi$  et  $\psi$  présentant une contribution plus importante aux alentours de  $\varphi=115^\circ$  et  $\psi=105^\circ$  (Figure 81), ce qui est cohérent avec la formation d'une structure hélicoïdale.

Ces résultats semblent indiquer que seules les parties les plus externes de ce type de polysaccharide sont susceptibles d'encapsuler des lipides, ce qui est cohérent avec ce qui a été proposé dans la littérature<sup>2</sup>. La formation de polymorphe au sein des chaînes internes nécessitent le déplacement d'une grande partie de l'amylopectine et il est probable que ceci demande un apport d'énergie trop important. En revanche les

chaines extérieures ne nécessitent le déplacement que d'une à deux dizaines de résidus, ce qui peut engendrer la formation rapide de complexes d'inclusion.

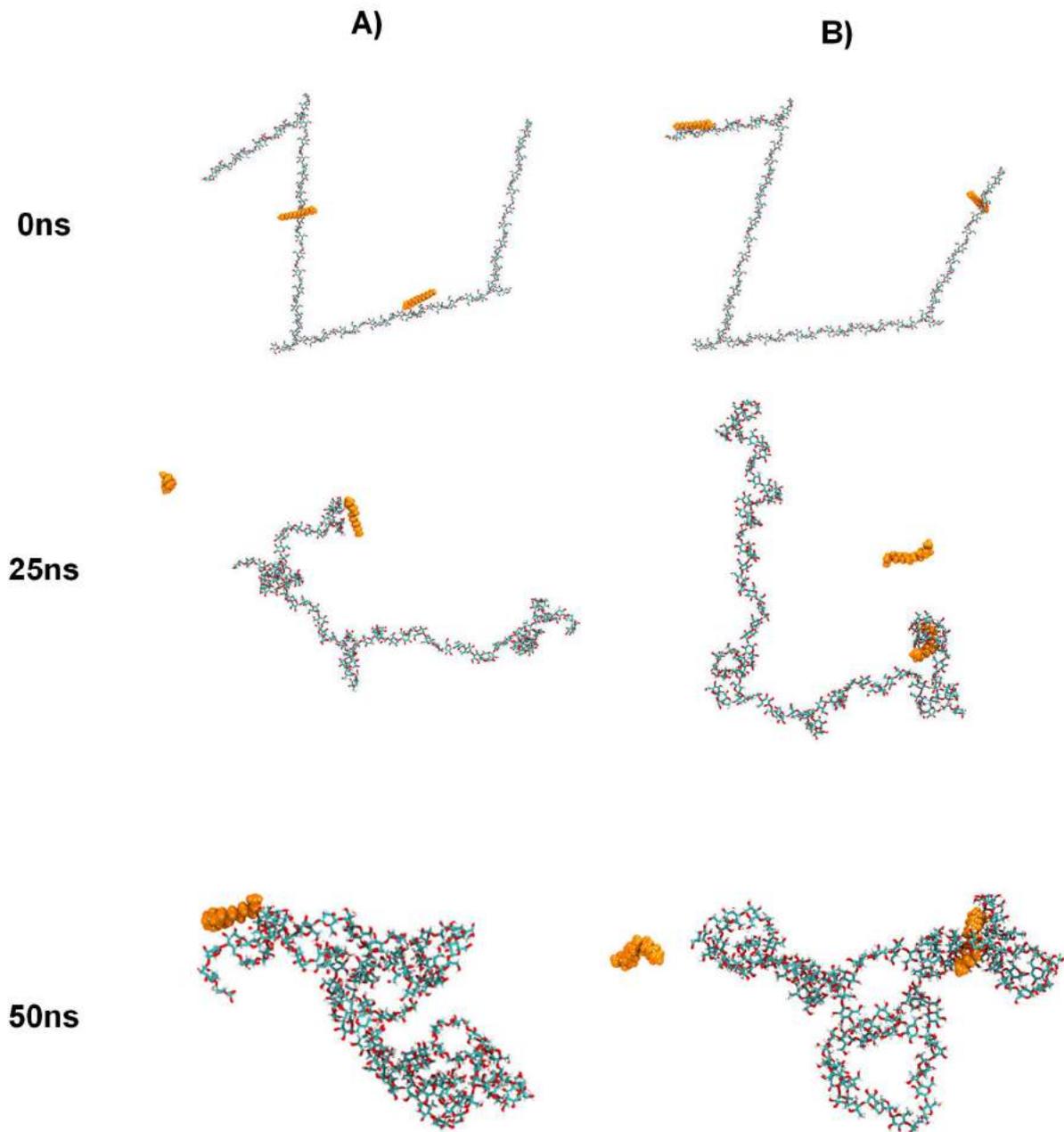


Figure 80: Structures extraites des trajectoires de calcul des systèmes contenant une molécule d'amylopectine de 100 résidus en présence de deux molécules de PAL proches des chaînes internes (A) ou externes (B).

Il faut aussi préciser qu'il est possible que la limite de notre méthode, à savoir la dynamique moléculaire « tout atome » soit atteinte ici. Pour espérer observer des déplacements d'un nombre de résidu aussi grand, des simulations utilisant des modèles de type gros-grain pourraient être mieux adaptées.

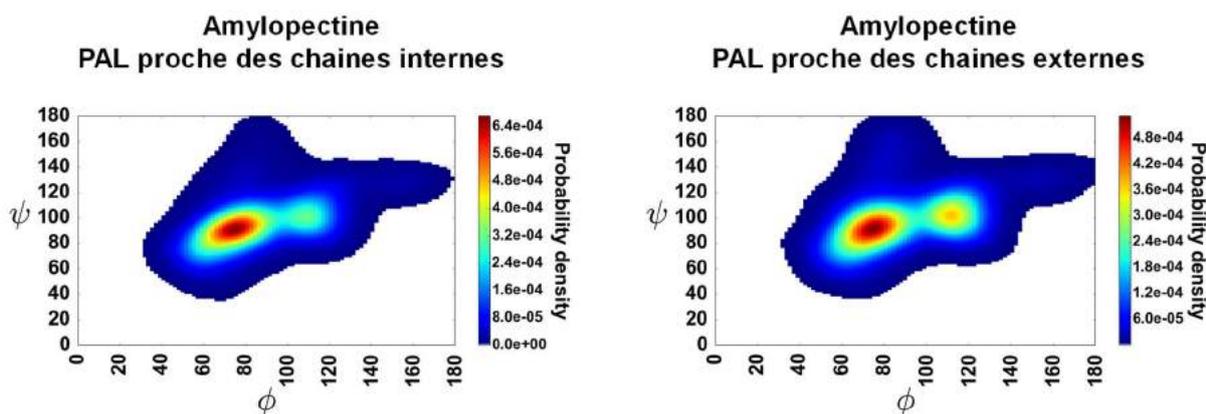


Figure 81 : Diagramme de type Ramachandran, représentant les angles dièdres  $\phi$  et  $\psi$  de la liaison glycosidique pour tous les résidus d'amylopectine en présence de deux molécules de PAL situées à proximité des chaînes internes ou externes.

Elle reste cependant adaptée à nos modèles de glucane, qui présentent des longueurs de chaînes plus courtes et beaucoup plus de ramifications, ce qui devrait engendrer une flexibilité plus grande. Ces structures sont à priori moins susceptibles que les structures d'amylose d'engendrer des complexes d'inclusion, cependant nous voulions vérifier s'il était possible d'observer une structuration locale du polysaccharide branché en présence de lipides.

Concernant le glucane100, les calculs ne permettent pas d'observer la formation d'une hélice similaire à celle observée sur l'amylose ou sur notre modèle d'amylopectine (Figure 82). En effet, les parties linéaires, ne contenant aucune ramification branchée en  $\alpha$ -(1 $\rightarrow$ 6), sont trop courtes pour permettre une telle structuration locale, ce qui est cohérent avec les résultats obtenus dans le premier chapitre de cette thèse. Dans ce dernier, nous avons démontré que lors de la formation du polymorphe  $V_X$ , des liaisons hydrogène sont établies entre le groupe hydroxyle présent sur le carbone  $C_6$  d'un résidu  $[n]$  et les groupes hydroxyles présents sur les carbones  $C_2$  et  $C_3$  situés sur le résidu  $[n+X]$ . La présence de ramifications branchées en  $\alpha$ -(1 $\rightarrow$ 6) empêche donc la formation d'un tel allomorphe et le nombre élevé de ramifications de nos modèles de glucane rend impossible leur observation. En revanche, les lipides peuvent engendrer une structuration locale au sein de notre modèle de glucane. Nos diagrammes de Ramachandran démontrent ceci au travers de l'intensité du maximum local situé aux angles  $\phi=115^\circ$  et  $\psi=105^\circ$  (Figure 83). Les calculs sur le glucane100 montrent que la présence d'une molécule de DIM tend à augmenter la distribution d'angles dièdres dans cette région de la carte de Ramachandran, et donc implique une structuration locale du polysaccharide. A contrario, la présence d'une molécule de DMT dans sa

conformation « membranaire » ne permet pas d'observer de changement particulier sur la structuration locale du glucane100 en comparaison avec les calculs ne contenant que le polysaccharide.

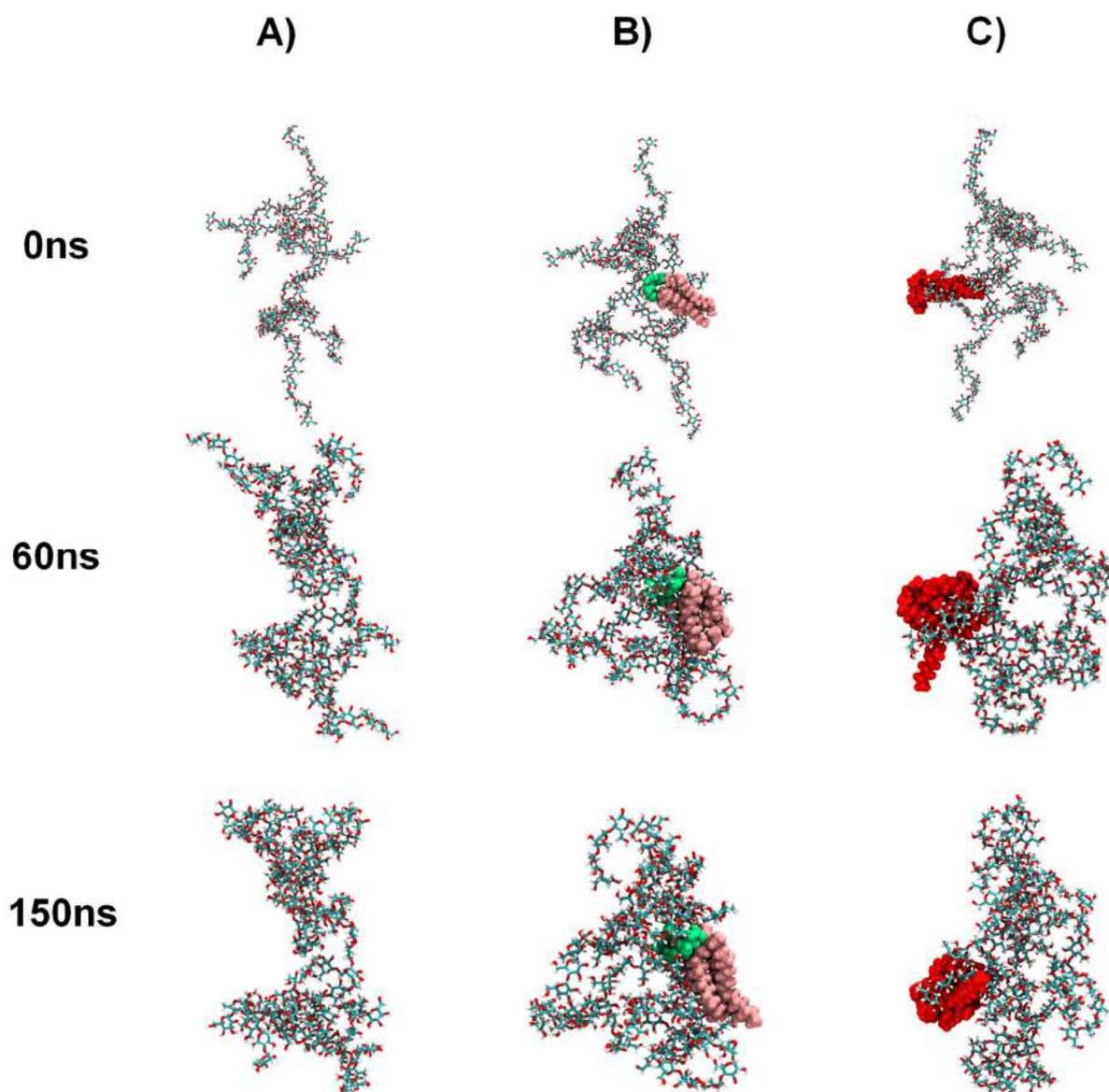


Figure 82: Structures extraites des trajectoires de calcul des systèmes contenant une molécule de glucane100 seul (A) en présence d'une molécule de DMT (B) ou d'une molécule de DIM (C).

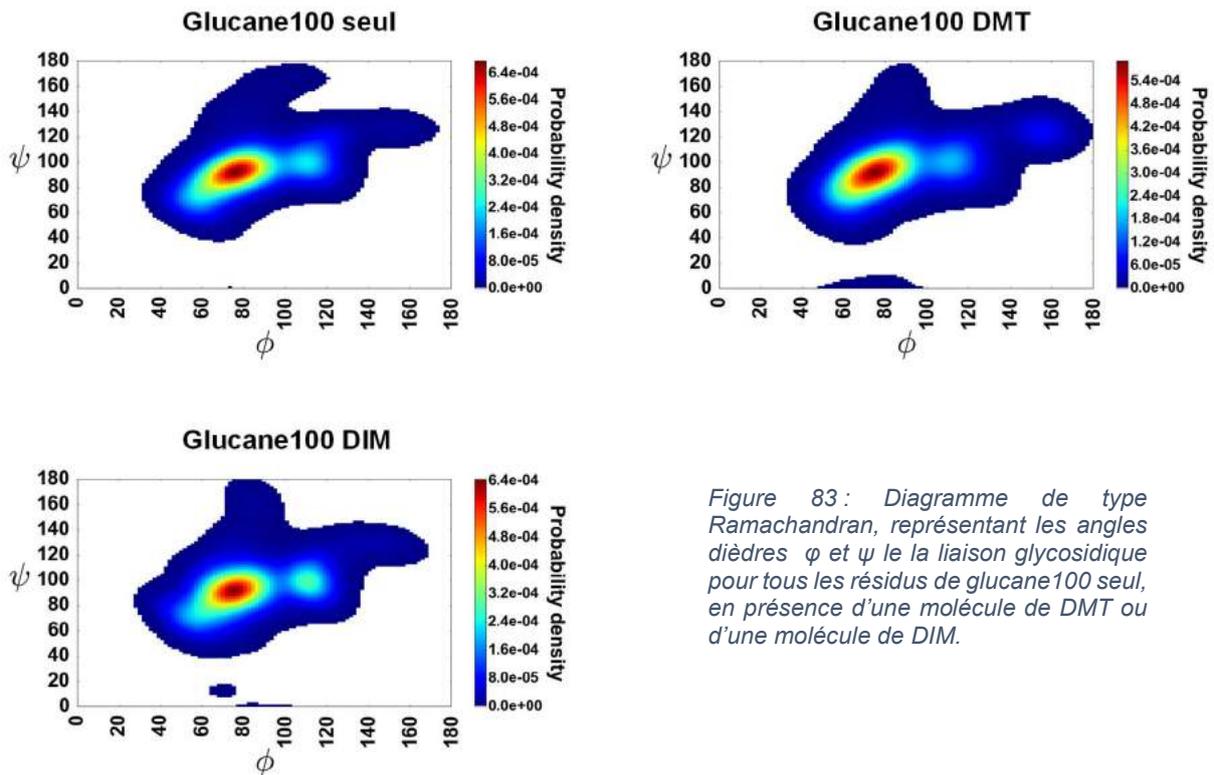


Figure 83 : Diagramme de type Ramachandran, représentant les angles dièdres  $\phi$  et  $\psi$  le la liaison glycosidique pour tous les résidus de glucane100 seul, en présence d'une molécule de DMT ou d'une molécule de DIM.

Dans ces simulations, nous pouvons observer que le DIM comme le DMT ont tendance à se situer à la surface du polysaccharide. Afin de vérifier si cette tendance n'était pas liée à un effet de taille du polysaccharide, nous avons réalisés des simulations similaires en présence d'un glucane deux fois plus grand (glucane200, Figure 84). Dans ces nouvelles conditions, le DMT reste à nouveau en surface du polysaccharide. De plus le diagramme de type Ramachandran ne permet pas de témoigner d'un effet de structuration locale du polysaccharide (Figure 85). L'ajout d'une seconde molécule de DMT, afin de retrouver le même rapport stœchiométrique entre le nombre de lipides et le nombre de résidus composant le polysaccharide, aboutit de nouveau à un positionnement des lipides en surface du polysaccharide. En revanche l'augmentation des contributions présentes en conformation hélicoïdale est observée (Figure 11), ce qui laisse penser que des lipides distribués au sein du polysaccharide peuvent amener une quantité de parties hélicoïdales plus importante.

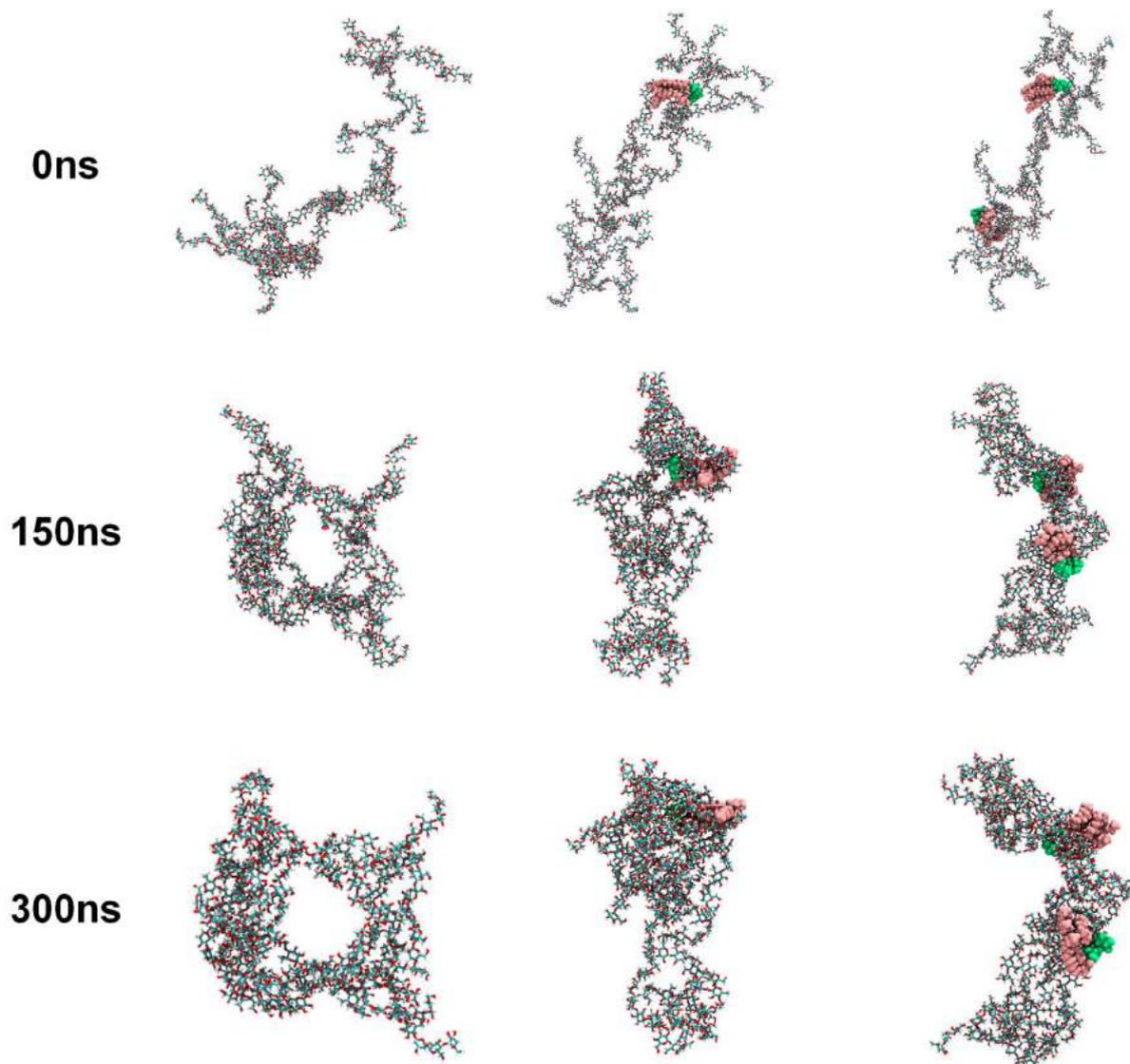


Figure 84: Structures extraites des trajectoires de calcul des systèmes contenant une molécule de glucane200 seul (A) en présence d'une (B) ou deux (C) molécules de DMT.

