HAL
open science

Modélisation gros grain de macromolécules végétales : champ de force paramétré par dynamique moléculaire et application à des assemblages cellulose-xylane

Liang Li

To cite this version:

Liang Li. Modélisation gros grain de macromolécules végétales : champ de force paramétré par dynamique moléculaire et application à des assemblages cellulose-xylane. Biotechnologies. AgroParisTech, 2013. Français. NNT : 2013AGPT0087 . tel-01137208

HAL Id: tel-01137208

https://pastel.hal.science/tel-01137208

Submitted on 30 Mar 2015

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

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

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

L'Institut des Sciences et Industries du Vivant et de l'Environnement

(AgroParisTech)

Spécialité : Génie des Procédés et des Produits

présentée et soutenue publiquement par

Liang LI

le 20 Décembre 2013

Modélisation gros grain de macromolécules végétales : champ de force paramétré par dynamique moléculaire et application à des assemblages cellulose-xylane.

Directeur de thèse : **Patrick PERRE**
Co-encadrement de la thèse : **Xavier FRANK**

Jury

M. André MERLIN, Professeur, LERMAB, Université de Lorraine	Président
M. Bernard KUREK, Directeur de recherche, FARE, INRA, URCA	Rapporteur
M. Jean-Yves DELENNE, Directeur de recherche, IATE INRA, CIRAD, SupAgro, UMII	Rapporteur
M. Patrick PERRE, Professeur, LGPM, Ecole Centrale Paris	Examinateur
M. Xavier FRANK, Chargé de recherche, IATE INRA, CIRAD, SupAgro, UMII	Examinateur
M. Karim MAZEAU, Directeur de recherche, CERMAV CNRS, Université Joseph Fourier	Examinateur
M. Olivier VITRAC, Chargé de recherche, I2MC, AgroParisTech, INRA	Examinateur

AgroParisTech
Laboratoire d'Etude des Ressources Forêt-Bois
14 Rue Girardet, 54000 Nancy France

REMERCIMENTS

Je tiens particulièrement mes remerciements sincères et distingués au directeur de la thèse Monsieur Patrick Perré, le directeur du LGPM de l'Ecole Centrale Paris, qui m'a recruté pour la réalisation de cette thèse et m'encadrait au sein de l'équipe 3B du LERFoB de l'AgroParisTech ENGREF. Je le remercie de tout cœur pour son soutien consistant de toute manière tout au long de la thèse, pour sa confiance accordée à moi, ainsi pour sa grande personnalité en tant qu'un directeur de thèse duquel je suis fier d'être étudiant.

Je remercie à Monsieur André Merlin, professeur de LERMAB, me honorait d'avoir présidé le jury. Je remercie les rapporteurs de cette thèse Bernard Kurek, directeur du laboratoire FARE INRA/URCA, et Jean-Yves Delenne, directeur de recherche INRA, pour la rapidité et la patience avec lesquelles ils ont lu le manuscrit et l'intérêt qu'ils ont porté à mon travail. Je remercie à Monsieur Karim Mazeau, directeur de recherche du CERMAV, de ses collaborations et directions indispensables pour la réalisation de la thèse. Je remercie à Monsieur Xavier Frank, encadrant de la thèse, de sa direction scientifique et de ses aides indispensables à l'optique de la correction du manuscrit la thèse. Je remercie à Monsieur Olivier Vitrac de ses commentaires de la thèse en tant qu'être membre du jury.

Mes remerciements vont aussi à Madame Meriem Fournier pour ses aides importants au niveau du dépôt de la thèse. J'adresse toute ma gratitude à Madame Corinne Martin qui m'aidait tout au long de la thèse, je la remercie pour le climat chaleur qu'elle apportait à notre équipe par ses multi-talents surtout par sa compétence pâtisserie.

Mes remerciements vont également à mes chers amis de l'équipe 3B pour leur accompagnement très agréable pendant mon séjour à Nancy. Je tiens particulièrement mes remerciements aux amis de mon bureau à Montpellier de leurs encouragements et supports. J'exprime mes remerciements à Nouredine Tamani de ses aides à l'optique de la relecture de la thèse ainsi ses conseils divers.

Je remercie à mes parents qui n'hésitent jamais de me supporter, sans condition et de toutes manières, ma carrière d'étude depuis toujours. Je les remercie de leurs encouragements et de la confiance qu'ils accordent à moi. Je remercie à Yan de son accompagnement et discussions qui ont été pour moi des réconforts et parfois de très bonnes déviations de mes agitations.

Résumé

La compréhension de la relation structure-propriétés des parois des cellules végétales s'appuie de plus en plus sur l'utilisation d'approches de modélisation moléculaire en général et de dynamique moléculaire en particulier. A ce jour, le poids numérique que représente une telle démarche à l'échelle de l'atome est la plupart du temps incompatible avec les puissances de calcul disponibles. C'est pourquoi des méthodes d'approximation sont indispensables pour pouvoir mettre en œuvre des simulations numériques à l'échelle de systèmes supramoléculaires réalistes. Dans le cadre de cette thèse, un modèle de dynamique moléculaire, dit « gros grain » a été mis au point à l'échelle du monomère de macromolécules pariétales. Les paramètres de ce modèle ont été calibrés à l'aide de simulations de dynamique moléculaire à l'échelle de l'atome. Ce modèle a fait l'objet de quatre applications : adsorption d'une chaine de xylane sur une surface de cellulose cristalline, arrachement d'une chaine de xylane adsorbée sur une surface de cellulose cristalline par une pointe AFM, adsorption d'une phase amorphe de xylane sur une surface de cellulose cristalline et adsorption d'une phase amorphe de xylane sur un monocristal de cellulose exposant trois surfaces différentes. Des effets de structuration au voisinage de la cellulose sont observés.

Mots clés : Dynamique moléculaire gros grain, dynamique moléculaire atomistique, inversion de Boltzmann, Force-Matching, xylane, cellulose.

Abstract

Nowadays, the understanding of plant cell walls' structure-properties relationship leans more and more on the use of molecular modeling approaches and of molecular dynamics in particular. To date, numerical weight of such an approach is usually out of the reach of available computing power if the atomic scale is used. As a consequence, building approximate methods is of crucial importance to perform numerical simulation of realistic supramolecular systems. Within the framework of this PhD, a "coarse grain" molecular dynamics model was built at plant cell wall macromolecule monomer's scale, it's parameters being fixed with the help of atom-scale molecular dynamics simulations. Then, several numerical studies were carried out: a single xylan chain was adsorbed on a crystalline cellulose surface, a single xylan chain was pulled from a crystalline cellulose surface with the help of the tip of an AFM cantilever, an amorphous xylan phase was adsorbed on a cellulose surface and an amorphous xylan phase was adsorbed on a cellulose crystal, which three surfaces were exposed. Local structuring effects were observed.

Key words: Coarse grain molecular dynamics, atomistic molecular dynamics, Boltzmann inversion, force matching, xylan, cellulose.