Afin de vérifier un effet synergique des lipides, et comme nous avons prévu de réaliser les expériences de complexation correspondantes, nous avons effectué un calcul dans lequel 5 molécules d'acide oléique étaient distribuées de manière arbitraire au sein du glucane200. Ces molécules ne contenant qu'une seule chaîne lipidique nous espérons qu'il serait plus facile à ces dernières d'être incluses dans le polysaccharide. C'est en effet ce qui est observé en présence de cinq molécules d'acide oléique (Figure 86). Ceci a pour effet d'augmenter significativement la quantité de parties hélicoïdales générées au sein du polysaccharide, comme en atteste l'augmentation de la distribution des angles dièdres aux alentours de  $\varphi=115^\circ$  et  $\psi=105^\circ$  (Figure 85). Une

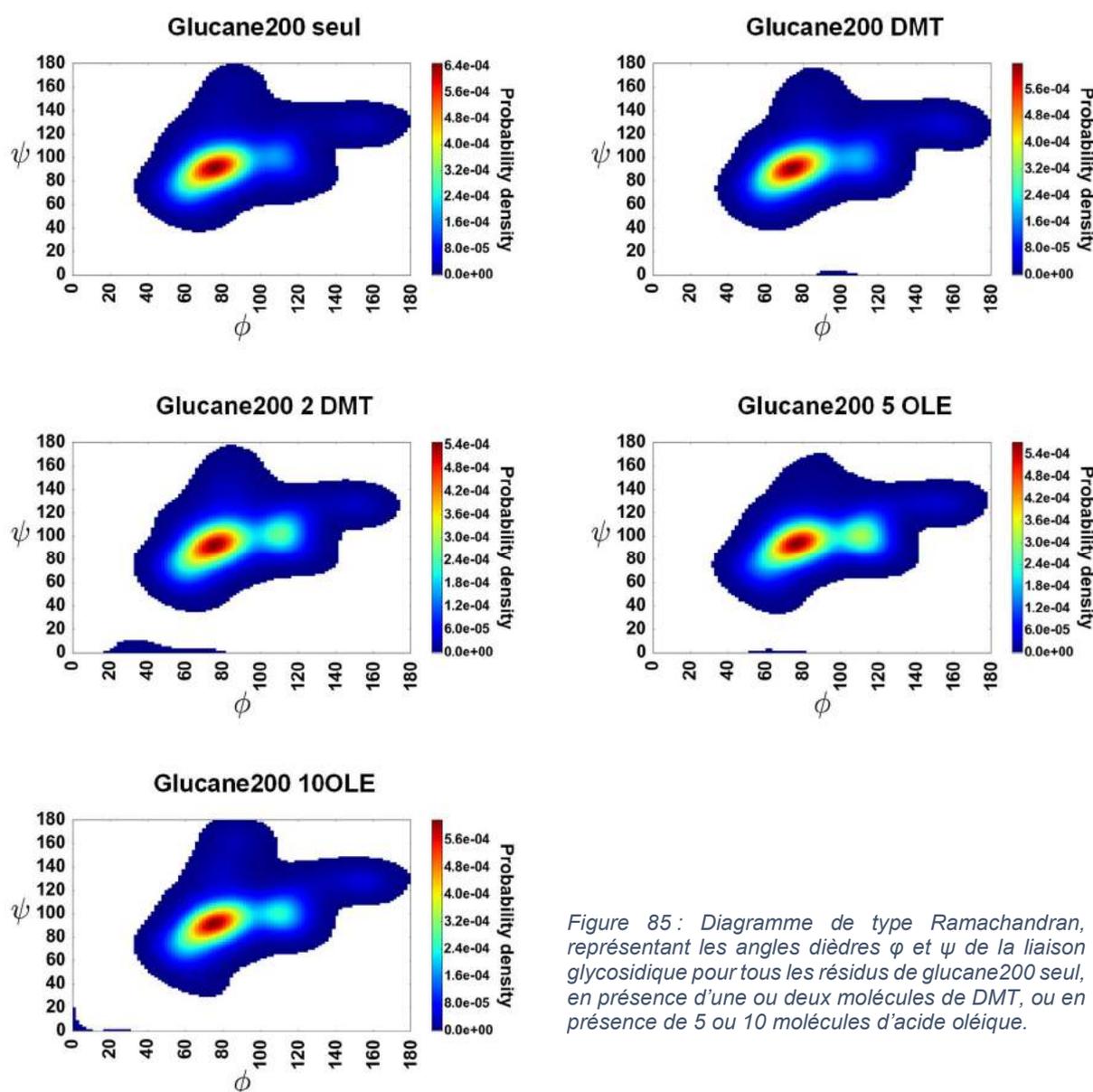


Figure 85 : Diagramme de type Ramachandran, représentant les angles dièdres  $\phi$  et  $\psi$  de la liaison glycosidique pour tous les résidus de glucane200 seul, en présence d'une ou deux molécules de DMT, ou en présence de 5 ou 10 molécules d'acide oléique.

trop forte concentration de lipides est cependant délétère pour la structuration intrinsèque du polysaccharide. En effet, le calcul contenant dix molécules d'acide oléique aboutit à l'agrégation de ces dernières et donc à une séparation de phase ne permettant plus la structuration du polysaccharide. Il en résulte une atténuation du maximum de distribution des angles dièdres à  $\phi=115^\circ$  et  $\psi=105^\circ$  (Figure 9). Ces calculs entrent par ailleurs en parfait accord avec le mécanisme de compétition que nous avons décrit dans le chapitre 2 de ce travail de thèse concernant l'amylose. Il est donc possible d'envisager un mécanisme similaire au sein de polysaccharides branchés, dans lesquels la formation de complexes d'inclusion est moins favorable qu'au sein de l'amylose en raison du DP moins élevé des parties du polysaccharide uniquement branchées en  $\alpha$ -(1 $\rightarrow$ 4).

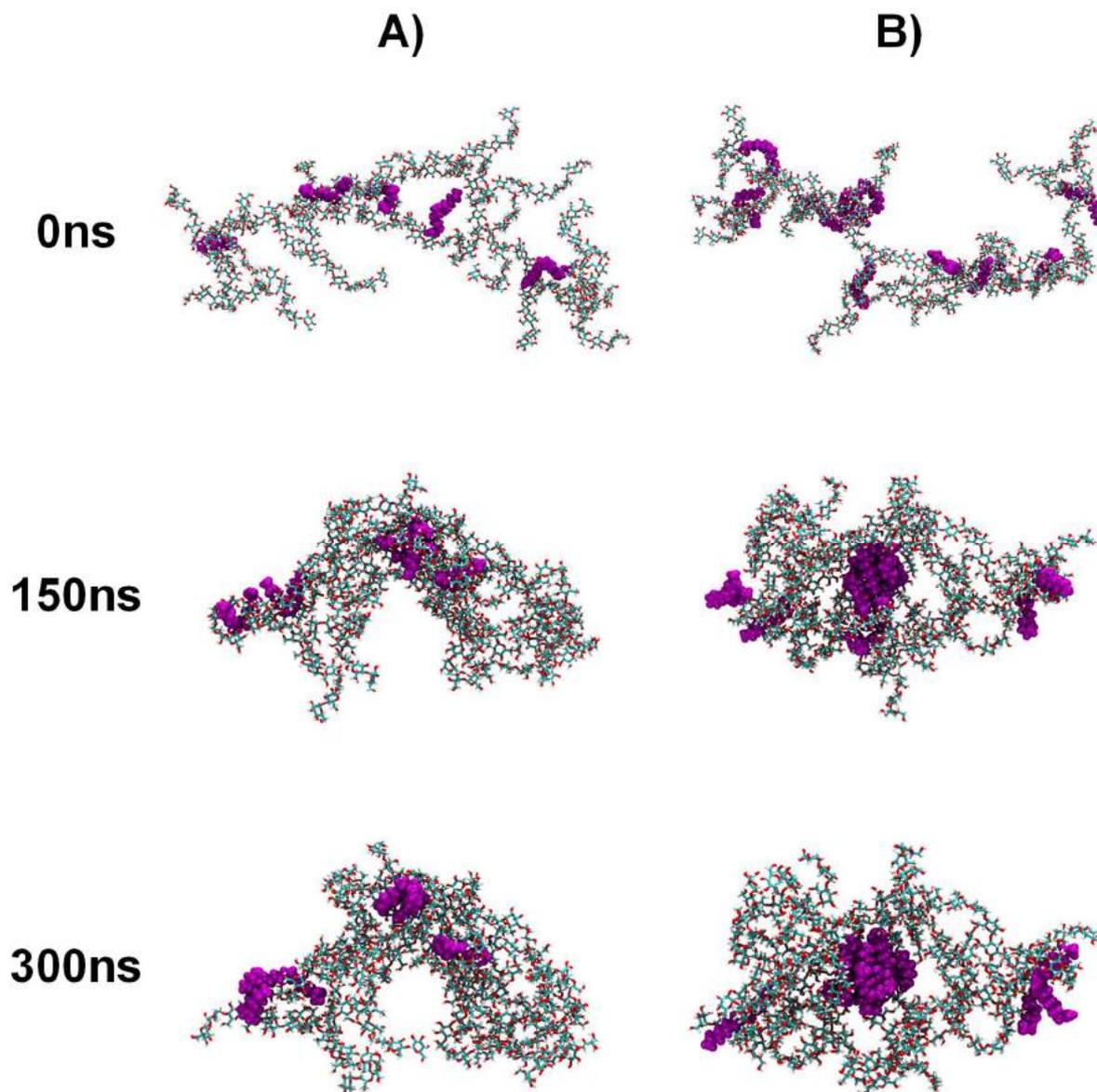


Figure 86: Structures extraites des trajectoires de calcul des systèmes contenant une molécule de glucane200 en présence de 5 (A) ou 10 (B) molécules d'acide oléique.

### 3. Conclusion, expériences futures.

Il nous a été impossible de caractériser expérimentalement la présence de polymorphe V au sein de polysaccharides ramifiés. Nous ne pouvons savoir à ce stade si cela est dû à notre protocole expérimental (basé sur la précipitation des complexes polysaccharide/lipides), notre méthode d'analyse (la RMN en phase solide) ou à l'impossibilité de former de tels complexes au sein de polysaccharides ramifiés. Les expériences menées par Beeren et co-auteurs sur l'amylopectine ont été réalisées en

milieu aqueux, là où nous nous efforçons de provoquer la précipitation de notre soluté. Afin de savoir si la précipitation en elle-même engendre une séparation de phases, des expériences pourraient être menées en utilisant des lipides qui ne précipitent pas à température ambiante, tels que l'acide oléique ou le DOPC. Si en utilisant ce type de lipide, il est toujours impossible d'observer une structuration des polysaccharides, il faudra alors envisager un autre protocole de synthèse et certainement d'autres méthodes d'étude de ces complexes.

Nos calculs théoriques montrent qu'il est possible d'observer de la structuration locale au sein des modèles de glucane en présence de lipides mycobactériens en milieu aqueux. Il est possible que des formes étendues des DMT et du DIM puissent aussi permettre d'observer une structuration plus nette, comme cela a été le cas dans le chapitre n°2 de cette thèse. De plus, ceux réalisés sur les systèmes contenant des molécules d'acide oléique démontrent cependant que la concentration est un des facteurs qu'il est nécessaire d'étudier aussi avec soin. Une concentration trop élevée de lipides (cf. simulation avec 10 molécules d'acide oléique) amènerait une séparation de phase (formation de micelles de lipides) ne permettant pas d'observer de structuration locale. Ceci est en accord avec la faible quantité de lipides (1-6%) retrouvés au sein de la capsule des mycobactéries<sup>3</sup>. En effet, l' $\alpha$ -glucane contient majoritairement des chaînes courtes et ramifiées, ce qui ne promeut pas la formation de polymorphe V. En revanche, même une faible quantité de lipides permet de structurer le polysaccharide en solution aqueuse, comme nous pouvons l'observer dans nos diagrammes de type Ramachandran. Ce dernier point peut permettre d'expliquer la structuration de la capsule autour de la paroi de la bactérie. Nous ne pouvons donc pas affirmer, ni infirmer de la présence de ces structures au sein de systèmes biologiques plus complexes.

Concernant la suite de cette étude, nous allons donc devoir améliorer notre protocole expérimental et/ou nos méthodes d'études, car ces deux paramètres ne semblent plus être adaptés à l'étude de ces complexes. D'un point de vue théorique, notre modèle chimique mais aussi notre méthodologie doivent être améliorés, que ce soit au niveau des structures primaires des polysaccharides, des structures initiales utilisées dans nos calculs ou des durées des trajectoires et ceci afin de pouvoir échantillonner un peu plus l'espace des configurations possibles au sein de ce type de complexes. Pour se

faire, il sera important de caractériser expérimentalement de manière plus approfondie la structure de l' $\alpha$ -glucane, mais ceci sort du contexte de ce travail de thèse.

## Références

- (1) Beeren, S. R.; Meier, S.; Hindsgaul, O. Probing Helical Hydrophobic Binding Sites in Branched Starch Polysaccharides Using NMR Spectroscopy. *Chem. - A Eur. J.* **2013**, *19* (48), 16314–16320. <https://doi.org/10.1002/chem.201302213>.
- (2) Beeren, S. R.; Hindsgaul, O. Nature's Dendrimer: Characterizing Amylopectin as a Multivalent Host. *Angew. Chemie - Int. Ed.* **2013**, *52* (43), 11265–11268. <https://doi.org/10.1002/anie.201305132>.
- (3) Lemassu, A.; Daffe, M. Structural Features of the Exocellular Polysaccharides of Mycobacterium Tuberculosis. *Biochem. J.* **1994**, *297* (2), 351–357. <https://doi.org/10.1042/bj2970351>.
- (4) Nakamura, Y. *Starch: Metabolism and Structure*; Nakamura, Y., Ed.; Springer Japan: Tokyo, 2015; Vol. 369. <https://doi.org/10.1007/978-4-431-55495-0>.
- (5) Rondeau-Mouro, C.; Veronese, G.; Buleon, A. High-Resolution Solid-State NMR of B-Type Amylose. *Biomacromolecules* **2006**. <https://doi.org/10.1021/bm060330x>.
- (6) Gidley, M. J.; Bociek, S. M. <sup>13</sup>C CP/MAS NMR Studies of Amylose Inclusion Complexes, Cyclodextrins, and the Amorphous Phase of Starch Granules: Relationships Between Glycosidic Linkage Conformation and Solid-State <sup>13</sup>C Chemical Shifts. *J. Am. Chem. Soc.* **1988**, *110* (12), 3820–3829. <https://doi.org/10.1021/ja00220a016>.
- (7) Rashid, A. M.; Batey, S. F. D.; Syson, K.; Koliwer-Brandl, H.; Miah, F.; Barclay, J. E.; Findlay, K. C.; Nartowski, K. P.; Khimyak, Y. Z.; Kalscheuer, R.; Bornemann, S. Assembly of α-Glucan by GlgE and GlgB in Mycobacteria and Streptomyces. *Biochemistry* **2016**, *55* (23), 3270–3284. <https://doi.org/10.1021/acs.biochem.6b00209>.



## Conclusion et perspectives



Dans ce travail de thèse, nous avons étudié des interactions entre polysaccharides et lipides extraits de divers systèmes biologiques en utilisant une stratégie combinant RMN  $^{13}\text{C}$  du solide, dynamique moléculaire et calculs de déplacements chimiques par méthode de chimie quantique.

Au commencement de ce travail de thèse en 2017, les interactions entre polysaccharides et lipides formaient déjà un vaste champ de recherche dans le cadre des deux polysaccharides composant l'amidon : l'amylose, composée principalement de résidus glucose branchés en  $\alpha$ -(1 $\rightarrow$ 4) et l'amylopectine, composée de résidus glucose branchés en  $\alpha$ -(1 $\rightarrow$ 4) et de ramifications branchées en  $\alpha$ -(1 $\rightarrow$ 6)<sup>1</sup>. L'amylose est notamment très connue pour sa capacité à former des complexes d'inclusion de structure hélicoïdale en présence de lipides, appelés polymorphe V<sup>2-7</sup>. L'amylopectine était plus connue pour les doubles hélices qu'elle formait en milieu aqueux, appelées polymorphe A ou B<sup>8-11</sup> selon les conditions de cristallisation. Il a aussi été démontré qu'il était possible de former des complexes d'inclusion en étudiant ce polysaccharide en présence de lipides<sup>12-14</sup>. Si les complexes d'inclusion étaient déjà largement caractérisés en présence d'acides gras ou de petites molécules, à notre connaissance, ils avaient rarement été caractérisés en présence de lipides contenant de multiples chaînes, tels que des phospholipides<sup>15</sup>.

Nous voulions étudier des interactions entre polysaccharides et lipides extraits de systèmes biologiques tels que des mycobactéries. Notre approche se devait donc d'être transférable entre des structures déjà connues, largement étudiées et pouvant former des phases ordonnées, voire cristallines et des systèmes beaucoup plus hétérogènes. Pour ceci, notre choix s'est porté sur la RMN à l'état solide couplée à la dynamique moléculaire et à la chimie quantique. La RMN du solide avait déjà été utilisée pour étudier les différentes structures de l'amylose en présence ou absence de lipide et permet notamment d'étudier des échantillons hétérogènes<sup>11,16</sup>. La dynamique moléculaire était relativement peu utilisée, mais il avait été démontré qu'elle était un outil pertinent pour l'étude de ce type de complexes<sup>17,18</sup>. Finalement les méthodes de chimie quantique avaient été peu utilisées sur des polysaccharides, et à priori jamais sur des systèmes contenant plus d'une dizaine de résidus<sup>19</sup>, malgré l'augmentation constante de la puissance de calcul disponible.

Le travail réalisé pendant ces trois ans peut être divisé en trois parties : 1) établissement d'une stratégie d'étude théorique, 2) étude théorique et expérimentale

de complexes d'amylose en présence d'acides gras, de phospholipides et de lipides mycobactériens et 3) étude préliminaire réalisée sur des systèmes contenant des polysaccharides ramifiés en présence de lipides d'origines diverses.

## Résumé des travaux réalisés

### Etablissement de la stratégie d'étude théorique.

Nous voulions établir une méthodologie théorique robuste et applicable à des systèmes contenant des polysaccharides et lipides complexes. Pour cela, nous avons étudié des complexes d'amylose en présence d'acide palmitique, profitant ainsi de la littérature expérimentale disponible sur le sujet. Notre approche utilisait la dynamique moléculaire classique couplée à des calculs de déplacements chimiques au niveau DFT.

Nous avons ainsi pu montrer que le champ de force que nous utilisons, GLYCAM06, nécessitait une modification afin de pouvoir générer des structures de polymorphe V de complexe amylose/acide palmitique stables et cohérentes avec les données expérimentales. L'utilisation de ce champ de force modifié nous a, par la suite, permis de démontrer que la taille de l'amylose influait sur la stabilité de la structure en polymorphe V.

Grâce aux calculs des déplacements chimiques  $^{13}\text{C}$  par une approche de chimie quantique, nous avons pu démontrer que la taille du polysaccharide influençait aussi le spectre RMN. Finalement, nous avons aussi prouvé qu'il était important d'inclure la dynamique du système dans les calculs RMN afin d'obtenir une meilleure reproduction des spectres expérimentaux. La stratégie théorique utilisée pour le reste du travail de thèse s'est donc basée sur l'utilisation de dynamique moléculaire classique afin d'étudier le comportement des systèmes d'intérêts et sur des calculs de RMN par DFT sur des structures issues des dynamiques moléculaires pour comparaison avec les spectres expérimentaux.

Cette stratégie a ensuite été appliquée sur un modèle de polymorphe B d'amylose, qui est caractérisé expérimentalement par deux jeux de signaux RMN, correspondant à deux types de résidus glucose différents au sein de la structure. Elle a permis de

montrer que le réseau de liaisons hydrogène établi était déterminant pour le calcul de RMN et que la différenciation de deux types de résidus ne pouvait être observée que lorsque l'on prenait en compte l'empilement cristallin dans le calcul.

### Etude de complexes d'amylose en présence de lipides contenant plusieurs chaînes hydrophobes.

Nous avons ensuite étudié des complexes d'amylose en présence de lipides classiques (acide palmitique, acide oléique, DPPC, POPC, DOPC) et uniquement retrouvés au sein de systèmes bactériens (BET, Myco. A., MMT, DMT et DIM) présentant d'une à quatre chaînes acyles, de longueur variable et plus ou moins insaturées ou méthylées.

L'étude expérimentale par RMN du solide, a permis de prouver que des complexes d'inclusion pouvait être formés en présence de phospholipides. Nous avons de plus mis en évidence un mécanisme de compétition entre l'auto-agrégation des lipides et la formation de complexes d'inclusion. Les équilibres en jeu dans cette compétition dépendent de la structure primaire de la chaîne hydrophobe du lipide étudié. En effet, la présence d'insaturations au sein des parties hydrophobes permet d'augmenter la quantité de polymorphe V dans l'amylose. La partie théorique relative à ces complexes nous a permis de démontrer la possibilité de formation de complexes d'inclusion intrinsèquement hétérogènes d'un lipide à l'autre. Cependant, le signal RMN qui en résulte atteste toujours d'une structuration hélicoïdale de l'amylose indiquant ainsi qu'il est possible de retrouver ces structures dans nos échantillons expérimentaux.