Table des matières

Introduction Générale 5

I. Bibliographie 7

Résumé du chapitre I 8

I.1 Dynamique moléculaire 9

I.1.1 Modélisation multi-échelle des matériaux 9

I.1.2 Principe de la dynamique moléculaire 10

I.1.3 Termes d'interaction, champs de force 11

I.1.4 Intégrateur Verlet-Vitesse 14

I.1.5 Ensembles statistiques, thermostats 16

I.1.6 Recherche de voisins 19

I.1.7 Conditions aux limites périodiques 22

I.2 Modélisation Gros Grain 24

I.2.1 Coarse-Graining 24

I.2.2 Champs de forces gros grain 25

I.3 Composition des cellules du bois 28

I.3.1 Xylologie du bois 28

I.3.2 Structure de la cellule du bois 30

I.3.3 Constituants chimiques de la cellule du bois 31

I.4 Modélisation des assemblages de polymères des parois végétales 35

I.4.1 Modèles atomistiques 35

I.4.2 Modèles gros grain 36

II. Modèle Gros Grain 39

Résume du chapitre II 40

II.1 Matériels 41

II.1.1 Modèles atomistiques 41

II.1.2 Logiciel de dynamique moléculaire 47

II.1.3 Logiciel de visualisation VMD 48

II.2 Définition des grains 49

II.2.1 Stratégie de conversion CG 49

II.2.2 Gain et Compromis 50

II.2.3 Modèles Gros Grain 51

II.2.4 Fixation de la cellulose 53

II.3 Potentiels intramoléculaires du xylane 55

II.3.1 Procédures de modélisation 55

II.3.2 Modèles du xylane de 5, 10 et 25 xyloses 55

II.3.3 Calibration de modèles de potentiels intramoléculaires 63

II.3.4 Validation des potentiels intramoléculaires 65

II.4 Potentiel intermoléculaire xylane-xylane 68

II.4.1 Dérivation de la distribution cible 69

II.4.2 Ajustement à posteriori du potentiel xylane-xylane 72

II.5 Potentiel intermoléculaire xylane-cellulose 75

II.5.1 Traitement de la trajectoire atomistique 75

II.5.2 Calibration initiale du modèle de potentiel 76

II.5.3 Optimisation du potentiel 79

II.6 Conclusions 83

III. Applications et Résultats 85

Résumé du chapitre III 86

III.1 Constructions des configurations initiales 87

III.1.1 Générations des chaînes de xylane 87

III.1.2 Générations des surfaces et cristaux de cellulose 96

III.1.3 Générations des systèmes xylane-cellulose 98

III.2 Interaction d'une chaine de xylane avec une surface de cellulose 107

III.2.1 Adsorption d'une chaîne de xylane sur une surface de cellulose 107

III.2.2 Arrachement d'une chaîne de xylane d'une surface de cellulose 117

III.3 Multicouche de xylane sur la cellulose 122

III.3.1 Phase amorphe de xylane sur une surface de cellulose 122

III.3.2 Phase amorphe de xylane sur un monocristal de cellulose 124

III.4 Conclusions 129

Conclusions et Perspectives 130

Conclusion 131

Perspectives 132

Références bibliographiques 135

Introduction Générale

La surconsommation de ressources fossiles non renouvelables, et en particulier de pétrole, fait de la recherche de sources alternatives d'énergie et de matériaux un enjeu industriel et environnemental majeur. Dans ce contexte, les bio-polymères sont actuellement l'objet d'un intérêt croissant dans la communauté scientifique. Parmi ces matériaux, la cellulose représente la première production mondiale en tonnage, les hémicelluloses étant d'une importance comparable. Cette production intensive est bien-évidemment dû au fait que ces polymères sont présents au sein de la paroi cellulaire des plantes, et en particulier, au sein du bois. Bien que l'étude des propriétés physico-chimiques de ces macromolécules aie fait l'objet de nombreux travaux, la compréhension, au niveau microscopique, de l'impact de la structure de ces molécules et de leurs interactions sur les propriétés macroscopiques des parois végétales est toujours incomplète.

Face aux obstacles propres à l'investigation expérimentale, la simulation numérique constitue un outil d'investigation complémentaire indispensable. Les approches numériques à l'échelle atomistique ont été fortement développées. Cependant, la taille des molécules à modéliser à l'échelle atomistique est incompatible avec la puissance de calcul disponible à ce jour, du fait du nombre gigantesque de degrés de liberté impliqué. Les possibilités des modèles atomistiques sont donc fortement contraintes.

La présente thèse s'inscrit parmi les travaux qui ont vocation à apporter des éléments de solution à ce problème, sous la forme d'une approche dite « gros grain ». Dans le cadre de ce type de modèle, des grains sont définis comme regroupements de plusieurs atomes, de façon à diminuer le nombre de degrés de liberté nécessaires pour décrire un même système. Il s'agit donc d'une modélisation basse résolution, par opposition aux modèles à haute résolution que sont les modèles atomistiques.

Un modèle gros grain sera donc défini et paramétré à l'aide de calculs à l'échelle atomistique. Ce modèle sera ensuite appliqué à différentes situations impliquant une interaction entre la cellulose cristalline et le xylane.

I. BIBLIOGRAPHIE

Résumé du chapitre I

Le chapitre I est consacré à l'exposé des méthodes génériques et de l'état de l'art.

En premier lieu, la dynamique moléculaire est présentée. Les différentes étapes d'une itération de dynamique moléculaire font chacune l'objet d'un développement. Les types de termes d'interaction impliqués, les algorithmes d'intégration du mouvement, les thermostats, ainsi que les méthodes de recherche de voisins sont successivement présentés. Enfin, le principe des approches gros grain, ainsi que deux méthodes de dérivation des potentiels gros grain sont détaillés.

En second lieu, la structure du matériau bois à différentes échelles, ainsi que ses constituants à l'échelle moléculaire, sont présentés.

Enfin, une revue de l'état de l'art de la modélisation moléculaire, à l'échelle atomistique et à l'échelle gros grain, des polymères de la paroi végétale et de leurs assemblages est présentée.

I.1 Dynamique moléculaire

I.1.1 MODELISATION MULTI-ECHELLE DES MATERIAUX

Une démarche de modélisation d'un matériau en vue de réaliser des simulations numériques peut prendre place à différentes échelles. Différentes échelles impliquent différentes approches numériques, en fonction de la physique qui est prédominante à l'échelle considérée et en fonction du niveau de finesse de description requis (Fig. I-1)

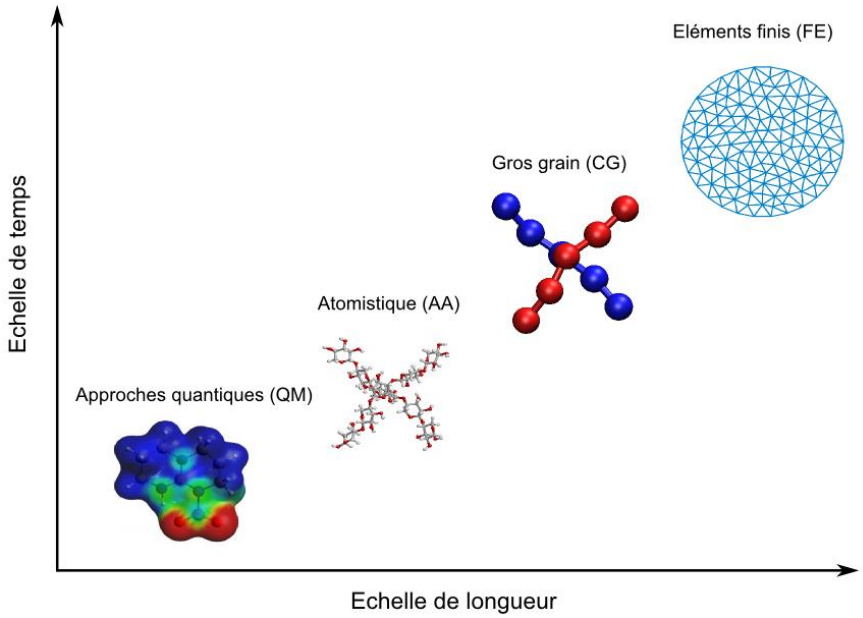

Figure I-1 Approches numériques à différentes échelles de longueur et de temps.