Par la suite, nous avons montré que la complexation de lipides contenant jusqu'à 4 chaînes hydrophobes était possible. Les structures primaires des lipides bactériens utilisés présentaient des similarités avec celles des phospholipides, et de ce fait nous avons pu vérifier que ces échantillons s'incluaient bien dans le mécanisme de compétition défini ci-dessus. L'utilisation des acides mycoliques et de leur analogue de synthèse, le BET, démontre qu'il est possible de réaliser des complexes similaires à ceux obtenus sur les acides gras et phospholipides. Le pourcentage de polymorphe V observé dans l'échantillon contenant du MMT est inférieur à celui retrouvé dans l'échantillon d'acides mycoliques, ce qui indique que l'ajout d'un groupe tréhalose ne promeut pas la formation de polymorphe V. Enfin l'ajout d'une seconde chaîne

mycolique au MMT, formant ainsi le DMT, permet une augmentation de la formation de polymorphe V dans nos échantillons. Les résultats théoriques obtenus sur ces lipides ont montré que la présence d'un groupement de haute polarité était importante pour la formation de ces complexes, alors que l'ajout d'un groupe tréhalose était délétère à cette dernière en raison des liaisons hydrogène établies entre celui-ci et l'amylose. Les lipides pourvus de 4 chaînes peuvent être présents sous une forme étendue qui implique de ce fait une surface hydrophobe plus grande, facilitant ainsi la formation de complexes d'inclusion. Enfin, nous avons été aussi capables de démontrer que malgré son hétérogénéité de structure, le lipide DIM pouvait être complexé avec de l'amylose, induisant une structuration hélicoïdale de cette dernière.

### [Etude d'interactions entre polysaccharides ramifiés et lipides divers.](#)

Finalement, des tentatives de complexation ont été réalisées en présence de polysaccharides ramifiés afin de se rapprocher autant que possible du système présent à la surface des mycobactéries.

Pour l'étude de ces polysaccharides branchés, le protocole que nous utilisons a dû être modifié, pour tenir compte des différences de solubilité du soluté étudié. Nous n'avons toutefois pas pu témoigner de la formation expérimentale de complexes d'inclusions au sein de tels polysaccharides.

Nous avons développé, théoriquement, des structures hypothétiques plausibles pouvant être représentatives de celles de l'amylopectine et de l' $\alpha$ -glucane. Via une approche de dynamique moléculaire, nous avons observé théoriquement la présence de complexes sur les chaînes les plus extérieures de notre modèle d'amylopectine. Concernant le glucane, les chaînes ne présentent pas un nombre de résidus suffisamment élevé pour pouvoir former des hélices de complexation stables. Cependant, une structuration hélicoïdale locale induite par les lipides a tout de même pu être observée au cours des simulations. La présence d'une trop grande quantité de lipides engendre en revanche la formation d'une séparation de phase, ces derniers s'agrégeant au milieu du polysaccharide et ne pouvant ainsi impliquer de changement de structure du polysaccharide. Ces résultats sont cohérents avec le mécanisme de compétition que nous avons postulé dans le chapitre 2 concernant l'amylose, et des

expériences de complexations complémentaires seront réalisées afin de vérifier la validité de nos calculs.

## Perspectives d'études.

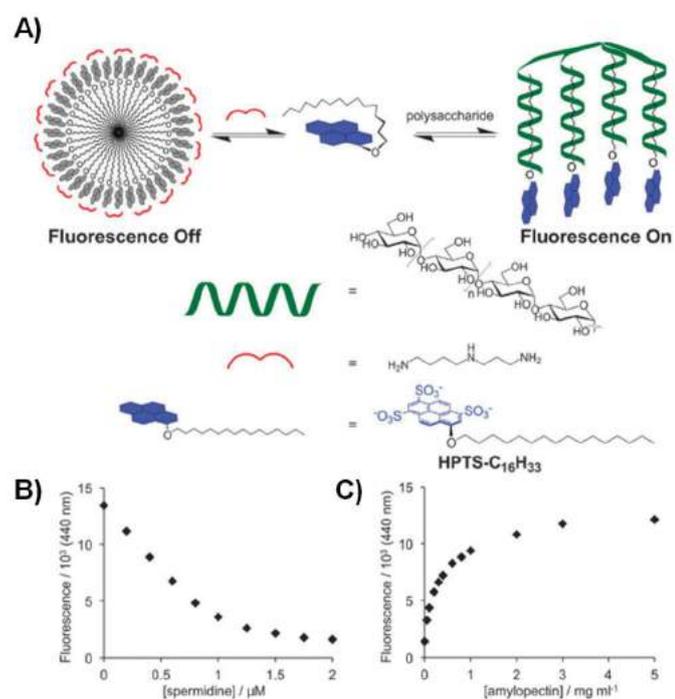
Comme nous l'avons déjà mentionné, l'étude de complexes d'inclusion amylose-lipides a surtout été réalisée en présence de lipides ne présentant qu'une seule chaîne aliphatique. Nous avons apporté la preuve que ce type de complexes pouvait être observé en présence de lipides bien plus complexes contenant deux à quatre chaînes. Ce type de complexe devrait ouvrir de nouvelles perspectives d'études concernant les complexes d'inclusion d'amylose dans l'industrie alimentaire, mais aussi dans le domaine pharmaceutique.

Dans la continuité de ce travail, nous aimerions réaliser des échantillons contenant des lipides issus de *Mycobacterium tuberculosis* autres que le DIM : DMT, MMT et acides mycoliques à longue chaînes carbonées - C60 à C90, présentant des fonctions oxygénées, ainsi que d'autres facteurs de virulence, tel que les glycolipides phénoliques. Comme nous l'avons déjà mentionné, des travaux réalisés à l'IPBS ont démontré qu'il était possible de retrouver des lipides issus de MTB dans la membrane des macrophages<sup>20</sup>. La principale difficulté rencontrée lors de ce type d'étude réside dans la façon de solubiliser ces derniers en milieu aqueux. Les complexes d'inclusion réalisés en présence d'amylose pourraient former un vecteur viable pour faciliter leur transport. Dans le cas où l'amylose seule ne présenterait pas une solubilité assez élevée, il a déjà été démontré que l'amylopectine pouvait former des complexes d'inclusion en milieu aqueux. Une autre alternative pourrait résider dans l'utilisation d'amylose propionylée, une forme légèrement modifiée du polysaccharide qui a déjà permis de vectoriser des médicaments hydrophobes sur modèle murin<sup>21</sup>. Ces complexes pourraient en outre être réalisés avec des facteurs de virulence lipidiques de MTB en vue de l'amélioration du vaccin BCG. Les DMTs sont connus pour activer des réponses immunitaires au travers d'une lectine appelée MINCLE<sup>22</sup>, et de ce fait des analogues de ces lipides ont été récemment proposés comme nouveaux adjuvants vaccinaux<sup>23</sup>. Encapsuler ces lipides pourrait donc permettre de faciliter leur biodisponibilité et donc d'augmenter l'efficacité du vaccin BCG.

Les études les plus importantes pour la suite du travail sur les polysaccharides mycobactériens pourraient concerner, à mon sens, la caractérisation structurale de l' $\alpha$ -glucane. Les informations dont nous disposons à son sujet sont des valeurs moyennes, et des distributions de longueur de chaîne et de taux de branchement. Il est important de savoir si sa structure primaire est homogène ou diverse afin de savoir si des complexes d'inclusion peuvent être envisagés en son sein. De plus, ce type d'informations permettrait d'établir des modèles théoriques plus précis en vue d'utilisation de dynamique moléculaire. La caractérisation par des méthodes biochimiques, telles que des digestions enzymatiques suivies de la caractérisation de des tailles des polysaccharides résultant, pourrait être envisagées.

A ce dessein, l'utilisation du lipide utilisé par Beeren et co-auteurs (HPTS) pourrait s'avérer d'une grande utilité<sup>12-14</sup>. Le HPTS présente la particularité d'être fluorescent et sa fluorescence disparaît en solution aqueuse quand de la spermidine y est ajoutée. Cette molécule semble stabiliser une forme micellaire du lipide qui disparaît lorsque l'on ajoute de l'amylopectine. En effet le polysaccharide encapsule le HPTS, qui retrouve ses propriétés de fluorescence. Les auteurs affirment que ce mécanisme n'est observé qu'en

présence de chaînes branchées en  $\alpha$ -(1 $\rightarrow$ 4). Des DP compris entre 8 et 15 seraient nécessaires afin de retrouver la fluorescence de l'échantillon. Utiliser ce lipide en présence d' $\alpha$ -glucane permettrait d'étudier leurs interactions. Par la suite, s'il s'avère que l'on puisse retrouver de la fluorescence en présence de ce polysaccharide, des



expériences de compétitions entre ce lipide et des lipides mycobactériens pourront être envisagées.

Du point de vue des méthodes RMN, nous avons évoqué lors du troisième chapitre qu'il était possible que l'étude en phase solide ne soit pas adaptée aux polysaccharides ramifiés. Des expériences de RMN en phase liquide, de type « Diffusion ordered spectroscopy » (DOSY) par exemple, pourraient être réalisées. Ces expériences visent à séparer le signal RMN de différentes espèces selon leur coefficient de diffusion et a déjà été utilisée pour l'étude de complexes d'inclusion<sup>14</sup>. Elles pourraient ainsi permettre d'approfondir la caractérisation du comportement dynamique de l' $\alpha$ -glucane en présence de lipides dans un environnement aqueux. De plus, la preuve d'interactions entre ce polysaccharide et les lipides de MTB permettrait d'envisager de nouveaux mécanismes de transport de ces derniers de la mycobactérie à l'hôte. Par la suite, ce type d'expérience pourraient aussi être envisagées en présence d'antibiotiques afin d'étudier leurs interactions avec le polysaccharide, permettant ainsi de proposer de nouvelles cibles thérapeutiques telles que les enzymes spécifiques impliquées dans sa biosynthèse.

Les travaux théoriques qui pourraient être entrepris par la suite concernant l'étude de ces interactions peuvent être séparés entre ceux qui concernent le modèle chimique étudié et ceux qui concernent la méthode d'étude utilisée.

Du point de vue du modèle chimique, nombre d'améliorations peuvent y être apportées, que ce soit sur des systèmes contenant de l'amylose ou sur des systèmes contenant des polysaccharides branchés.

Concernant l'amylose, nous avons négligé les interactions entre plusieurs chaînes d'amylose. Ces interactions pourraient être responsables d'une modification de structure et à fortiori du signal RMN obtenu, notamment au niveau des carbones C<sub>2</sub>, C<sub>3</sub> et C<sub>5</sub>, carbones dont nous n'arrivons toujours pas à reproduire le déplacement chimique de manière précise. Dans cette lignée, des calculs utilisant la méthode GIPAW pourraient aussi être envisagés sur des complexes de type polymorphe V à condition d'y inclure les conditions cristallines. Il a été récemment démontré que le déplacement chimique du C<sub>1</sub> pouvait dépendre de l'allomorphe étudié et il serait intéressant de voir si cet effet est reproductible et interprétable d'un point de vue théorique en terme d'environnement local de ce carbone. Enfin, si le signal RMN du

polymorphe B atteste de la présence de deux résidus aux environnements différents, nous pouvons imaginer un effet similaire sur le polymorphe A. En effet le signal du C<sub>1</sub> de ce dernier se présente sous la forme d'un triplet ce qui laisse sous-entendre qu'il existerait 3 environnements différents au sein de cette structure. Une approche employant la méthode GIPAW pourrait permettre de les caractériser.

Concernant les polysaccharides ramifiés, notre étude est préliminaire et incomplète et de ce fait nous pourrions envisager de diversifier notre modèle chimique. Le groupement tréhalose présent sur les MMTs et DMTs semble se comporter comme une ancre dans nos simulations contenant de l'amylose. Un modèle d' $\alpha$ -glucane présentant plus de résidus et des DMT permettrait notamment de pouvoir observer si ces lipides s'agrègent aussi facilement que les molécules d'acide oléique, ou si, au contraire, ils restent piégés dans le polysaccharide en raison des liaisons hydrogènes établies entre le tréhalose et le polysaccharide. De plus, si nos calculs n'ont pas permis d'identifier une structuration locale du polysaccharide en présence de ces lipides, une conformation étendue de ces derniers pourrait peut-être en impliquer une. Les lipides présentés ici ne sont aussi pas les seuls retrouvés dans la capsule de MTB<sup>24</sup>. Le développement de nouveaux modèles de lipides issus de MTB et leur capacité à former des complexes en présence d'amylose ou d' $\alpha$ -glucane forme aussi un champ de recherche très intéressant dans le cadre de la recherche de vaccins, comme nous l'avons évoqué précédemment.

Il existe aussi une question biologique sous-jacente à l'étude de l' $\alpha$ -glucane, à savoir son implication dans le transport de lipides du bacille à l'hôte, ou sa capacité à piéger de petites molécules telles que les antibiotiques utilisés pour combattre la maladie. Si sa capacité à piéger de petites molécules peut être étudié simplement de manière théorique en milieu aqueux, le transport de lipides pourrait impliquer le développement d'un modèle d'interaction entre ce polysaccharide et des membranes. Le polysaccharide pourrait alors ne plus être considéré comme un soluté présent dans le milieu aqueux mais comme une phase facilitant l'extraction de lipides de la membrane. Des calculs de type potential mean force, steered molecular dynamics ou de coefficient de diffusion permettraient ainsi de savoir s'il est plus facile pour un lipide d'être extrait d'une membrane si la phase externe est composée d'une majorité de polysaccharides ou s'il s'agit d'un simple milieu aqueux. Une récente étude de l'institut a démontré qu'il était possible de retrouver du DIM dans les membranes des macrophages infectés par

MTB, de tels calculs pourraient ainsi indiquer un rôle potentiel de notre polysaccharide dans le transfert de lipides de la bactérie à l'hôte.

Outre les modifications du modèle chimique, des modifications de notre méthode d'étude pourraient aussi constituer des perspectives intéressantes.

Concernant les calculs que nous avons réalisés en utilisant la méthode GIPAW, les interactions avec les molécules d'eau ont été complètement négligées. Or il est techniquement envisageable de les inclure dans notre calcul de RMN à condition de pouvoir échantillonner leur dynamique. Pour ceci, une approche de type dynamique moléculaire *ab initio* en conditions périodiques pourrait être utilisée, à condition de disposer de la puissance de calcul nécessaire.

Il est difficile de ne pas mentionner les méthodes de type « machine learning » pour imaginer une suite à ce projet. Lors de ce travail de thèse, nous avons pu générer une quantité non négligeable de calculs RMN sur des polymères composés de résidus glucose. Ces données pourraient permettre d'établir un lien entre l'environnement structural d'un résidu donné (angles glycosidiques, rotamères des groupements hydroxyles, liaisons hydrogènes etc...) et le déplacement chimique calculé de chacun des carbones le composant. Une approche de ce type permettrait de générer des spectres RMN sur des polysaccharides de plusieurs centaines voire milliers de résidus en un temps raisonnable et donc de pouvoir établir d'autres liens entre des structures issues de travaux de dynamique moléculaire et l'expérience.

Finalement ce travail de thèse démontre au travers de l'étude de ces complexes que les méthodes d'études théoriques et expérimentales sont complémentaires et qu'il est important de continuer à connecter ces domaines, même si cela peut parfois paraître comme un véritable défi. Pour ma part, étant plus familier avec les approches théoriques, c'est l'utilisation de ces dernières que je vais approfondir par la suite. Etant principalement intéressé par des objets biologiques, le dialogue avec les sciences expérimentales me sera quoi qu'il en soit nécessaire.

## Références

- (1) Nakamura, Y. *Starch: Metabolism and Structure*; Nakamura, Y., Ed.; Springer Japan: Tokyo, 2015; Vol. 369. <https://doi.org/10.1007/978-4-431-55495-0>.
- (2) Rundle, R. E.; Foster, J. F.; Baldwin, R. R. On the Nature of the Starch-Iodine Complex. *J. Am. Chem. Soc.* **1943**, *66* (5), 2116–2120.
- (3) Rundle, R. E.; Edwards, C. The Configuration of Starch in the Starch-Iodine Complex. IV. An X-Ray Dif- Fraction Investigation of Butanol-Precipitated Amylose. *J. Am. Chem. Soc.* **1943**, *65* (1942), 2200–2203.
- (4) Bluhm, T. L.; Zugenmaier, P. Detailed Structure of the Vh-Amylose-Iodine Complex: A Linear Polyiodine Chain. *Carbohydr. Res.* **1981**, *89* (1), 1–10. [https://doi.org/10.1016/S0008-6215\(00\)85224-6](https://doi.org/10.1016/S0008-6215(00)85224-6).
- (5) Nimz, O.; Gessler, K.; Usón, I.; Sheldrick, G. M.; Saenger, W. Inclusion Complexes of V-Amylose with Undecanoic Acid and Dodecanol at Atomic Resolution: X-Ray Structures with Cycloamylose Containing 26 D-Glucoses (Cyclohexaicosaoose) as Host. *Carbohydr. Res.* **2004**. <https://doi.org/10.1016/j.carres.2004.02.030>.
- (6) Le, C. A. K.; Choïnard, L.; Wouessidjewe, D.; Putaux, J. L. Polymorphism of Crystalline Complexes of V-Amylose with Fatty Acids. *Int. J. Biol. Macromol.* **2018**, *119*, 555–564. <https://doi.org/10.1016/j.ijbiomac.2018.07.163>.
- (7) Le, C. A. K. *Inclusion Complexes of Amylose: Morphogenesis, Crystal Structure and Release of Bioactive Molecules*, 2018.
- (8) Buleon, A.; Duprat, F.; Booy, F. P.; Chanzy, H. Single Crystals of Amylose with a Low Degree of Polymerization. *Carbohydr. Polym.* **1984**, *4* (3), 161–173. [https://doi.org/10.1016/0144-8617\(84\)90009-2](https://doi.org/10.1016/0144-8617(84)90009-2).
- (9) Imberty, A.; Chanzy, H.; Perez, S.; Bulkon, A.; Tran, V. The Double-Helical Nature of the Crystalline Part of A-Sarch. *J. Mol. Biol.* **1988**, *201*, 365–378.
- (10) Imberty, A.; Recherche, C. De; Perez, S.; Physico-chimie, L. De; France, N. A Revisit to the Three-Dimensional Structure of B-Type Starch. *Biopolymers* **1988**, *27*, 1205–1221.

- (11) Rondeau-Mouro, C.; Veronese, G.; Buleon, A. High-Resolution Solid-State NMR of B-Type Amylose. *Biomacromolecules* **2006**. <https://doi.org/10.1021/bm060330x>.
- (12) Beeren, S. R.; Hindsgaul, O. Nature's Dendrimer: Characterizing Amylopectin as a Multivalent Host. *Angew. Chemie - Int. Ed.* **2013**, *52* (43), 11265–11268. <https://doi.org/10.1002/anie.201305132>.
- (13) Beeren, S. R.; Hindsgaul, O. A Fluorescence Assay That Detects Long Branches in the Starch Polysaccharide Amylopectin. *Chem. Commun.* **2014**, *50* (13), 1530–1532. <https://doi.org/10.1039/c3cc48163h>.
- (14) Beeren, S. R.; Meier, S.; Hindsgaul, O. Probing Helical Hydrophobic Binding Sites in Branched Starch Polysaccharides Using NMR Spectroscopy. *Chem. - A Eur. J.* **2013**, *19* (48), 16314–16320. <https://doi.org/10.1002/chem.201302213>.
- (15) Sang, S.; Chen, Y.; Zhu, X.; Narsimhan, G.; Hu, Q.; Jin, Z.; Xu, X. Effect of Egg Yolk Lipids on Structure and Properties of Wheat Starch in Steamed Bread. *J. Cereal Sci.* **2019**, *86* (October 2018), 77–85. <https://doi.org/10.1016/j.jcs.2019.01.009>.
- (16) Lebail, P.; Buleon, A.; Shiftan, D.; Marchessault, R. H. Mobility of Lipid in Complexes of Amylose-Fatty Acids by Deuterium And<sup>13</sup>C Solid State NMR. *Carbohydr. Polym.* **2000**. [https://doi.org/10.1016/S0144-8617\(00\)00180-6](https://doi.org/10.1016/S0144-8617(00)00180-6).
- (17) López, C. A.; De Vries, A. H.; Marrink, S. J. Amylose Folding under the Influence of Lipids. *Carbohydr. Res.* **2012**. <https://doi.org/10.1016/j.carres.2012.10.007>.
- (18) Tusch, M.; Krüger, J.; Fels, G. Structural Stability of V-Amylose Helices in Water-DMSO Mixtures Analyzed by Molecular Dynamics. *J. Chem. Theory Comput.* **2011**, *7* (9), 2919–2928. <https://doi.org/10.1021/ct2005159>.
- (19) Toukach, F. V.; Ananikov, V. P. Recent Advances in Computational Predictions of NMR Parameters for the Structure Elucidation of Carbohydrates: Methods and Limitations. *Chem. Soc. Rev.* **2013**, *42* (21), 8376–8415. <https://doi.org/10.1039/c3cs60073d>.
- (20) Augenstreich, J.; Haanappel, E.; Ferré, G.; Czaplicki, G.; Jolibois, F.; Destainville, N.; Guilhot, C.; Milon, A.; Astarie-Dequeker, C.; Chavent, M. The

- Conical Shape of DIM Lipids Promotes Mycobacterium Tuberculosis Infection of Macrophages. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (51), 25649–25658. <https://doi.org/10.1073/pnas.1910368116>.
- (21) Gao, W.; Liu, Y.; Jing, G.; Li, K.; Zhao, Y.; Sha, B.; Wang, Q.; Wu, D. Rapid and Efficient Crossing Blood-Brain Barrier: Hydrophobic Drug Delivery System Based on Propionylated Amylose Helix Nanoclusters. *Biomaterials* **2017**, *113*, 133–144. <https://doi.org/10.1016/j.biomaterials.2016.10.045>.
- (22) Ishikawa, E.; Ishikawa, T.; Morita, Y. S.; Toyonaga, K.; Yamada, H.; Takeuchi, O.; Kinoshita, T.; Akira, S.; Yoshikai, Y.; Yamasaki, S. Direct Recognition of the Mycobacterial Glycolipid, Trehalose Dimycolate, by C-Type Lectin Mincle. *J. Exp. Med.* **2009**, *206* (13), 2879–2888. <https://doi.org/10.1084/jem.20091750>.
- (23) Kallerup, R. S.; Franzyk, H.; Schiøth, M. L.; Justesen, S.; Martin-Bertelsen, B.; Rose, F.; Madsen, C. M.; Christensen, D.; Korsholm, K. S.; Yaghmur, A.; et al. Adjuvants Based on Synthetic Mycobacterial Cord Factor Analogues: Biophysical Properties of Neat Glycolipids and Nanoself-Assemblies with DDA. *Mol. Pharm.* **2017**, *14* (7), 2294–2306. <https://doi.org/10.1021/acs.molpharmaceut.7b00170>.
- (24) Neyrolles, O.; Guilhot, C. Recent Advances in Deciphering the Contribution of Mycobacterium Tuberculosis Lipids to Pathogenesis. *Tuberculosis* **2011**, *91* (3), 187–195. <https://doi.org/10.1016/j.tube.2011.01.002>.