A petite échelle, la matière n'est correctement décrite que par la mécanique quantique. Les approches quantiques (QM, pour « Quantum Mechanics ») prennent en compte les modifications des orbitales électroniques des atomes lorsque ceux-ci sont en interaction. Ce type de simulation représente un poids numérique important et ne sera pas développé ici. Afin de simuler des systèmes moléculaires plus volumineux, on s'appuie sur une approximation classique, dite modélisation atomistique (AA, pour « All-Atom ») qui sera développée plus en

détail dans cette thèse. A l'échelle supramoléculaire, la prise en compte des détails à l'échelle de l'atome peut constituer une barrière, encore une fois du fait du poids numérique. Une autre méthode d'approximation peut alors être mise en place, une approche gros grain (CG, pour « Coarse Grain »). Il s'agit de traiter des ensembles d'atomes comme on traite des atomes dans les approches atomistiques. Cette voie est la clé de voûte de cette thèse. Enfin, à une échelle supérieure, la matière peut être considérée comme continue, et la description moléculaire n'est plus pertinente. C'est à ces échelles que sont appliquées les approches classiques du type éléments finis (FE, pour « Finite Elements »).

Deux échelles de la modélisation moléculaire, atomistique et CG, seront décrites dans cette section.

I.1.2 PRINCIPE DE LA DYNAMIQUE MOLECULAIRE

L'approximation classique à l'échelle de l'atome consiste à considérer que les atomes sont des points matériels interagissant avec les atomes voisins via des forces conservatives, c'est-à-dire dérivant de potentiels.

A partir de ce postulat de base, diverses approches ont été développées. On peut citer l'approche Monte Carlo [Metropolis and Ulam 1949, Metropolis *et al.* 1953, Flory 1969], une approche stochastique visant à générer un ensemble aléatoires de configurations du système compatible avec les contraintes de l'équilibre thermodynamique qui lui est imposé. Plus simplement, disposer de potentiels d'interaction permet également de rechercher les états d'énergie minimale du système, ce qui permet d'identifier les états les plus représentés statistiquement. Cette approche est largement utilisée dans le cadre des explorations conformationnelles de molécules.

La dynamique moléculaire (MD, pour « Molecular Dynamics »), consiste quant à elle à générer une trajectoire à partir d'un état initial du système [Alder and Wainwright 1957, Alder and Wainwright 1959]. Autrement dit, à partir de la position initiale et de la vitesse de tous les atomes constituant le système, et moyennant une intégration correcte des équations du mouvement, il est théoriquement possible de prédire ces mêmes variables à tout instant ultérieur. Les systèmes étant, d'une part, très idéalisés, et, d'autre part, fortement sensibles aux conditions initiales, les trajectoires en elles-mêmes n'ont pas de sens physique. Seules les grandeurs moyennes et corrélations, extraites des trajectoires, seront utiles. Les trajectoires résultantes peuvent être utilisées de façon à déterminer les propriétés macroscopiques du système moléculaire moyennant l'hypothèse ergodique : la valeur moyenne d'une grandeur calculée de manière statistique est identique à la moyenne des mesures prises dans le temps de simulation si un grand nombre d'itérations est imposé. Les informations à mesurer peuvent être extraites des trajectoires après ou pendant la simulation.

Le mouvement des atomes du système obéit donc à une équation très simple (Eq.I-1) :

$$m_i\vec{a}_i = \vec{f}_i(\vec{r}_i) = -\vec{\nabla}U_i(\vec{r}_i)$$ Équation I-1

On note m_i la masse de l'atome i, $\vec{a}_i$ son accélération et $\vec{r}_i$ sa position. La force exercée sur l'atome i par tous les autres atomes est notée $\vec{f}_i$ et $U_i(\vec{r}_i)$ est le potentiel d'interaction de l'atome i avec tous les autres.

Une itération de dynamique moléculaire comporte trois grandes étapes de calcul : le calcul des forces d'interaction entre les particules, l'intégration des équations de mouvement et enfin les contraintes extérieures imposées au système, en particulier celles qui fixent l'ensemble statistique. Chacun de ces points est développé ci-après.