# Annexes



## Annexe 1. Outils de mécanique quantique

---

Nous devons à Richard Feynman cette célèbre phrase: « I think I can safely say that nobody understands quantum mechanics ». Etant chimiste théoricien de formation, ce chapitre ne vise aucunement à démentir ce fait.

Cependant, la mécanique quantique fait aujourd'hui partie des outils utilisés communément par les chimistes pour étudier des systèmes moléculaires de tailles diverses. Si personne ne peut comprendre la mécanique quantique, il était important pour moi d'essayer de trouver du sens derrière les équations actuellement utilisées. Je me dois de remercier W. Koch et M. Haulthausen pour leur excellent livre *A Chemist's guide to Density Functional Theory*<sup>1</sup>, qui m'a aidé à appréhender de manière plus rationnelle des méthodes de calcul que je considérais opaques, et dont je me suis inspiré pour la rédaction de cette annexe.

S'il me paraissait difficile de rédiger cette annexe sans y introduire les équations discutées, j'espère y avoir apporté assez de sens pour qu'elle reste accessible à des non-physiciens de formation.

### 1. Equation de Schrödinger

Afin d'accéder aux différentes propriétés électroniques d'un système moléculaire de géométrie fixe, on se doit de résoudre l'équation de Schrödinger indépendante du temps :

$$\hat{\mathcal{H}}|\Psi\rangle = E|\Psi\rangle \quad (1.74)$$

$\Psi$  est appelée la fonction d'onde du système et dépend des coordonnées  $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N, \vec{R}_1, \vec{R}_2, \dots, \vec{R}_M\}$  des N électrons et M noyaux du système. Il faut préciser que la grandeur  $\{\vec{x}_i\}$  contient les trois coordonnées spatiales  $\{\vec{r}_i\}$  ainsi que l'information

sur l'état de spin  $\{s_i\}$  de l'électron «  $i$  » considéré. La fonction d'onde contient donc toute l'information du système.

$\hat{\mathcal{H}}$  est l'opérateur hamiltonien du système, défini comme tel :

$$\hat{\mathcal{H}} = \hat{\mathcal{T}}_R + \hat{\mathcal{V}}_R + \hat{\mathcal{T}}_r + \hat{\mathcal{V}}_r + \hat{\mathcal{V}}_{r,R} \quad (1.75)$$

Les deux premiers termes  $\hat{\mathcal{T}}_R$  et  $\hat{\mathcal{V}}_R$  réfèrent respectivement à l'énergie cinétique et potentielle nucléaire, les deux termes suivants  $\hat{\mathcal{T}}_r$  et  $\hat{\mathcal{V}}_r$  à l'énergie cinétique et potentielle électronique et le dernier terme  $\hat{\mathcal{V}}_{r,R}$  est un terme d'énergie potentielle mixte représentant les interactions électron-noyau. Enfin,  $E$  représente l'énergie associée à la fonction d'onde, solution de cette équation.

Une des premières simplifications usuellement faites en vue de la résolution de cette équation est l'approximation de Born-Oppenheimer<sup>2</sup>. Cette approximation profite de la différence de masse entre les électrons et les noyaux pour considérer ces derniers comme étant fixes. Ceci rend le terme cinétique nucléaire nul, et permet de négliger les couplages quantique entre électrons et noyaux, rendant le terme potentiel nucléaire constant. De ce fait, la plupart du temps, seule l'équation de Schrödinger électronique est résolue, dont l'opérateur hamiltonien correspondant est le suivant :

$$\hat{\mathcal{H}}_{elec} = \hat{\mathcal{T}}_r + \hat{\mathcal{V}}_r + \hat{\mathcal{V}}_{r,R} \quad (1.76)$$

Ses termes sont définis comme tel :

$$\hat{\mathcal{T}}_r = - \sum_i^N \frac{\hbar^2}{2m_e} \nabla_i^2 \quad (1.77)$$

$$\hat{V}_r = \frac{1}{4\pi\epsilon_0} \cdot \frac{1}{2} \sum_i^N \frac{q_e^2}{|r_i - r_j|} \quad (1.78)$$

$$\hat{V}_{r,R} = - \frac{1}{4\pi\epsilon_0} \sum_i^N \sum_n^M \frac{Z_n q_e^2}{|r_i - R_n|} \quad (1.79)$$

D'une molécule à une autre, les paramètres qui vont changer seront le nombre d'électrons considérés  $N$  et le potentiel  $\hat{V}_{r,R}$ , car il dépend du nombre de noyaux  $M$  du système. Si ce potentiel décrit l'interaction entre les électrons et les différents noyaux, il arrive aussi de le trouver sous le nom de potentiel externe, ou  $\hat{V}_{ext}$ . Cette appellation permet notamment d'inclure d'autres effets (champs magnétique, électrique etc.) au sein de l'hamiltonien du système étudié. Un opérateur de ce type s'applique de la façon suivante sur les fonctions ( $\Psi$ ) considérées :

$$\langle \hat{\mathcal{H}} \rangle = \int \dots \int \Psi^* \hat{\mathcal{H}} \Psi d\vec{x}_1 d\vec{x}_2 \dots d\vec{x}_N = \langle \Psi | \hat{H} | \Psi \rangle \quad (1.80)$$

En d'autres termes, un opérateur est une fonction dont l'argument est lui-même une fonction et qui associera un nombre à cette dernière. En mathématique, une telle fonction est appelée fonctionnelle. L'application de l'opérateur  $\hat{\mathcal{H}}_{elec}$  sur la fonction d'onde du système permet d'obtenir l'énergie électronique  $E_{elec}$  du système

En pratique, la solution exacte, dont on dit qu'elle correspond à l'état fondamental d'un système, est inatteignable pour des systèmes polyatomiques, principalement car la forme de l'énergie d'interaction électronique est inconnue. Cependant, il existe une méthode permettant de s'approcher de la solution exacte : le principe variationnel. Le principe variationnel stipule que toute solution de l'équation ne correspondant pas à l'état fondamental présentera une énergie supérieure à ce dernier. L'objectif ici est donc de trouver la fonction d'onde  $\Psi_0$  qui donnera l'énergie  $E_0$  la plus basse. Cette

fonction d'onde est appelée l'état fondamental du système. Il faut donc échantillonner toutes les fonctions d'onde possibles et garder celle qui permettra d'obtenir l'énergie la plus basse, afin de se rapprocher le plus possible de  $E_0$ .

## 2. Méthode de résolution de l'équation

On peut recenser deux types de méthodes jusqu'ici développées afin de trouver la meilleure approximation de  $E_0$  en sortie de l'opérateur considéré : celles dont l'argument de l'hamiltonien est une fonction d'onde et celles dont l'argument est la densité électronique. Ces dernières font partie de la Théorie de la Fonctionnelle de la Densité (DFT). Par soucis de clarté et de compréhension, il est préférable de s'attarder en premier lieu sur les méthodes basées sur la fonction d'onde.

### I. Méthodes basées sur la fonction d'onde

Echantillonner toutes les fonctions existantes possibles est en pratique impossible. Pour répondre à cette affirmation, la première méthode mise en place a été développée entre 1927 et 1935 par D. R. Hartree<sup>3,4</sup>. Celle-ci comportant dans un premier temps une infraction au principe d'exclusion de Pauli, elle fut révisée par J. C. Slater<sup>5</sup> et V. A. Fock<sup>6</sup>. Elle est connue aujourd'hui sous le nom d'approximation de Hartree-Fock (HF) et elle a servi de fondation à toutes les méthodes basées sur la fonction d'onde aujourd'hui connues. Cette approximation consiste à considérer qu'une fonction d'onde contenant  $N$  électrons peut être définie comme un produit antisymétrique de  $N$  fonctions contenant un seul électron. La définition mathématique de ce produit a été formulée par Slater en 1929<sup>5</sup> et s'appelle un déterminant de Slater (DS):

$$\Psi_{DS}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\vec{x}_1) & \chi_2(\vec{x}_1) & \dots & \chi_N(\vec{x}_1) \\ \chi_1(\vec{x}_2) & \chi_2(\vec{x}_2) & \dots & \chi_N(\vec{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\vec{x}_N) & \chi_2(\vec{x}_N) & \dots & \chi_N(\vec{x}_N) \end{vmatrix} \quad (2.81)$$

Ici, les éléments du déterminant sont appelés spin-orbitales et sont composés d'une partie spatiale  $\phi(\vec{r})$  et d'une partie déterminant l'état de spin du système  $\sigma(s)$ . Cette dernière peut prendre deux valeurs correspondant aux fonctions de spin  $\alpha$  ou  $\beta$ . Si ces deux fonctions de spin sont naturellement orthogonales, les spin-orbitales sont habituellement définies de façon à l'être aussi.

Maintenant qu'il existe une définition de la forme des fonctions que l'on va considérer, à savoir des DS, il faut trouver celui qui donnera l'énergie la plus basse après application de l'opérateur hamiltonien.

Il a pu être montré que l'énergie calculée à partir de l'application de l'opérateur hamiltonien sur un DS prendra la forme suivante :

$$E_{DS} = \langle \Psi_{DS} | \hat{\mathcal{H}} | \Psi_{DS} \rangle = \sum_i^N (i | \hat{h} | i) + \frac{1}{4\pi\epsilon_0} \cdot \frac{1}{2} \sum_i^N \sum_j^N (ii | jj) - (ij | ji) \quad (2.82)$$

Le terme  $(i | \hat{h} | i)$  désigne la somme de l'énergie cinétique et potentielle, telles que définies dans les équations (1.77) et (1.78), appliquées à la spin-orbitale  $\chi_i$ . Le second et le troisième terme sont appelés intégrale de Coulomb et intégrale d'échange respectivement :

$$(ii | jj) = \iint |\chi_i(\vec{x}_1)|^2 \frac{1}{|\vec{r}_2 - \vec{r}_1|} |\chi_j(\vec{x}_2)|^2 d\vec{x}_1 d\vec{x}_2 \quad (2.83)$$

$$(ij | ji) = \iint |\chi_i(\vec{x}_1)\chi_j(\vec{x}_1)| \frac{1}{|\vec{r}_2 - \vec{r}_1|} |\chi_i(\vec{x}_2)\chi_j(\vec{x}_2)| d\vec{x}_1 d\vec{x}_2 \quad (2.84)$$

Ici il faut pointer un fait qui sera de nouveau discuté lors de la description des méthodes basées sur la densité électronique. La double somme de l'équation (2.9) permet aux indices  $i$  et  $j$  de prendre la même valeur. En soit cela revient à décrire l'interaction d'un électron avec lui-même, ou Self-interaction, ce qui n'a pas de sens physique. Cependant, dans le cas présent, l'intégrale de Coulomb (2.10) et d'échange (2.11) deviennent identiques, et de ce fait annulent la contribution de la double somme.

Les seules entités que nous pouvons faire varier afin de trouver l'énergie la plus basse sont les spin-orbitales. Comme précisé ci-dessus, celles-ci sont par définition orthogonales, et de ce fait, la méthode mathématique utilisée afin de trouver le minimum en énergie est appelée méthode des multiplicateurs de Lagrange. Cette méthode nous amène ici aux équations de Hartree-Fock :

$$\hat{F}\chi_i = \varepsilon_i\chi_i \quad (2.85)$$

$\hat{F}$  est appelé l'opérateur de Fock et comprend les termes suivants :

$$\hat{F}_i = \hat{T}_i + \hat{V}_i + \sum_j^N (\hat{J}_j(\vec{r}_i) + \hat{K}_j(\vec{r}_i)) \quad (2.86)$$

$\hat{T}_i$  et  $\hat{V}_i$  sont les opérateurs décrivant l'énergie cinétique et l'énergie potentielle due à l'interaction entre noyaux et électrons, et les opérateurs  $\hat{J}_j$  et  $\hat{K}_j$  sont les opérateurs coulombien et d'échange menant respectivement aux observables définies par les équations (2.83) et (2.84).

L'opérateur coulombien décrit le potentiel moyen ressenti par l'électron  $i$  du système dû à la distribution de charge moyenne générée par un autre électron présent dans la spin-orbitale  $j$ .

L'opérateur d'échange n'a pas de sens physique a proprement parler et est généré par le caractère antisymétrique du DS. Son action revient à échanger les variables intrinsèques de deux spin-orbitales de même spin.

La somme de ces deux derniers opérateurs telle que définie dans l'équation (13) est appelée le potentiel de Hartree-Fock, ou  $\hat{V}_{HF}$ .

La définition de ce dernier est importante, car elle justifie l'utilisation de la méthode appelée Self Consistent Field (SCF)<sup>7</sup> afin de trouver de manière itérative le meilleur jeu ou vecteur de spin-orbitales. En effet, l'opérateur de Fock est par définition lié aux spin-orbitales de départ à travers  $\hat{V}_{HF}$ . De ce fait, pour trouver le meilleur vecteur, en pratique on utilise un premier jeu de fonctions. Ensuite on résout les équations de Hartree-Fock afin d'améliorer ce vecteur de fonctions. On remplace le set initial par ce dernier, et on réitère l'opération jusqu'à ce que la différence d'énergie entre un vecteur et le suivant soit en dessous d'un seuil prédéfini. On considère alors qu'il y a convergence de l'énergie et la fonction d'onde obtenue est une bonne approximation de la fonction d'onde de l'état fondamental.

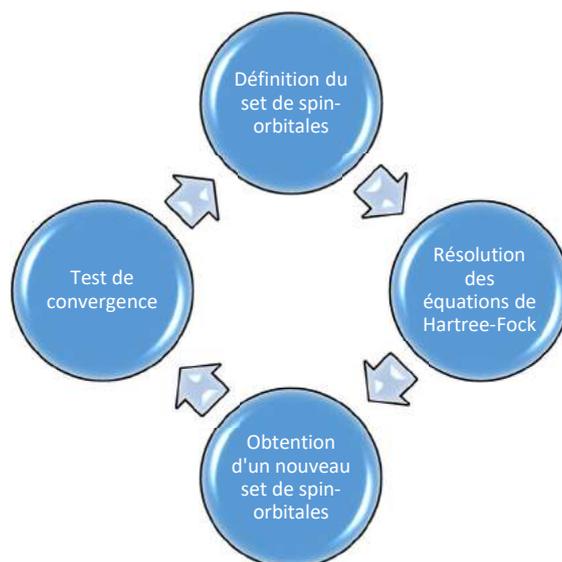


Figure 88: Cycle de la procédure SCF

Il faut préciser que l'approximation réalisée initialement nous offre une solution qui n'est pas la solution exacte de notre système. Physiquement parlant, le déterminant

de Slater représente la fonction d'onde exacte d'un système d'électrons n'interagissant pas les uns avec les autres, soumis à un potentiel extérieur  $\hat{V}_{HF}$ . De ce fait, en se basant sur le principe variationnel défini plus haut, la solution donnée par cette méthode sera toujours supérieure à la solution exacte du système. Ceci nous amène à la définition de l'énergie de corrélation, définie par Löwdin en 1959<sup>8</sup> :

$$E_{Corr} = E_0 - E_{HF} \quad (2.87)$$

De nombreuses méthodes ont par la suite été développées afin de calculer cette énergie. Parmi les plus populaires, on peut citer la méthode développée par Moller et Plesset<sup>9</sup> en 1934, qui se base sur la théorie de la perturbation de Rayleigh-Schrödinger et est nommée selon l'ordre considéré (second ordre : MP2, 4ème ordre : MP4 etc...). D'autres méthodes populaires<sup>10,11</sup> utilisées pour introduire la corrélation électronique sont celles incluant de l'interaction de configuration (CI) ou des approches de type Coupled Cluster (CC). En principe ces approches devraient permettre d'accéder à la solution exacte concernant le système considéré. Elles sont cependant extrêmement coûteuses en ressources computationnelles et restent donc limitées à des systèmes de petite taille. Les méthodes basées sur la fonction d'onde ne seront pas plus décrites, car elles n'ont pas été utilisées au sein de ce travail de thèse.

## II. Méthodes basées sur la densité électronique

Un autre moyen de résoudre l'équation de Schrödinger électronique réside dans l'utilisation de la densité électronique Cette grandeur est définie comme tel :

$$\rho(\vec{r}_1) = N \int \dots \int |\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)|^2 ds_1 d\vec{x}_2 \dots d\vec{x}_N \quad (2.88)$$

Elle décrit la probabilité de trouver un des  $N$  électrons avec un état de spin arbitraire du système dans l'élément de volume présent aux coordonnées  $d\vec{r}_1$  en sachant que les autres électrons ont des positions arbitraires. Etant donné que les  $N$  électrons du système sont tous identiques, l'intégrale multiple est tout simplement multipliée par le nombre d'électrons. On peut noter deux propriétés intéressantes de la densité : i) elle s'annule quand  $\vec{r}$  tend vers l'infini, ii) son intégrale sur l'ensemble de l'espace est égale au nombre d'électrons présents dans le système. A cette notion peut venir s'ajouter la notion de *densité de paire*, qui représente la probabilité de trouver simultanément deux électrons dont les états de spin  $s_1$  et  $s_2$  sont définis dans les éléments de volume  $d\vec{r}_1$  et  $d\vec{r}_2$  :

$$\rho_2(\vec{x}_1, \vec{x}_2) = N(N - 1) \int \cdots \int |\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)|^2 d\vec{x}_3 d\vec{x}_4 \cdots d\vec{x}_N \quad (2.89)$$

Cette quantité présente la particularité de contenir l'information relative à la corrélation électronique sous deux formes : la corrélation de Fermi et la corrélation de Coulomb.

La corrélation de Fermi est due à l'antisymétrie de la fonction d'onde utilisée. Par définition celle-ci amène l'égalité suivante :

$$\rho(\vec{x}_1, \vec{x}_2) = -\rho(\vec{x}_2, \vec{x}_1) \quad (2.90)$$

Si on échange deux variables de la densité de paire, on obtient son opposé. Imaginons que nous voulions échanger les coordonnées de deux électrons de même spin et au sein du même élément de volume. Nous obtiendrions l'égalité suivante :

$$\rho(\vec{x}_1, \vec{x}_1) = -\rho(\vec{x}_1, \vec{x}_1) \quad (2.91)$$

Cette égalité n'est possible que si la densité de paire considérée est nulle. En d'autres termes, il est impossible de trouver deux électrons de même spin exactement au même endroit de l'espace, ce qui est en accord avec le principe d'exclusion de Pauli. Deux électrons ne peuvent donc pas être considérés « sans interaction », de par la corrélation de Fermi.

La corrélation de Coulomb est due à la charge intrinsèque des électrons et se présente sous la forme d'un terme inversement proportionnel à la distance entre les deux électrons considérés.

A ce stade il faut noter que la corrélation de Fermi est incluse dans la méthode HF, au travers du DS qui est lui-même antisymétrique. Cependant la corrélation de Coulomb y est complètement négligée.