I.1.3 TERMES D'INTERACTION, CHAMPS DE FORCE

En pratique, les potentiels d'interaction entre atomes sont décrits par des modèles dotés de paramètres ajustables. Il existe également des potentiels tabulés, mais il n'en est pas fait usage dans cette thèse. On appelle champ de forces l'ensemble des potentiels paramétrés pour la modélisation de l'énergie potentielle d'un système moléculaire particulier. Les potentiels paramétrés d'un champ de force sont déterminés à partir de résultats expérimentaux ou de calculs de mécanique quantique, voire les deux, et font souvent l'objet d'ajustements à posteriori. La mise au point d'un champ de force atomistique est une tâche complexe.

Les potentiels d'interaction peuvent être classés en deux grandes catégories : potentiels intramoléculaires et potentiels intermoléculaires (Eq.I-2) :

$$U_{total} = U_{intra} + U_{inter}$$ Équation I-2

$$U_{total} = U_{liaison} + U_{angle} + U_{dièdre} + U_{vdw} + U_{es}$$

Les interactions intramoléculaires s'exercent entre atomes d'une même molécule et topologiquement distants au plus d'un nombre limité de liaisons covalentes, trois en l'occurrence, dans le cas des potentiels de torsion, qui contraignent les angles dièdres. Les interactions intermoléculaires s'exercent entre atomes de molécules différentes et entre atomes d'une même molécule topologiquement distants au moins d'un nombre fixé de liaisons. Ces interactions sont détaillées ci-après. De nombreux modèles peuvent être utilisés pour décrire chacun de ces types d'interaction. Afin de simplifier le propos, seuls les modèles utilisés dans le cadre de cette thèse seront présentés.

I.1.3.1 Potentiels intramoléculaires

POTENTIEL D'ELONGATIONDE LIAISON

Il s'agit d'une interaction très forte contraignant la longueur des liaisons entre atomes voisins au sein d'une même molécule (Fig. I-2).

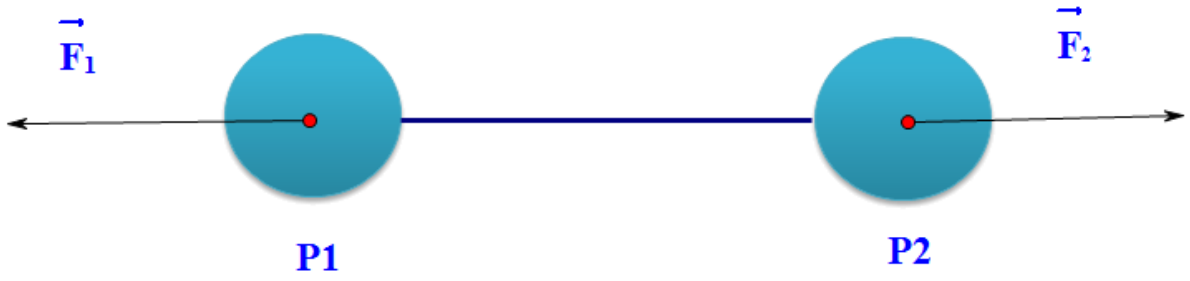

Figure I-2 Interaction d'élongation de liaison.

Le potentiel harmonique (Eq. I-3) est un modèle populaire pour ce type d'interaction :

$$U_{liaison}(r) = \frac{1}{2}k(r - r_0)^2 \qquad \text{Équation I-3}$$

Où r est la distance entre deux atomes, r_0 est la distance d'équilibre, k est l'intensité de l'interaction.

POTENTIEL DE FLEXION

Ce type d'interaction contraint l'angle formé par deux liaisons (Fig.I-3). Il s'agit donc d'une interaction entre trois atomes.