Notre densité de paire peut maintenant être réécrite de façon à faire apparaître ces deux types de corrélation :

$$\rho_2(\vec{x}_1, \vec{x}_2) = \rho(\vec{r}_1) \cdot \rho(\vec{r}_2) \cdot [1 + f(\vec{x}_1, \vec{x}_2)] \quad (2.92)$$

La densité de paire est donc la densité électronique de deux électrons non-interagissant pondérée par  $f(\vec{x}_1, \vec{x}_2)$  qu'on appelle le facteur de corrélation. Ici, nous retrouvons le problème de Self-interaction décrit plus haut dans le cas de la méthode HF. En effet, cette définition de la densité de paire intègre pour  $N^2$  électrons au lieu de  $N(N-1)$ . Afin d'y remédier, on la définit en utilisant une probabilité conditionnelle. Il s'agit de la probabilité de trouver un électron de spin  $s_2$  dans un élément de volume  $d\vec{r}_2$ , sachant qu'il y a un électron de spin  $s_1$  présent dans un élément de volume  $d\vec{r}_1$  :

$$P(\vec{x}_1, \vec{x}_2) = \frac{\rho_2(\vec{x}_1, \vec{x}_2)}{\rho(\vec{r}_1)} \quad (2.93)$$

La différence entre cette probabilité et la densité électronique sans interaction est appelée le *trou d'échange-corrélation* :

$$h_{XC} = P(\vec{x}_1, \vec{x}_2) - \rho(\vec{r}_2) = \frac{\rho_2(\vec{x}_1, \vec{x}_2)}{\rho(\vec{r}_1)} - \rho(\vec{r}_2) \quad (2.94)$$

Il peut être montré que cette quantité intègre pour  $-1$  ce qui signifie que le trou « contient » exactement un électron. Cette quantité peut être séparée en deux autres entités appelées le trou de Fermi et le trou de Coulomb, dont les formes mathématiques ont été élégamment représentées dans le cas de la molécule de dihydrogène par Baerends et Gritsenko en 1997<sup>12</sup> :

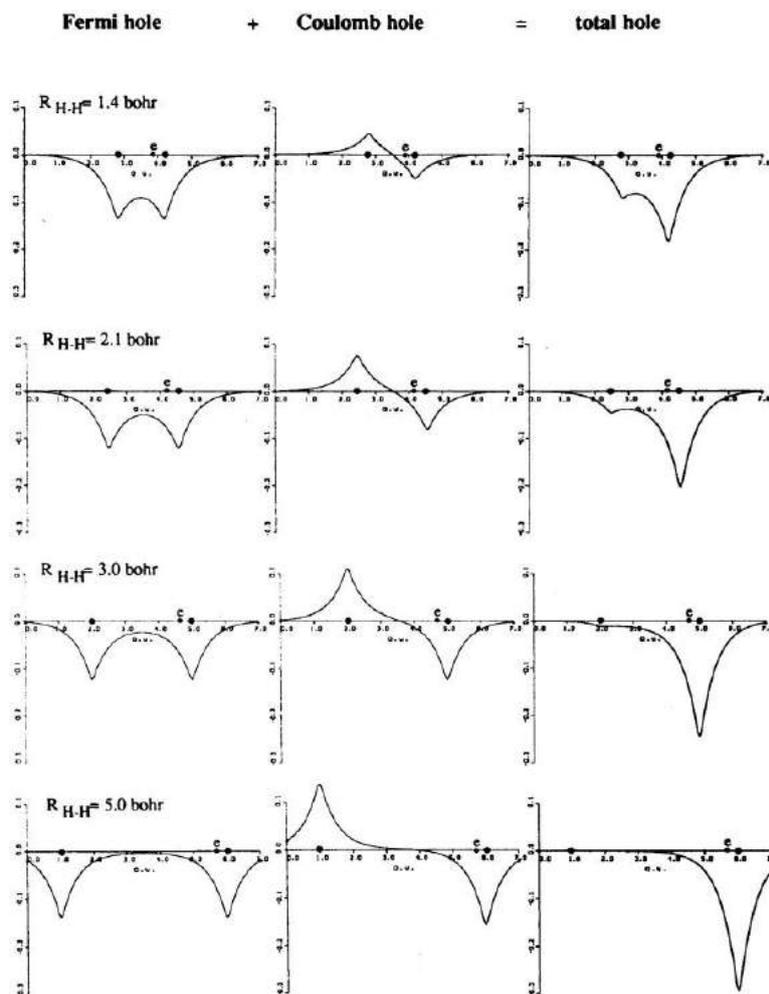


Figure 89: Représentations des trous de Coulomb et de Fermi dans le cas du dihydrogène. Figure extraite de Baerends and Gritsenko, 1997.

Cette grandeur est de première importance, car elle permet de compenser l'effet de self-interaction et de prendre en compte les corrélations de Fermi et de Coulomb dans les calculs.

Les problématiques et notions importantes liées à la densité électronique étant maintenant définies, nous pouvons maintenant nous concentrer sur les méthodes utilisées afin de résoudre l'équation de Schrödinger électronique en utilisant la densité électronique au lieu de la fonction d'onde.

Parmi les premiers essais, on peut citer le modèle de Thomas-Fermi<sup>13,14</sup> développé en 1927. L'énergie cinétique du système était traitée de façon quantique, mais l'énergie potentielle correspondant aux interactions électroniques et mixtes était traitée de manière purement classique. Le calcul de l'énergie cinétique était basé sur un modèle de gaz d'électron uniforme. En 1951<sup>15</sup>, Slater a essayé d'ajouter à ce modèle la contribution du trou d'échange-corrélation en approximant sa forme à celle d'une sphère. Le modèle qui en résulte contient de ce fait le traitement de la self-interaction et une approximation du terme d'échange et de la corrélation de Coulomb.

La véritable révolution concernant la DFT est arrivée en 1964 à travers un article de Hohenberg et Kohn<sup>16</sup>. Dans cet article sont définis les deux théorèmes fondateurs de la DFT telle qu'on la connaît aujourd'hui :

- Théorème n°1 : « On considère un système contenant un nombre arbitraire d'électrons contenus dans une boîte et se déplaçant sous l'action d'un potentiel extérieur  $\hat{V}_{ext}$ . Ce potentiel extérieur, et donc l'opérateur hamiltonien du système, est une fonctionnelle unique de la densité de l'état fondamental ».

Ceci nous permet de définir l'énergie de l'état fondamental comme étant dépendante de la densité de l'état fondamental et de la séparer en trois composantes obtenues à travers les termes définis dans les équations (4) (5) et (6) :

$$E_0[\rho_0] = E_c[\rho_0] + E_{p,elec}[\rho_0] + E_{p,mixte}[\rho_0] \quad (2.95)$$

La seule quantité qui est dépendante du système étudié est l'énergie potentielle mixte. De ce fait, Hohenberg et Kohn définissent une fonctionnelle qui est valide quel que soit le système considéré et permettant d'obtenir de manière exacte la somme de l'énergie cinétique électronique et de l'énergie d'interaction électron-électron :

$$F_{HK}[\rho] = E_c[\rho] + E_{p,elec}[\rho] = \langle \Psi | \hat{T}_r + \hat{V}_r | \Psi \rangle \quad (2.96)$$

Cette fonctionnelle réussit l'exploit de relier l'énergie de la densité utilisée comme argument à celle de la fonction d'onde de l'état correspondant. Cependant, il est en pratique impossible d'accéder à la fonction d'onde d'un état depuis la densité électronique. Finalement,  $F_{HK}$  contient en soit toute l'information de corrélation électronique manquante dans la théorie de HF, et de ce fait résout en théorie de façon exacte l'équation de Schrödinger concernant les termes purement électronique. Il est amusant de pointer du doigt le fait suivant : la démonstration de ce théorème se fait en utilisant un simple raisonnement par l'absurde.

- Théorème n°2 : « Il existe une fonctionnelle universelle  $E[\rho]$  qui définit l'énergie associée à la densité  $\rho$  considérée, valide pour tout potentiel extérieur et nombre d'électrons du système. Pour un potentiel extérieur et un nombre d'électrons fixe, le minimum de cette fonctionnelle est l'énergie de l'état fondamental  $E_0$  et la densité associée  $\rho_0$  est celle de l'état fondamental ».

Si le premier théorème permettait de justifier de l'utilisation de la densité pour trouver l'énergie du système considéré, ce second théorème précise que l'énergie de l'état fondamental, ainsi que toutes ses propriétés, ne seront obtenues qu'en utilisant la densité de l'état fondamental. Il s'agit encore une fois du principe variationnel décrit plus en amont, cette fois appliquée à la DFT.

Ces deux théorèmes ont permis de justifier l'utilisation de la DFT afin de trouver l'état fondamental d'un système. Cependant, aucune indication n'a jusque-là été fournie concernant la forme de  $F_{HK}$ . C'est un an plus tard en 1965 que Kohn et Sham publieront un article<sup>17</sup> dans lequel sera décrit une méthode d'approximation de cette fonctionnelle. L'idée principale de cette méthode est de calculer de manière exacte la

plus grande partie de cette fonctionnelle, et d'approximer les parties inconnues de manière empirique. Pour ce faire, on peut décomposer  $F_{HK}$  de la manière suivante :

$$F_{HK}[\rho] = E_c[\rho] + E_{p,elec}[\rho] = E_{c,ni}[\rho] + J[\rho] + E_{c,nc}[\rho] + E_{p,nc}[\rho] \quad (2.97)$$

Commençons par décrire  $E_{c,ni}[\rho]$  qui représente l'énergie cinétique d'un système non-interagissant. Cette partie de l'énergie totale peut-être obtenue en utilisant un raisonnement similaire à celui qui est décrit dans la méthode HF. Nous avons vu qu'un DS représente la fonction d'onde exacte d'un système de N électrons interagissant avec un potentiel extérieur  $\hat{V}_{HF}$ . Par un raisonnement analogue, nous pouvons construire un Hamiltonien similaire à l'opérateur de Fock décrit équation (2.86) :

$$\hat{F}^{KS} = \hat{T} + \hat{V}^{KS}(\vec{r}) \quad (2.98)$$

Il s'agit d'un Hamiltonien exact pour des particules sans interactions. Le potentiel  $\hat{V}^{KS}(\vec{r})$  est choisi de manière à ce que la densité obtenue à partir des spin-orbitales du DS utilisé soit égale à la densité électronique de l'état fondamental :

$$\rho(\vec{r}) = \sum_i^N |\varphi_i^{KS}(\vec{r})|^2 = \rho_0(\vec{r}) \quad (2.99)$$

Les orbitales considérées sont appelées orbitales de Kohn-Sham et sont donc solution de l'équation suivante :

$$\left( -\frac{\hbar^2}{2m_e} \nabla_i^2 + \hat{V}^{KS}(\vec{r}) \right) \varphi_i^{KS} = \varepsilon_i \varphi_i^{KS} \quad (2.100)$$

A partir de ces orbitales, l'énergie cinétique du système de particules non-interagissant est tout simplement calculée de la manière suivante :

$$E_{c,ni} = -\frac{\hbar^2}{2m_e} \sum_i^N \langle \varphi_i^{KS} | \nabla^2 | \varphi_i^{KS} \rangle \quad (2.101)$$

Nous avons donc un moyen de calculer l'énergie cinétique exacte, nous savons aussi calculer la répulsion coulombienne classique  $J[\rho]$ , il reste donc deux inconnues, la partie d'énergie cinétique négligée  $E_{c,nc}[\rho]$  et la corrélation électronique  $E_{p,nc}[\rho]$ . Kohn et Sham décidèrent de rassembler ces deux derniers termes sous une même forme appelée énergie d'échange-corrélation notée  $E_{XC}$  et valeur propre du potentiel d'échange corrélation  $\hat{V}_{XC}$ .

Nous avons donc une grandeur inconnue qui contient une partie de l'énergie cinétique du système, la self-interaction, la corrélation d'échange et la corrélation de coulomb non classique, qui est indépendante du système considéré. Ce potentiel peut donc être considéré comme un mélange entre une représentation du trou d'échange-corrélation et une correction de l'énergie cinétique du système.

La fonctionnelle de l'énergie d'un système quelconque peut donc être réécrite comme suivant :

$$E_c[\rho] = F_{HK}[\rho] + E_{p,mixte}[\rho] = E_{c,ni}[\rho] + J[\rho] + E_{XC}[\rho] + E_{p,mixte}[\rho] \quad (2.102)$$

A ce jour, toutes les fonctionnelles de la densité qui ont été publiées ont visé à trouver une bonne approximation de  $E_{XC}[\rho]$ . On peut noter deux approches :

- L'approche « Local Density Approximation » (LDA) : Cette approche se base sur le gaz d'électron uniforme déjà mentionné plus haut. Elle définit l'énergie d'échange-corrélation comme suivant :

$$E_{XC}^{LDA}[\rho] = \int \rho(\vec{r}) \varepsilon_{XC}[\rho] d\vec{r} \quad (2.103)$$

$\varepsilon_{XC}[\rho]$  est l'énergie d'échange-corrélation d'une seule particule. On peut aussi trouver cette approche sous le nom de « Local Spin-Density Approximation » (LSDA), auquel cas la densité sera ici séparée en deux densités incluant chacune l'information de spin  $\{\rho_\alpha, \rho_\beta\}$ . Ce type de fonctionnelle ne sera pas décrit ici.

- L'approche « Generalized Gradient Approximation » (GGA) : Celle-ci est similaire à l'approche LDA, mais est définie en fonction du gradient de la densité :

$$E_{XC}^{GGA}[\rho] = \int \rho(\vec{r}) \varepsilon_{XC}[\rho_\alpha, \rho_\beta, \nabla\rho_\alpha, \nabla\rho_\beta] d\vec{r} \quad (2.104)$$

L'inclusion des gradients respectifs de ces densités permet de tenir compte de l'hétérogénéité de la densité.

En pratique,  $E_{XC}^{GGA}[\rho]$  est séparée en deux termes qui sont l'énergie d'échange  $E_X^{GGA}[\rho]$  et l'énergie de corrélation  $E_C^{GGA}[\rho]$ . Ces deux termes n'ont aucun sens physique quand ils sont séparés. Concernant les fonctionnelles permettant de calculer  $E_X^{GGA}[\rho]$  on peut citer, entre autre, la fonctionnelle de Becke (B88)<sup>18</sup>, la fonctionnelle de Perdew et Wang (PW91)<sup>19</sup> ou celle de Perdew, Burke et Ernzerhof (PBE)<sup>20</sup>. Ces deux dernières ont été décrites en même temps que leurs fonctionnelles permettant de calculer  $E_C^{GGA}[\rho]$ , elles aussi appelées PW91 et PBE respectivement. Une dernière fonctionnelle dont il nous faut parler est la fonctionnelle LYP, développée par Lee Yang et Parr en 1988<sup>21</sup>, car c'est celle qui sera utilisé dans cette thèse. En principe, toutes les combinaisons de ces deux types de fonctionnelles sont utilisables.

En 1993, Becke proposa une fonctionnelle hybride appelée B3<sup>22</sup>, basée sur un

mélange entre énergie d'échange exacte d'un système sans interactions (de type HF), énergie définie par utilisation de la méthode LDA et énergie obtenue par les fonctionnelles B88 et PW91 :

$$E_{XC} = E_{XC}^{LSDA} + a_0(E_X^{Exact} - E_X^{LSDA}) + a_X E_X^{B88} + a_C E_C^{PW91} \quad (2.105)$$

Du fait des trois paramètres empiriques  $\{a_0, a_X, a_C\}$  cette fonctionnelle est appelée B3PW91. Il existe une version alternative proposée par Stephens<sup>23</sup> et ses collaborateurs en 1994 :

$$E_{XC} = E_{XC}^{LSDA} + a_0(E_X^{Exact} - E_X^{LSDA}) + a_X E_X^{B88} + a_C E_C^{LYP} + (1 - a_C) E_C^{VWN} \quad (2.106)$$

Cette fonctionnelle s'appelle B3LYP et c'est celle qui sera le plus utilisée dans cette thèse pour les calculs de chimie quantique effectués. Cette fonctionnelle a permis d'obtenir de bons résultats dans les études précédentes et est toujours utilisée aujourd'hui dans le cadre de calculs de tenseur d'écrantage chimiques en RMN. Il existe cependant d'autres fonctionnelles hybrides ayant démontré leur efficacité, telle que la fonctionnelle PBE0<sup>24</sup>, qui est une fonctionnelle hybride dérivée de la fonctionnelle PBE. Nous pouvons aussi mentionner des exemples de fonctionnelles plus sophistiquées, telles que les fonctionnelles dites « méta-GGA », qui dépendent de la dérivée seconde de la densité. Enfin, il existe aussi des fonctions de correction de dispersion qui ont été récemment développées afin d'inclure la corrélation électronique à longue distance dans le calcul d'énergie<sup>25,26</sup>.

Le dernier élément qu'il reste à décrire est ce qu'on appelle la base d'orbitales du système considéré. Pour rappel, dans le cadre de la DFT, nous voulons trouver des orbitales  $\varphi_i^{KS}$  qui sont solutions de l'équation suivante :

$$\hat{F}^{KS} \varphi_i^{KS} = \varepsilon_i \varphi_i^{KS} \quad (2.107)$$

Cette équation est l'équivalent de l'équation (2.12), dans le cadre de la méthode Kohn-Sham. La résolution purement numérique de cette équation étant fastidieuse à réaliser, il est nécessaire de trouver une définition des orbitales  $\varphi_i^{KS}$  pratique à utiliser dans le cadre de méthodes numériques. La solution la plus populaire à ce problème réside dans l'utilisation d'orbitales moléculaires obtenues par la méthode dite « Linear Combination of Atomic Orbitals », développée par Roothaan<sup>27</sup>. Dans cette méthode on définit les orbitales comme étant une combinaison linéaire de fonctions  $\{\eta_\mu\}$ , que l'on appelle fonctions de base :

$$\varphi_i = \sum_{\mu=1}^L c_{\mu i} \eta_\mu \quad (2.108)$$

Ici,  $\mu$  est le numéro de la fonction de base considérée, et  $i$  est le numéro de l'orbitale du système que l'on caractérise. Cette équation ne permet de reproduire les orbitales  $\varphi_i$  que si le nombre total de fonctions utilisées « L » est infini, ce qui n'est en pratique pas possible. De ce fait, afin de reproduire au mieux le comportement de la densité électronique au voisinage du noyau, les fonctions de base devraient être de la forme des orbitales de type Slater ou « Slater-type orbitals » (STO) :

$$\eta^{STO} = N \cdot r^{n-1} \cdot e^{-\zeta r} \cdot Y_{lm}(\theta, \varphi) \quad (2.109)$$

Ici,  $N$  est un facteur de normalisation,  $n$  correspond au nombre quantique principal,  $l$  au nombre quantique secondaire,  $m$  au nombre quantique magnétique et  $\zeta$  est un facteur empirique dépendant de la charge du noyau considéré.  $Y$  est l'harmonique sphérique décrivant la partie angulaire de la fonction, dont  $\theta$  et  $\varphi$  sont les angles de la représentation en coordonnées sphériques. Ces fonctions qui ont été introduites par Slater en 1930 se rapprochent naturellement du comportement de la densité électronique sur les noyaux<sup>28</sup>, présentant notamment une cuspide quand  $r \rightarrow 0$  (Figure 90). Malheureusement, l'utilisation de ce type de fonction rend le calcul des intégrales présentes dans le terme d'échange et dans le terme de corrélation difficile à mettre en

place d'un point de vue informatique. De ce fait, en pratique on utilise plutôt des fonctions de base de type gaussienne ou « Gaussian-type orbitals » (GTO) :

$$\eta^{GTO} = N \cdot x^l \cdot y^m \cdot z^n \cdot e^{-\alpha r^2} \quad (2.110)$$

$N$  étant toujours un facteur de normalisation et  $\alpha$  étant un facteur décrivant la largeur de la gaussienne. Les exposants  $\{l, m, n\}$  servent ici à classer le type de gaussienne utilisée en fonction du type d'orbitale auxquelles elles sont assimilées. De cette façon si on définit  $L = l + m + n$ , on peut construire des orbitales de type s ( $L=0$ ), p ( $L=1$ ) et d ( $L=2$ ).

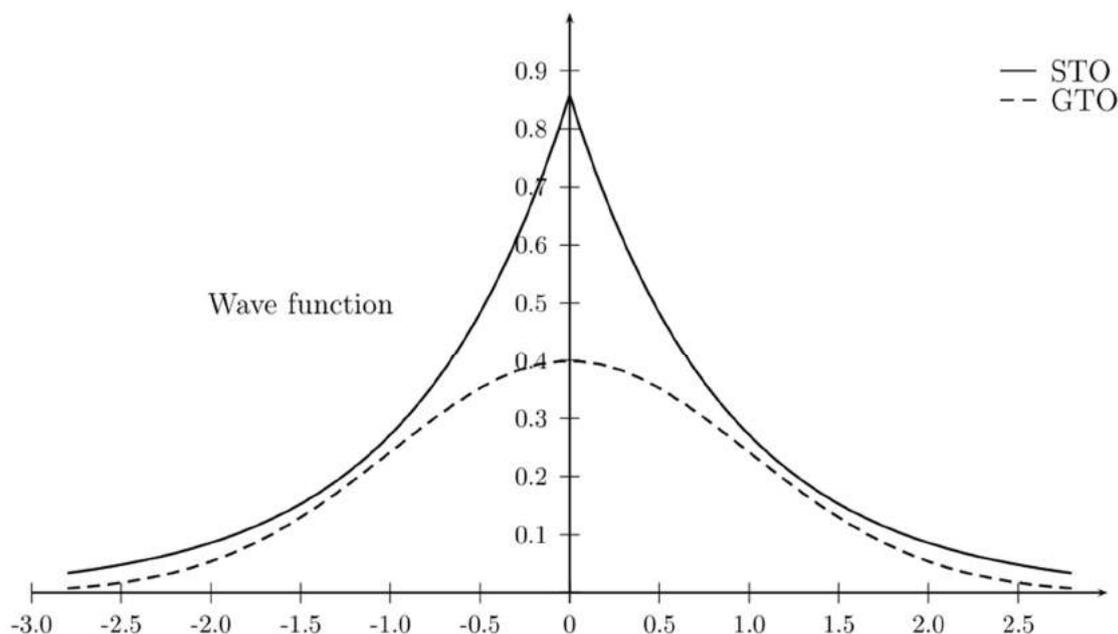


Figure 90: Comparaison de la forme de la partie radiale des STO et GTO aux alentours du noyau considéré.

D'un point de vue informatique, il existe des algorithmes robustes et très efficaces qui permettent de calculer les intégrales d'échange et de corrélation, lorsque l'on considère des GTO en tant que fonctions de bases. Cependant, le comportement de ces fonctions au niveau du noyau est incorrect, et l'on ne retrouve pas la cuspide quand  $r \rightarrow 0$  (Figure 90). Afin de remédier à ce problème, on utilise généralement des fonctions appelées fonction gaussiennes contractées ou « contracted gaussian functions » (CGF) qui sont des sommes pondérées de gaussiennes<sup>29</sup> :

$$\eta^{CGF} = \sum_a^A d_a \eta^{GTO} \quad (2.111)$$

Ces fonctions présentent deux avantages : Elles permettent de s'approcher du comportement des STO (Figure 91), et permettent de réduire aussi le temps total de calcul.

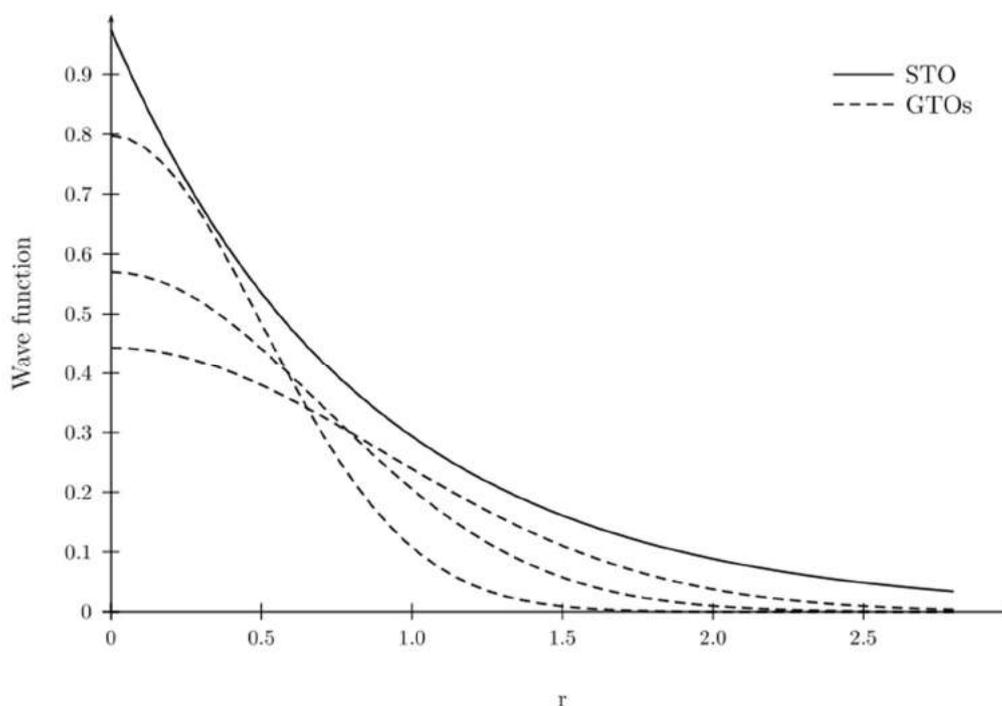


Figure 91: Reconstruction d'une STO en utilisant (de bas en haut) une, deux ou trois GTOs. Extrait de Hoggan et al (2011)

Les jeux de bases couramment utilisés sont généralement nommés selon leur composition en termes de gaussiennes. Dans le cas où on utilise une seule fonction de base pour modéliser toutes les orbitales d'un atome considéré (orbitales de cœur ainsi qu'orbitales de valence), on parle de base minimale. L'exemple le plus connu de base minimale est la base STO-3G, développée par Hehre, Ditchfield, Stewart et John Pople en 1970<sup>30</sup>. Chaque fonction de base utilisée est composée de 3 gaussiennes

(3G) utilisées afin de reproduire une fonction analogue à une STO (STO-3G), il s'agit donc d'une base de type CGF, malgré son nom.

Parmi les jeux de bases un peu plus précis et sophistiqués sont retrouvées les jeux de bases dits « double-zeta ». Ici deux changements apparaissent par rapport aux bases minimales : 1) on différencie les orbitales de cœur et les orbitales de valence et 2) on utilise deux fonctions pour définir chaque type d'orbitales. En pratique, on considère cependant les orbitales de cœur comme étant inertes, et on utilise une seule fonction pour les décrire, on parle alors d'un jeu de base « split-valence ». Un des exemples les plus connus est la base 6-31G, dans laquelle les orbitales de cœur sont décrites par une CGF de 6 gaussiennes, et les orbitales de valence sont décrites par deux fonctions, l'une étant une CGF composée de 3 gaussiennes et l'autre étant une simple GTO<sup>31</sup>.

Enfin, il est d'usage, dans les jeux de fonctions de bases encore un peu plus sophistiqués, d'ajouter des fonctions de polarisation aux fonctions déjà présentes. Les fonctions de polarisation sont d'un moment angulaire supérieur à celles présentes dans les jeux « split-valence ». Par exemple, on rajoute des fonctions de type « d » pour décrire les orbitales « p » des atomes de la seconde ligne du tableau périodique, ou des fonctions de type « p » pour la description des atomes d'hydrogène. L'ajout de ces fonctions permet aux orbitales déjà présentes d'être légèrement modifiées et donc de s'adapter plus facilement à leur environnement proche. Un exemple typique de base « split-valence » présentant des fonctions de polarisation est la base 6-31G (d,p), qui présentent des fonctions de polarisation de type « d » et « p » sur les atomes lourds et les hydrogènes, respectivement.

Il existe bien d'autre type de fonctions bien plus complexes, de type triple-zêta ou quadruple-zêta, dans lesquels on utilise plusieurs fonctions, elles-mêmes augmentées par des fonctions de polarisation pour décrire les différents types d'orbitales du système. La logique de fabrication de ces jeux de fonctions reste la même que celle décrite dans les paragraphes précédents, aussi nous n'en parlerons pas plus car elles ne sont pas utilisées dans ce manuscrit.

En conclusion de cette partie, nous avons vu que nous avons comme objectif de résoudre l'équation de Schrödinger électronique. Pour ce faire il existe deux types d'approches : celles basées sur la fonction d'onde et celles basées sur la densité

électronique. Si les méthodes basées sur la fonction d'onde permettent un formalisme exact, elles sont très coûteuses en ressources calculatoires. Comme nos systèmes d'études sont des systèmes relativement grands (100~1000+ atomes), nous nous sommes tournés vers les méthodes basées sur la densité électronique. Parmi les fonctionnelles utilisées pour décrire l'énergie d'échange-corrélation selon la définition de Kohn et Sham, nous avons décidé d'utiliser la fonctionnelle B3LYP combinée à la base d'orbitale 6-31g(d,p)<sup>32</sup>.

## Références

- (1) Koch, W.; Holthausen, M. C. *Wolfram Koch , Max C . Holthausen A Chemist ' s Guide to Density Functional Theory*; 2001; Vol. 3. <https://doi.org/10.1002/3527600043>.
- (2) Born, M.; Oppenheimer, R. Zur Quantentheorie Der Molekeln. *Ann. Phys.* **1927**.
- (3) Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Math. Proc. Cambridge Philos. Soc.* **1928**, *24* (1), 89–110. <https://doi.org/10.1017/S0305004100011919>.
- (4) Hartree, D. R.; Hartree, W. Self-Consistent Field, with Exchange, for Beryllium. *Proc. R. Soc. London. Ser. A - Math. Phys. Sci.* **1935**, *150* (869), 9–33. <https://doi.org/10.1098/rspa.1935.0085>.
- (5) Slater, J. C. The Theory of Complex Spectra. *Phys. Rev.* **1929**, 1293–1322.
- (6) Fock, V. Näherungsmethode Zur Lösung Des Quantenmechanischen Mehrkörperproblems. *Zeitschrift für Phys.* **1930**, *61*, 126–148.
- (7) Slater, J. C. The Self Consistent Field and the Structure of Atoms. *Phys. Rev.* **1928**, *32*, 339–348.
- (8) Lowdin, P. Correlation Problem in Many-Electron Quantum Mechanics. *Adv. Chem. Phys.* **1958**, *II*, 207–322.
- (9) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.
- (10) Dreuw, A.; Head-Gordon, M. Single-Reference Ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105* (11), 4009–4037. <https://doi.org/10.1021/cr0505627>.
- (11) Bartlett, R. J.; Musiał, M. Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79* (1), 291–352. <https://doi.org/10.1103/RevModPhys.79.291>.
- (12) Baerends, E. J.; Gritsenko, O. V. A Quantum Chemical View of Density Functional Theory. *J. Phys. Chem. A.* **1997**, *101* (30), 5383–5403.

- (13) Thomas, L. H. The Calculation of Atomic Fields. *Math. Proc. Cambridge Philos. Soc.* **1927**, *23* (5), 542–548. <https://doi.org/10.1017/S0305004100011683>.
- (14) Fermi, E. Una Metodo Statistico per La Determinazione Di Alcune Prioprietà Dell'atomo. *Rend. Acad. Dei Lincei* **1927**, *6*, 602–607.
- (15) Slater, J. C. A Simplification of the Hartree-Fock Method. *Phys. Rev.* **1951**, *81* (3), 385–390. <https://doi.org/10.1103/PhysRev.81.385>.
- (16) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), 864–871. <https://doi.org/10.1007/BF01198136>.
- (17) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), 1133–1138.
- (18) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098–3100. <https://doi.org/10.1063/1.1749835>.
- (19) Wang, Y.; Perdew, J. P. Spin Scaling of the Electron-Gas Correlation Energy in the High-Density Limit. *Phys. Rev. B* **1991**, *43* (11), 8911–8916. <https://doi.org/10.1103/PhysRevB.43.8911>.
- (20) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for Mixing Exact Exchange with Density Functional Approximations. *J. Chem. Phys.* **1996**, *105* (22), 9982–9985. <https://doi.org/10.1063/1.472933>.
- (21) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**. <https://doi.org/10.1103/PhysRevB.37.785>.
- (22) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**. <https://doi.org/10.1063/1.464913>.
- (23) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623–11627. <https://doi.org/10.1021/j100096a001>.
- (24) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without

- Adjustable Parameters : The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170.
- (25) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27* (15), 1788–1799. <https://doi.org/10.1002/jcc>.
- (26) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15). <https://doi.org/10.1063/1.3382344>.
- (27) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23* (2), 69–89. <https://doi.org/10.1103/RevModPhys.23.69>.
- (28) Slater, J. C. Atomic Shielding Constants. *Phys. Rev.* **1930**, *36* (1), 57–64. <https://doi.org/10.1103/PhysRev.36.57>.
- (29) Hoggan, P. E.; Ruiz, M. B.; Özdogan, T. MOLECULAR INTEGRALS OVER SLATER-TYPE ORBITALS. FROM PIONEERS TO RECENT DEVELOPMENTS. In *Quantum Frontiers of Atoms and Molecules*; 2011; pp 64–90. <https://doi.org/10.4037/ajcc2009686>.
- (30) Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. Self-Consistent Molecular Orbital Methods. IV. Use of Gaussian Expansions of Slater-type Orbitals. Extension to Second-Row Molecules. *J. Chem. Phys.* **1970**, *52* (5), 2763–2768. <https://doi.org/10.1063/1.1673374>.
- (31) Hehre, W. J.; Ditchfield, K.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56* (5), 2257–2261. <https://doi.org/10.1063/1.1677527>.
- (32) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *J. Chem. Phys.* **1980**, *72* (1), 650–654. <https://doi.org/10.1063/1.438955>.

## Annexe 2. Article n°1 : Informations supplémentaires

---

Voici les informations supplémentaire concernant l'article intitulé « Structures and NMR spectra of short amylose-lipid complexes. Insight using molecular dynamics and DFT quantum chemical calculations » présentées sous leur version finale pour publication dans *Carbohydrate Polymers* avant édition par le journal.

## Supplementary Materials

### **Structures and NMR spectra of short amylose-lipid complexes. Insight using molecular dynamics and DFT quantum chemical calculations.**

Adrien Schahl<sup>1,2</sup>, Valérie Réat<sup>2</sup>, Franck Jolibois<sup>1\*</sup>

1) LPCNO, Université Fédérale de Toulouse Midi-Pyrénées, INSA-CNRS-UPS, 135 avenue de Rangueil, 31077 Cedex 4 Toulouse, France

2) Institut de Pharmacologie et de Biologie Structurale, UMR 5089, CNRS - Université de Toulouse, UPS BP 64182, 205 route de Narbonne, 31077 Toulouse, Cedex04, France

\* Corresponding author: mail to [franck.jolibois@univ-tlse3.fr](mailto:franck.jolibois@univ-tlse3.fr)

Stability of the helical forms: The left-handed V helix is a form of amylose that has been intensively characterised (Eliasson & Krog, 1985; Godet et al., 1993; Sarko & Winter, 1974). While a right-handed form can be envisaged regarding structural considerations only, it has rarely been observed experimentally, to the best of our knowledge (Bulpin, Welsh, Morris, & England, 1982). Moreover, the so-called V<sub>6</sub> helix (with 6 glucose residues per turn) seems to be more commonly detected experimentally than other kind of helices with more glucose residues per turn. In order to explain these observations, the relative stabilities of 19-residue right- and left-handed V<sub>6</sub> helices have been computed at the quantum chemical level (see computational details), showing that the left-handed form is more stable than the right-handed form by 60 kcal.mol<sup>-1</sup>. This energy difference can be attributed to a difference in the hydrogen bonding network created upon helix formation (Figure S1). In the left-handed V<sub>6</sub> helix, hydrogen bonds are observed between the (HO)<sub>6</sub> hydroxyl group at residue [n] with the (HO)<sub>2</sub> (H-bond acceptor) and (HO)<sub>3</sub> (H-bond donor) hydroxyl groups at residue [n+6], where (HO)<sub>2</sub>, (HO)<sub>3</sub> and (HO)<sub>6</sub> correspond to the hydroxyl groups of carbons 2, 3 and 6, respectively (Figure 1). A hydrogen bond is also formed between (HO)<sub>3</sub> (H-bond acceptor) at residue [n] and (HO)<sub>2</sub> (H-bond donor) at residue [n+1]. These features are in perfect agreement with the experimental structure obtained by X-Ray diffraction (Gessler et al., 1999). In the case of the right-handed V<sub>6</sub> helix, hydrogen bonds are observed between the (HO)<sub>6</sub> hydroxyl group at residue [n] with (HO)<sub>2</sub> (H-bond acceptor) and HO<sub>3</sub> (H-bond donor) at residues [n+5] and [n+6], respectively. Thus, the third hydrogen bond observed in the left-handed structure cannot be formed between successive residues in the right-handed structure.

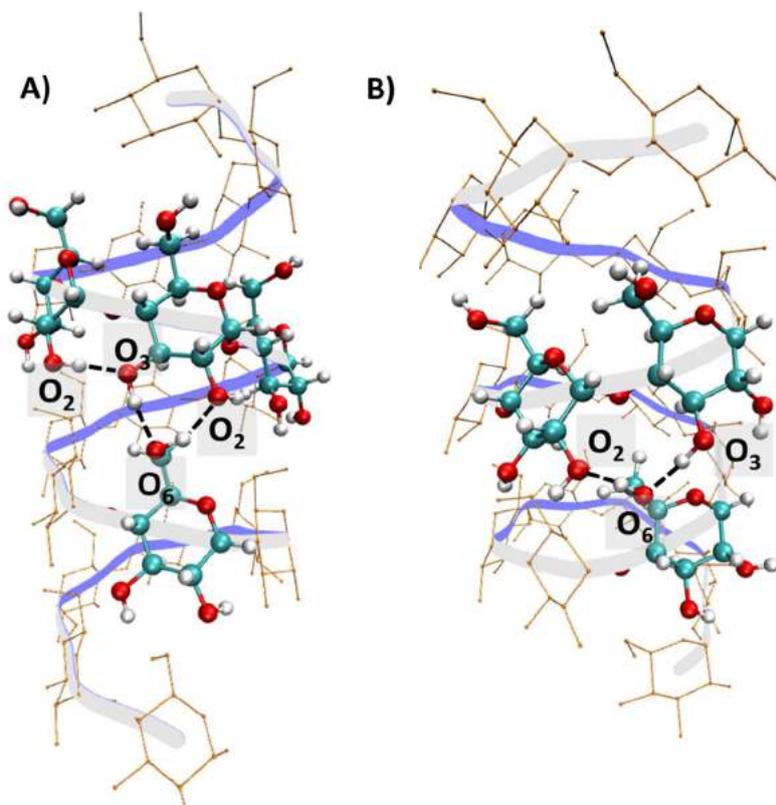


Figure S1: Hydrogen bond network upon helix formation. A) Left-handed  $V_6$  helix; B) Right-handed  $V_6$  helix.

Comparison with the left-handed  $V_7$  helix shows an energy difference of  $40 \text{ kcal.mol}^{-1}$ . Solvent effects have been accounted for using a SMD continuum solvation model (Marenich, Cramer, & Truhlar, 2009) for geometry optimisations. In this case, the  $V_6$  helix remains more stable than the  $V_7$  helix, with the relative energy difference reduced to  $26 \text{ kcal.mol}^{-1}$ . These theoretical results corroborate the experimental observations.

#### References :

Bulpin, B. P. V, Welsh, E. J., Morris, E. R., & England, B. (1982). Physical Characterization of Amylose-Fatty Acid Complexes in Starch Granules and in Solution. *Starch-Stärke*, 10, 335–339.

Eliasson, A. C., & Krog, N. (1985). Physical properties of amylose-monoglyceride complexes. *Journal of Cereal Science*, 3(3), 239–248.

Gessler, K., Uson, I., Takaha, T., Krauss, N., Smith, S. M., Okada, S., ... Saenger, W. (1999). V-Amylose at atomic resolution: X-ray structure of a cycloamylose with 26 glucose residues (cyclomaltohexaicosaoase). *Proceedings of the National Academy of Sciences*, 96(8), 4246–4251.

Godet, M. C., Buléon, A., Tran, V., & Colonna, P. (1993). Structural features of fatty acid-amylose complexes. *Carbohydrate Polymers*, 21(2–3), 91–95

Marenich, A. V, Cramer, C. J., & Truhlar, D. G. (2009). Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B*, 113(18), 6378–6396.

Sarko, A., & Winter, W. T. (1974). Crystal and Molecular Structure of the Amylose-DMSO Complex. *Biopolymers*, 13(7), 1461–1482.1016/0144-8617(93)90003-M

Figure S2:  $[n]-[n+6]$  distance variations along the amylose chain in the case of the analysis of the stability of a preformed  $V_6$  helix of 19 residues complexed with a palmitic acid molecule. The variations are plotted clockwise as a function of time for the whole NVT molecular dynamics simulation and for each residue along the amylose chain (from  $n=1$  to  $n=13$ ). The yellow circle corresponds to the experimental value for the V-helix conformation (see main text for details)

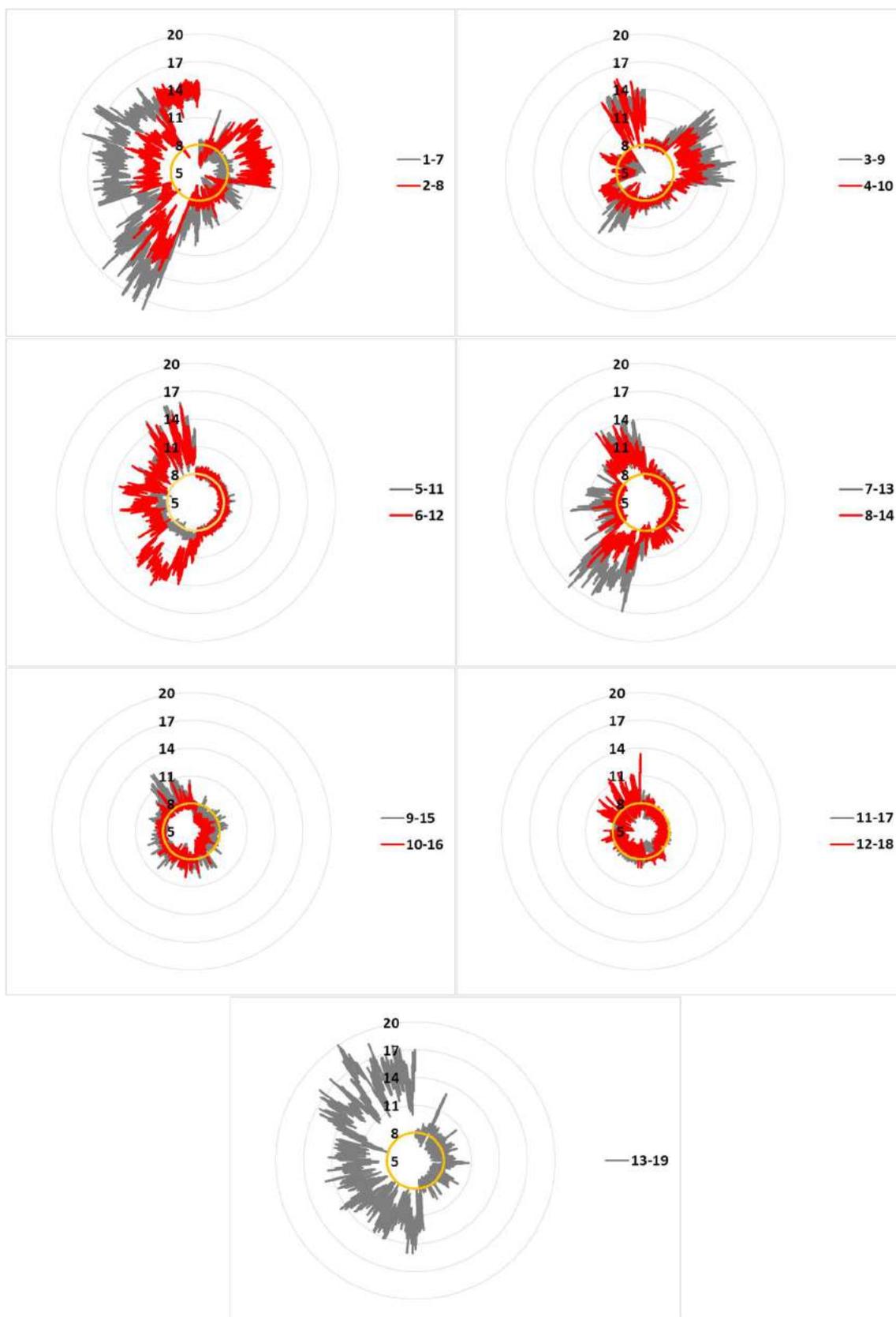
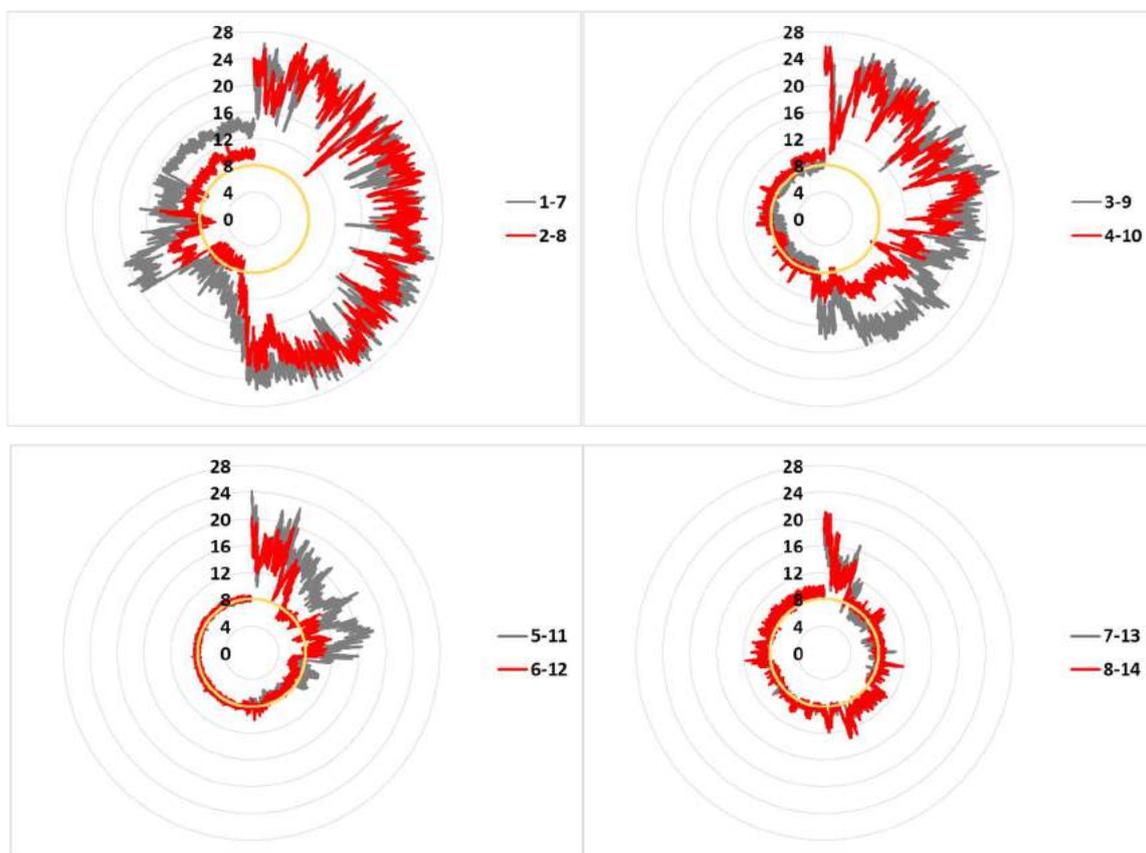


Figure S3:  $[n]-[n+6]$  distance variations along the amylose chain in the case of the analysis of the complexation process between a 19 residues amylose and a palmitic acid molecule. The variations are plotted clockwise as a function of time for the whole NVT molecular dynamics simulation and for each residue along the amylose chain (from  $n=1$  to  $n=13$ ). The yellow circle corresponds to the experimental value for the V-helix conformation (see main text for details)



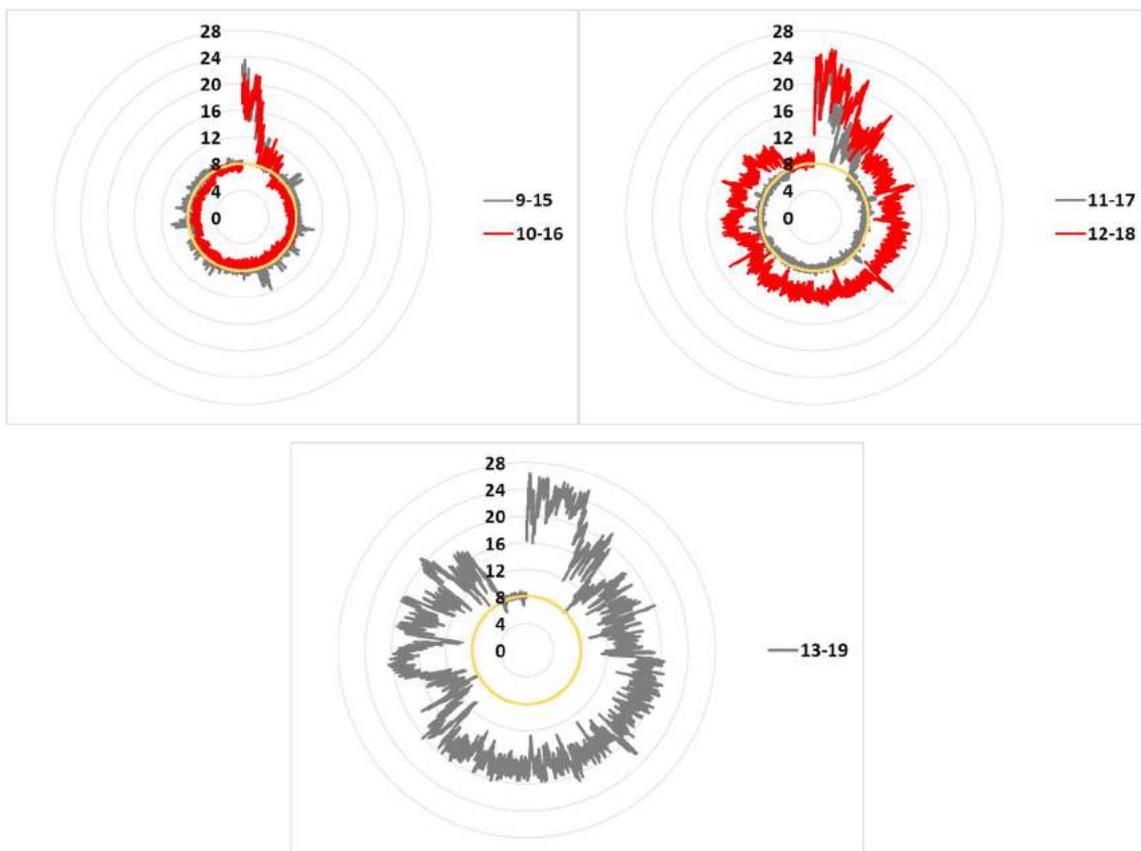


Figure S4: Corrected theoretical  $^{13}\text{C}$  NMR spectra of amylose-palmitic acid complexes with a 19-residue amylose from 50 snapshots extracted from the last 10 ns of complexation analysis molecular dynamics simulation (Red Line), the first (Black Plain Line) and the last (Black Dotted line) 10 ns of stability analysis molecular dynamics simulation.

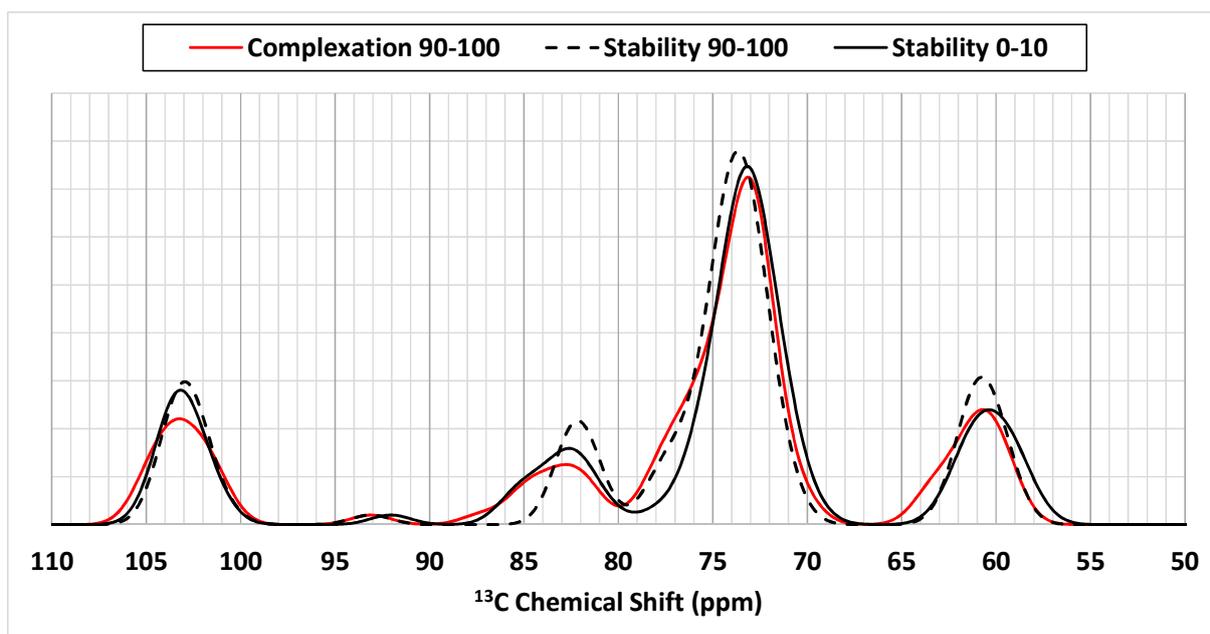


Figure S5: Deconvolution of theoretical  $^{13}\text{C}$  NMR spectra of amylose-palmitic acid complexes with 19-residue amylose obtained from 50 snapshots extracted from the first 10 ns of stability analysis molecular dynamics simulation. Black line: Spectrum calculated using the full set of 19 residues; Contributions to the spectrum for residues 1 to 2, 3 to 10 and 11 to 19 are shown using blue, red and dotted red lines, respectively.

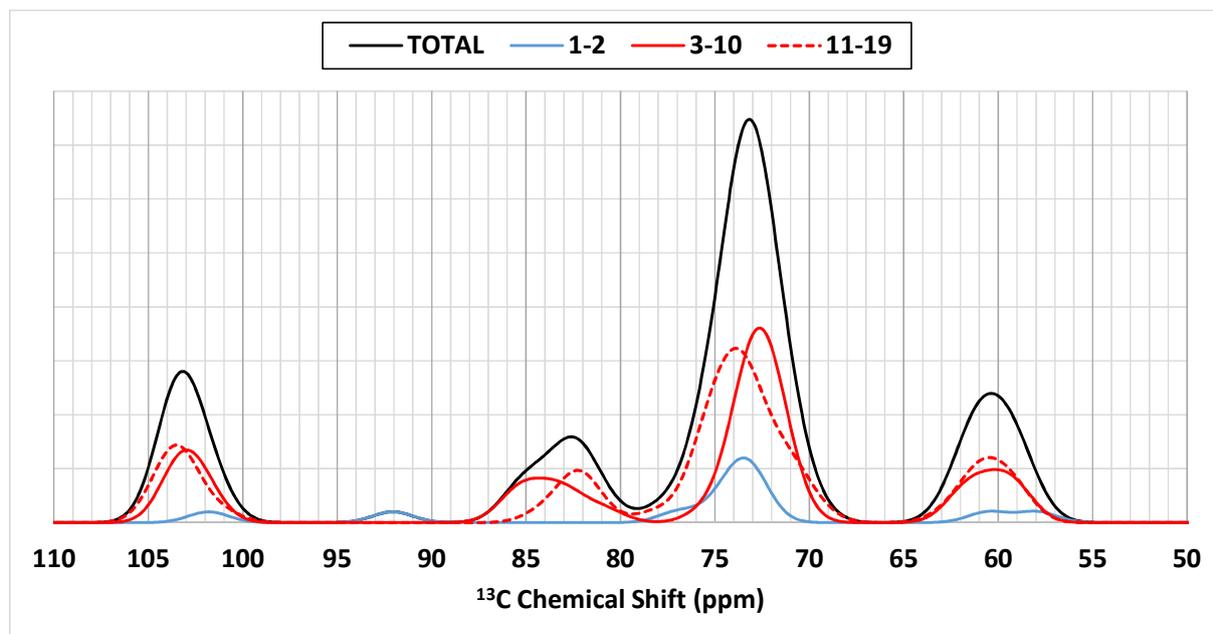
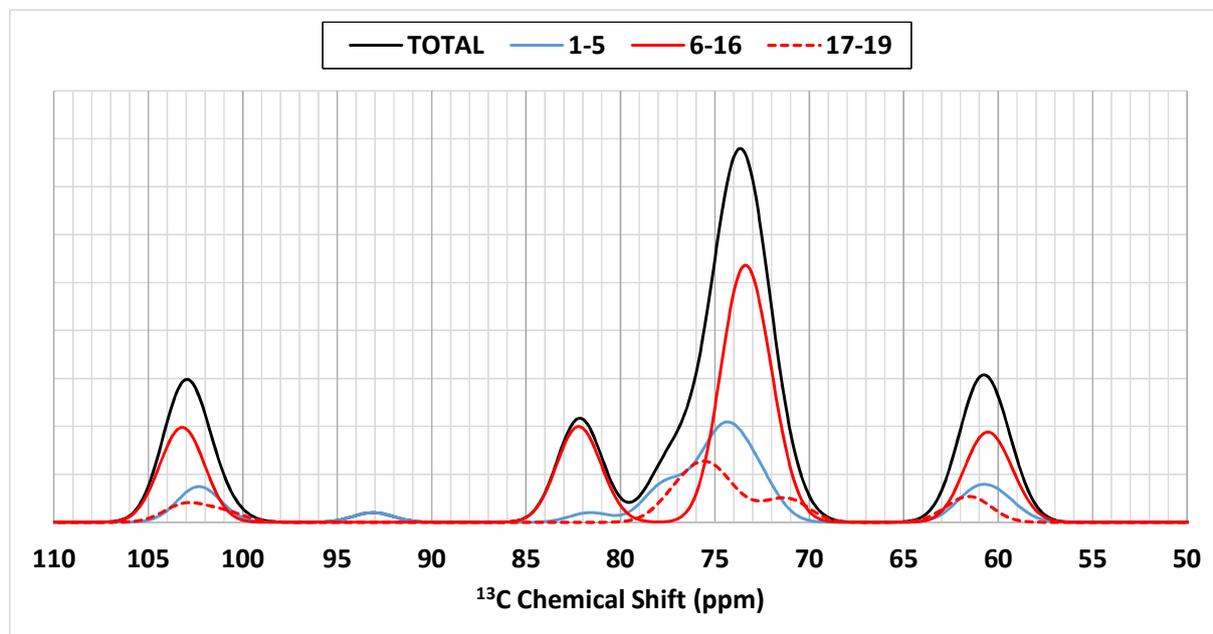


Figure S6: Deconvolution of theoretical  $^{13}\text{C}$  NMR spectra of amylose-palmitic acid complexes with 19-residue amylose obtained from 50 snapshots extracted from the last 10 ns of stability analysis molecular dynamics simulation. Black line: Spectrum calculated using the full set of 19 residues; Contributions to the spectrum for residues 1 to 5, 6 to 16 and 17 to 19 are shown using blue, red and dotted red lines, respectively.



## Annexe 3. Article n°2 : Informations supplémentaires

---

Voici les informations supplémentaire concernant l'article intitulé « Diversity of hydrogen bond network and its impact on NMR parameters of amylose B polymorph. A study using molecular dynamics and DFT calculations within periodic boundary conditions » présentées sous leur version finale pour soumission dans *Journal of Physical Chemistry A*.

## **ELECTRONIC SUPPLEMENTARY INFORMATION**

### **Diversity of hydrogen bond network and its impact on NMR parameters of amylose B polymorph. A study using molecular dynamics and DFT calculations within periodic boundary conditions.**

Adrien Schahl,<sup>a,b</sup> Iann Gerber,<sup>a</sup> Valérie Réat<sup>b</sup> and Franck Jolibois<sup>\*a</sup>

1) LPCNO, Université Fédérale de Toulouse Midi-pyrénées, INSA-CNRS-UPS, 135 avenue de Rangeuil, 31077 Cedex 4 Toulouse, France

2) Institut de Pharmacologie et Biologie Structurale, UMR 5089, CNRS – Université de Toulouse, UPS BP 64182, 205 route de Narbonne, 31077 Toulouse, Cedex 04, France

\*Corresponding Author: Email: [franck.jolibois@univ-tlse3.fr](mailto:franck.jolibois@univ-tlse3.fr)

Table S1:  $\varphi$  and  $\psi$  angles after geometry optimisations and for the crystal structure.  $N_{link}$  : number of the link between two glucose residues

$N_{link}$	Opt <sub>HBN_1</sub>		Opt <sub>HBN_2</sub>		Crystal	
	$\varphi$	$\psi$	$\varphi$	$\psi$	$\varphi$	$\psi$
1	97,6	106,5	97,6	94,8	83,9	94,3
2	91,9	102,4	92,3	99,3	84,2	94,0
3	90,8	99,7	84,6	96,7	83,9	94,3
4	88,2	103,4	88,0	97,7	84,0	94,4
5	95,7	95,5	100,9	94,3	83,9	94,3
1'	104,1	90,2	93,8	90,7	83,9	94,3
2'	91,9	97,3	92,0	94,7	84,2	94,0
3'	98,8	86,0	99,6	86,1	83,9	94,3
4'	93,7	97,3	100,2	93,2	84,3	94,0
5'	94,5	99,3	103,7	99,1	83,9	94,3

Table S2: Selected dihedral angles for full (all atoms) geometry optimisations. Prime numbers refer to the second amylose chain in the double helix.

	Residue number	C <sub>1</sub> -C <sub>2</sub> -O <sub>2</sub> -H <sub>O2</sub>	C <sub>2</sub> -C <sub>3</sub> -O <sub>3</sub> -H <sub>O3</sub>	C <sub>4</sub> -C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub>	C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub> -H <sub>O6</sub>
<b>Opt<sub>HBN_1</sub></b>	1	195,0	197,9	47,0	287,9
	2	192,4	198,4	47,8	289,6
	3	275,8	184,1	55,9	285,5
	4	236,6	181,9	38,5	285,1
	5	243,2	189,9	40,2	281,0
	6	245,5	174,2	62,2	287,5
	1'	195,7	197,9	47,0	287,2
	2'	191,5	198,4	47,8	285,8
	3'	195,8	184,1	55,9	277,5
	4'	211,6	193,5	38,5	287,5
	5'	236,6	191,6	40,2	286,6
	6'	262,7	172,5	62,2	288,2
<b>Opt<sub>HBN_2</sub></b>	1	87,4	309,0	54,0	169,3
	2	5,4	306,9	52,1	160,8
	3	358,6	307,0	55,7	144,9
	4	34,5	304,1	56,7	155,5
	5	27,1	309,1	59,0	143,0
	6	25,1	302,2	59,8	55,4
	1'	314,4	313,9	55,5	151,1
	2'	28,4	307,7	56,2	158,6
	3'	27,5	310,8	61,2	151,1
	4'	29,2	308,6	60,2	55,9
	5'	24,3	311,0	61,1	55,3
	6'	21,9	301,5	60,4	54,9

Table S3: Selected dihedral angles for constrained (Hydrogen atoms only) geometry optimisations. Prime numbers refer to the second amylose chain in the double helix.

	Residue number	C <sub>1</sub> -C <sub>2</sub> -O <sub>2</sub> -H <sub>O2</sub>	C <sub>2</sub> -C <sub>3</sub> -O <sub>3</sub> -H <sub>O3</sub>	C <sub>4</sub> -C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub>	C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub> -H <sub>O6</sub>
<b>Opt<sub>HBN_1</sub>(H)</b>	1	192,0	212,7	67,5	284,4
	2	182,0	212,9	51,7	286,4
	3	299,8	206,1	67,6	282,7
	4	279,8	205,9	51,8	286,6
	5	276,6	205,1	67,5	280,7
	6	279,8	175,6	51,7	181,1
	1'	192,2	209,0	67,5	284,4
	2'	279,7	206,2	51,7	287,1
	3'	277,0	205,8	67,6	282,2
	4'	280,1	205,3	51,7	288,4
	5'	276,7	204,9	67,5	283,5
	6'	280,3	178,2	51,8	260,0
<b>Opt<sub>HBN_2</sub>(H)</b>	1	312,1	308,1	67,5	144,1
	2	313,2	307,9	51,7	134,7
	3	314,4	307,7	67,6	142,8
	4	46,2	299,1	51,8	134,5
	5	36,6	300,4	67,5	143,6
	6	45,8	286,2	51,7	59,7
	1'	312,7	307,7	67,5	144,2
	2'	47,3	299,3	51,7	134,1
	3'	36,7	299,7	67,6	143,8
	4'	49,0	299,2	51,7	61,2
	5'	37,3	299,7	67,5	50,3
	6'	45,6	285,5	51,7	60,1

Table S4:  $\varphi$  and  $\psi$  angles averages extracted from MD simulation. Prime numbers refer to the second amylose chain in the double helix.  $N_{link}$  number of the link between two glucose residues. Sta.dev. = Standard deviation.

$N_{link}$	$\varphi$	Sta.dev.	$\psi$	Sta.dev.
1	101,0	23,6	104,0	12,5
2	91,4	14,5	93,6	10,0
3	98,1	13,4	94,1	7,7
4	93,7	15,1	92,8	7,1
5	86,3	12,8	91,4	7,5
6	87,3	11,2	92,5	6,8
7	83,1	7,8	92,9	6,6
8	83,8	8,4	91,8	6,9
9	83,7	8,1	92,8	6,8
10	83,7	8,5	93,6	7,2
11	81,0	8,5	92,0	6,5
12	83,6	8,5	93,9	7,3
13	83,7	11,0	91,8	6,8
14	80,8	7,4	92,4	7,2
15	90,3	11,4	92,2	7,2
16	78,3	6,4	94,9	6,7
17	110,6	9,1	95,4	8,0
18	86,0	11,5	90,2	8,0
19	80,7	24,7	87,2	15,5
1'	104,4	13,7	96,9	10,3
2'	89,4	15,2	90,3	9,1
3'	76,7	8,7	89,9	6,9
4'	81,0	11,4	92,3	6,6
5'	82,8	7,7	93,8	6,5
6'	82,8	7,7	92,9	6,7
7'	84,0	9,2	92,2	6,5
8'	82,7	8,1	91,2	6,8
9'	85,7	10,4	92,4	6,6
10'	85,5	10,2	93,0	7,2
11'	86,9	13,5	93,0	6,8
12'	83,4	9,2	94,4	7,5
13'	78,4	8,4	90,9	6,9
14'	96,7	12,0	102,3	8,4
15'	74,1	6,6	92,3	6,1
16'	87,4	10,4	106,1	9,1
17'	79,5	7,8	91,1	8,4
18'	108,6	11,1	96,5	9,8
19'	82,0	13,4	86,6	10,4

Table S5 Selected dihedral angles averages extracted from MD simulation. Prime numbers refer to the second amylose chain in the double helix.  $N_{link}$  number of the link between two glucose residues. Sta.dev. = Standard deviation.

$N_{link}$	C <sub>1</sub> -C <sub>2</sub> -O <sub>2</sub> -H <sub>O2</sub>	Sta.dev	C <sub>2</sub> -C <sub>3</sub> -O <sub>3</sub> -H <sub>O3</sub>	Sta.dev	C <sub>4</sub> -C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub>	Sta.dev	C <sub>5</sub> -C <sub>6</sub> -O <sub>6</sub> -H <sub>O6</sub>	Sta.dev
1	181,8	70,4	215,2	67,3	85,2	63,3	189,4	65,8
2	198,1	70,8	221,2	62,8	132,6	63,2	180,8	49,2
3	175,7	55,7	214,4	53,1	130,5	62,8	166,7	33,4
4	168,6	58,2	215,5	62,0	69,5	39,1	148,9	28,8
5	163,5	64,2	233,7	71,9	59,9	23,7	147,7	19,0
6	132,7	55,2	243,1	67,4	63,4	31,4	147,3	28,0
7	117,6	51,1	252,5	72,8	57,2	20,8	157,3	37,2
8	99,7	30,8	251,3	78,7	74,8	48,0	153,1	28,5
9	102,1	31,4	253,0	75,7	57,6	18,5	145,3	45,5
10	104,2	37,7	252,5	73,7	65,9	38,0	140,9	35,7
11	101,0	30,7	248,6	79,9	69,9	41,2	112,5	48,5
12	112,9	52,2	252,9	72,9	94,6	55,4	127,4	70,0
13	110,4	49,7	251,3	71,5	60,4	23,6	134,6	43,0
14	110,4	38,2	244,3	79,8	54,4	11,9	186,4	47,1
15	160,1	94,2	246,6	68,2	118,2	65,0	177,0	38,4
16	126,4	61,5	243,8	79,0	64,6	30,3	223,0	36,1
17	132,1	53,3	202,9	31,9	170,4	40,4	201,1	32,2
18	162,0	34,0	179,7	96,1	61,4	38,3	206,0	24,7
19	291,1	26,4	242,1	72,0	94,8	60,9	166,2	59,9
20	169,1	83,0	240,2	63,8	143,6	65,8	176,9	87,2
1'	168,6	58,7	195,4	54,7	123,3	62,0	161,1	73,7
2'	172,0	45,1	216,1	67,3	98,1	56,8	162,4	82,6
3'	144,7	36,3	237,1	74,6	62,7	38,4	117,8	40,4
4'	138,5	39,9	255,4	51,2	69,3	41,2	121,9	42,5
5'	113,9	37,2	261,9	56,2	56,3	14,7	140,4	18,5
6'	99,0	29,8	240,9	88,7	55,5	13,3	140,0	19,2
7'	105,2	40,7	246,3	77,6	57,6	19,3	143,2	24,8
8'	116,8	52,0	249,4	78,1	55,5	11,8	140,3	20,6
9'	109,7	35,0	245,8	76,5	62,8	30,5	141,7	34,0
10'	119,5	57,6	244,6	76,4	60,5	25,4	144,2	34,4
11'	113,4	39,0	238,8	72,7	58,7	27,2	134,4	35,6
12'	142,0	79,3	241,8	81,8	83,4	52,0	118,2	59,4
13'	164,5	93,6	256,3	70,3	64,0	33,2	140,9	50,7
14'	113,8	52,8	226,3	63,0	98,3	63,4	158,6	53,4
15'	131,2	50,9	248,1	68,3	54,0	8,0	69,1	13,0
16'	134,2	46,5	228,9	72,5	169,5	15,2	50,0	48,5
17'	180,1	76,6	245,3	66,9	57,3	18,0	124,5	75,3
18'	146,5	43,9	204,8	39,4	179,3	47,1	185,4	108,1
19'	170,7	41,2	226,0	69,4	93,1	60,5	171,9	81,0
20'	265,8	67,0	242,2	69,3	130,9	69,1	164,6	90,1

Table S6: Chemical shielding obtained for full (all atoms) geometry optimisations methods. Prime number refers to the numbering of the second chain of amylose in the double helix.

	Residue number	C1	C2	C3	C4	C5	C6
<b>Opt<sub>HBN_1</sub></b>	1	98.5	117.8	116.8	109.2	119.0	127.7
	2	88.9	118.3	117.4	112.0	119.8	127.1
	3	88.8	119.0	116.7	120.6	116.1	129.5
	4	88.0	119.2	119.2	113.7	117.7	128.6
	5	88.3	119.5	118.6	112.8	118.1	128.1
	6	87.7	119.3	119.9	114.1	118.4	128.5
	1'	98.3	118.5	116.1	113.1	117.8	127.3
	2'	91.0	118.7	117.2	113.6	118.0	126.3
	3'	89.8	119.5	117.6	116.1	118.7	125.9
	4'	91.0	119.2	118.7	113.0	120.5	127.4
	5'	88.6	119.5	118.5	120.8	117.3	129.7
	6'	88.7	119.1	118.7	112.2	118.6	128.7
<b>Opt<sub>HBN_2</sub></b>	1	93.7	116.8	116.0	115.5	118.7	129.7
	2	89.9	117.5	115.0	115.1	118.4	128.9
	3	90.1	117.7	114.3	116.7	118.8	129.6
	4	91.4	116.7	115.2	115.3	119.1	130.6
	5	90.7	117.5	113.8	114.1	119.2	131.4
	6	89.9	116.7	118.4	122.7	116.2	130.6
	1'	97.5	117.9	115.2	116.0	118.3	130.6
	2'	91.3	117.2	114.7	115.8	119.0	129.6
	3'	91.1	117.4	113.4	116.1	119.0	130.6
	4'	91.3	118.0	116.8	114.4	118.2	130.1
	5'	90.2	118.8	116.5	113.1	118.3	130.8
	6'	89.1	116.8	118.1	122.7	115.9	130.9

Table S7: Chemical shielding obtained for constrained (hydrogen atoms only) geometry optimisations methods. Prime number refers to the numbering of the second chain of amylose in the double helix

	Residue number	C1	C2	C3	C4	C5	C6
<b>Opt<sub>HBN_1(H)</sub></b>	1	97.7	123.1	112.1	116.6	122.2	128.0
	2	91.3	123.2	112.4	117.0	120.9	126.2
	3	90.9	124.0	111.5	121.0	121.6	127.9
	4	89.7	123.3	111.9	116.4	120.7	126.1
	5	89.0	123.6	112.0	116.7	121.6	128.1
	6	89.9	122.9	114.3	116.5	118.5	127.9
	1'	97.9	123.4	112.0	116.6	122.2	128.1
	2'	89.8	123.4	112.0	116.4	120.8	126.1
	3'	90.3	123.7	111.9	116.3	121.8	127.9
	4'	89.9	123.4	111.9	116.1	120.6	126.0
	5'	90.3	123.7	111.9	121.0	121.5	129.7
	6'	89.9	122.9	113.4	116.2	118.5	129.6
<b>Opt<sub>HBN_2(H)</sub></b>	1	97.6	120.4	112.1	116.9	123.0	130.4
	2	91.1	120.7	112.3	117.9	120.7	128.9
	3	91.5	121.1	112.5	116.8	121.5	130.8
	4	90.0	121.1	114.6	118.3	120.2	128.9
	5	88.8	120.9	113.9	117.0	121.2	131.0
	6	89.1	121.4	114.3	123.9	119.1	128.2
	1'	97.8	120.8	112.5	117.0	122.8	130.6
	2'	90.1	121.4	114.8	118.1	120.3	129.1
	3'	90.1	121.0	114.0	116.9	121.4	130.9
	4'	89.6	121.3	115.1	117.6	120.3	127.3
	5'	89.7	121.5	114.1	117.7	121.5	131.2
	6'	89.2	121.4	114.2	124.1	119.2	130.4

Table S8: Molecular approach. Experimental chemical shift and average theoretical chemical shielding used for theory vs experience comparison and parameters of the linear fit. Sugar 1 and 2 refer to experimental data. OptHBN\_1, OptHBN\_2 or OptHBN\_1 (H only), OptHBN\_2 (H only) refer to geometry optimisations on all atoms or on only hydrogen atoms for the HBN\_1 and HBN\_2 structures, respectively. Mol. Dyn. refers to calculation performed on structures extracted from molecular dynamics simulations.

	Sugar 1	Opt <sub>HBN_1</sub>	Opt <sub>HBN_2</sub>	Opt <sub>HBN_1</sub> (H)	Opt <sub>HBN_2</sub> (H)	Mol. Dyn.
C1	<b>99.3</b>	89.1	90.5	90.1	89.9	80.9
C2	<b>71.3</b>	119.0	117.4	123.4	121.1	111.0
C3	<b>75.1</b>	118.0	115.6	112.3	113.7	112.6
C4	<b>72.6</b>	113.0	115.2	116.5	117.4	110.0
C5	<b>69.8</b>	118.3	118.3	120.9	120.9	114.2
C6	<b>61.3</b>	127.9	130.3	127.6	129.8	127.9
R <sup>2</sup>		0.965	0.987	0.965	0.994	0.974
Slope		-0.954	-0.978	-0.944	-0.941	-0.821
Intercep		183.9	187.0	183.6	183.5	164.8
	Sugar 2	Opt <sub>HBN_1</sub>	Opt <sub>HBN_2</sub>	Opt <sub>HBN_1</sub>	Opt <sub>HBN_2</sub>	Mol. Dyn.
C1	<b>100.2</b>	89.1	90.5	90.1	89.9	80.9
C2	<b>72.1</b>	119.0	117.4	123.4	121.1	111.0
C3	<b>75.1</b>	118.0	115.6	112.3	113.7	112.6
C4	<b>74.3</b>	113.0	115.2	116.5	117.4	110.0
C5	<b>70.2</b>	118.3	118.3	120.9	120.9	114.2
C6	<b>61.3</b>	127.9	130.3	127.6	129.8	127.9
R <sup>2</sup>		0.976	0.993	0.958	0.990	0.984
Slope		-0.976	-0.998	-0.957	-0.955	-0.839
Intercep		187.0	189.9	185.7	185.8	167.4

Table S9: Periodic Boundary conditions approach. Experimental chemical shift and theoretical chemical shielding (calculated for  $\text{sugar}_{\text{interface}}$  and  $\text{sugar}_{\text{solvent}}$  residues in HBN\_1 and HBN\_2 optimized structures or for their averaged over the structures) used for theory vs experience comparison associating (a) sugar1 with  $\text{sugar}_{\text{interface}}$  and sugar2 with  $\text{sugar}_{\text{solvent}}$  and (b) sugar1 with  $\text{sugar}_{\text{solvent}}$  and sugar2 with  $\text{sugar}_{\text{interface}}$ . Parameters for each linear fit.

		Exp.	HBN_1 (a)	HBN_1 (b)	HBN_2 (a)	HBN_2 (b)	Average 1 (a)	Average 2 (b)
Sugar 1	C1	99.3	-113.59	-114.11	-112.7	-112.3	-113.2	-113.1
	C2	71.3	-134.32	-131.10	-133.6	-133.8	-133.9	-131.7
	C3	75.1	-130.84	-130.77	-133.6	-131.4	-132.2	-131.0
	C4	72.6	-132.11	-133.23	-131.5	-131.4	-131.8	-132.9
	C5	69.8	-135.04	-134.40	-132.9	-134.4	-134.0	-135.3
	C6	61.3	-142.50	-142.15	-140.7	-141.4	-141.6	-141.0
Sugar 2	C1	100.2	-114.11	-113.59	-112.3	-112.7	-113.2	-112.4
	C2	72.1	-131.10	-134.32	-133.8	-133.6	-132.4	-134.5
	C3	75.1	-130.77	-130.84	-131.4	-133.6	-131.1	-131.6
	C4	74.3	-133.23	-132.11	-131.4	-131.5	-132.3	-131.4
	C5	70.2	-134.40	-135.04	-134.4	-132.9	-134.4	-133.4
	C6	61.3	-142.15	-142.50	-141.4	-140.7	-141.8	-141.0
R <sup>2</sup>			0.9861	0.9853	0.9854	0.9855	0.9939	0.9936
Slope			1.4	1.1	1.3	1.3	1.4	1.4
Intercept			254.1	254.1	250.2	250.2	253.6	253.6

Table S10: Periodic Boundary conditions approach. Experimental chemical shift and theoretical chemical shielding (calculated for  $\text{sugar}_{\text{interface}}$  and  $\text{sugar}_{\text{solvent}}$  residues in HBN\_3 and HBN\_4 optimized structures) used for theory vs experience comparison associating (a) sugar1 with  $\text{sugar}_{\text{interface}}$  and sugar2 with  $\text{sugar}_{\text{solvent}}$  and (b) sugar1 with  $\text{sugar}_{\text{solvent}}$  and sugar2 with  $\text{sugar}_{\text{interface}}$ . Parameters for each linear fit.

		Exp.	HBN_3 (a)	HBN_3 (b)	HBN_4 (a)	HBN_4 (b)
Sugar 1	C1	99.3	-113.8	-113.5	-112.4	-113.1
	C2	71.3	-133.8	-132.6	-134.5	-131.7
	C3	75.1	-132.7	-131.8	-131.6	-131.0
	C4	72.6	-134.2	-132.9	-131.4	-132.9
	C5	69.8	-134.9	-136.2	-133.4	-135.3
	C6	61.3	-141.3	-141.6	-141.0	-141.0
Sugar 2	C1	100.2	-113.5	-113.8	-113.1	-112.4
	C2	72.1	-132.6	-133.8	-131.7	-134.5
	C3	75.1	-131.8	-132.7	-131.0	-131.6
	C4	74.3	-132.9	-134.2	-132.9	-131.4
	C5	70.2	-136.2	-134.9	-135.3	-133.4
	C6	61.3	-141.6	-141.3	-141.0	-141.0
R <sup>2</sup>			0.9939	0.9890	0.9890	0.9911
Slope			1.4	1.1	1.3	1.3
Intercept			255.6	255.1	252.3	252.5

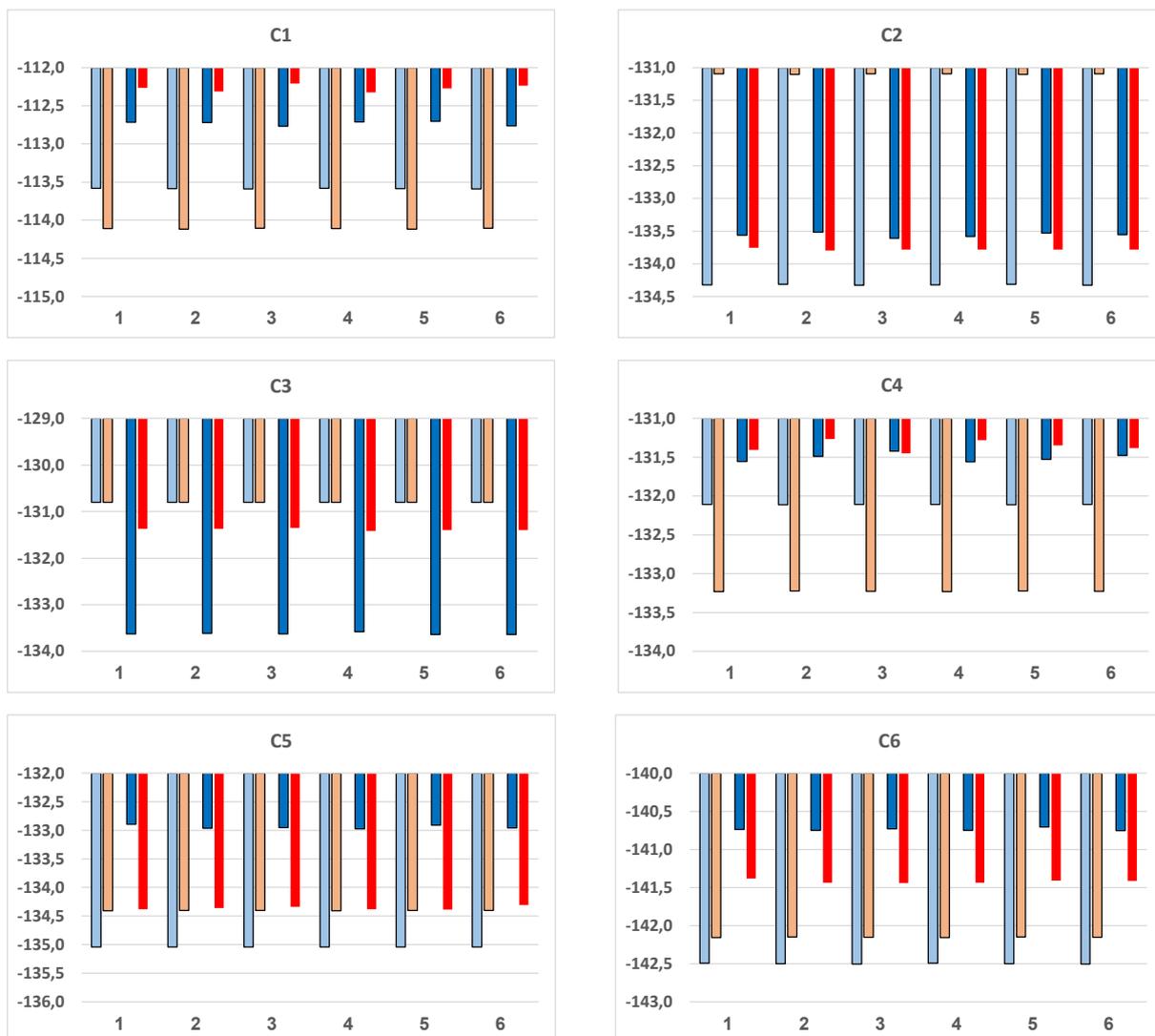


Figure S1: Carbon chemical Shielding obtained after GIPAW calculations on HBN\_1 and HBN\_2 optimised structures. Light blue: HBN\_1, Sugar<sub>interface</sub>. Light orange: HBN\_1, Sugar<sub>solvent</sub>. Dark Blue: HBN\_2 Sugar<sub>interface</sub>. Red : HBN\_2, Sugar<sub>solvent</sub>.