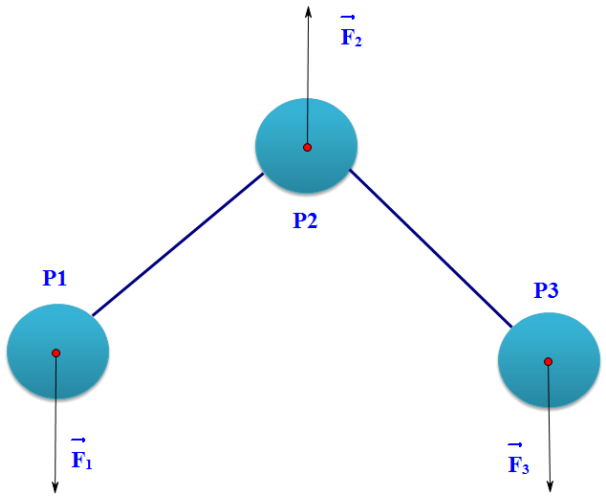

Figure I-3 Interaction de flexion.

Dans le cadre de cette thèse, ce type d'interaction est modélisé par un potentiel de type différence de cosinus au carré (Eq. I-4) :

$$U_{angle}(\theta) = k_{ang} \cdot (cos\theta - cos\theta_0)^2$$ Équation I-4

Où θ est l'angle entre deux liaisons reliant trois particules consécutives, k_{ang} est l'intensité du potentiel, θ_0 est l'angle à équilibre.

POTENTIEL DE TORSION

Ce type d'interaction contraint l'angle dièdre formé par trois liaisons entre quatre atomes consécutifs (Fig. I-4). L'angle dièdre est défini comme l'angle entre le plan $P_1P_2P_3$ et le plan $P_2P_3P_4$.

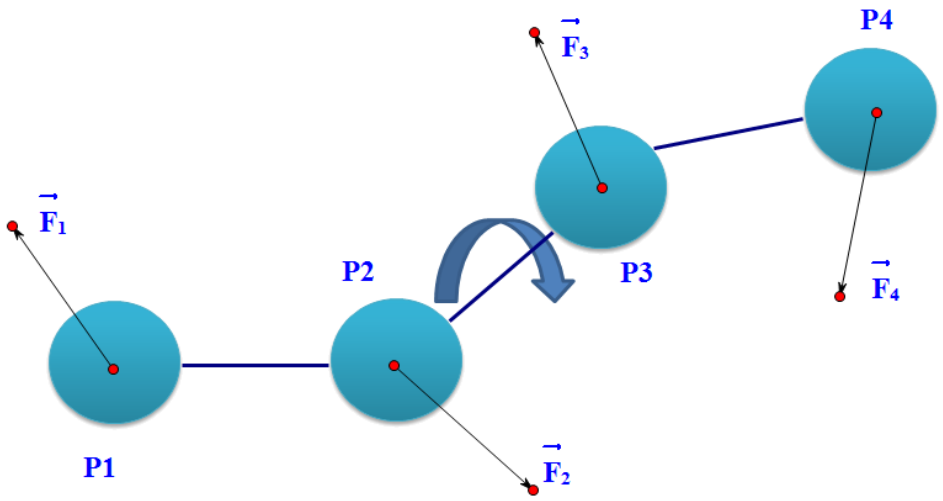

Figure I-4 Interaction de torsion.

Dans le cadre de cette thèse, ce type d'interaction est modélisé par un potentiel sous la forme d'une série de Fourier (Eq. I-5) :

$$U_{dièdre}(\varphi) = \sum_{n=1}^{n=3} \frac{U_n}{2} \cdot (1 - \cos(n \cdot \varphi - \varphi_n))$$ Équation I-5

Où φ est l'angle dièdre, et pour chaque terme de la série, φ_n est l'angle de phase qui dépend de l'origine choisie pour les angles dièdres, n est la périodicité du potentiel, U_n est l'intensité.

I.1.3.2 Potentiels intermoléculaires

POTENTIEL DE MORSE

Le potentiel de Morse (Eq. I-6) est un modèle initialement utilisé pour la description des molécules diatomiques, mais son usage a été étendu aux interactions intermoléculaires.

$$U(r) = D_e \cdot (1 - e^{-\alpha(r-r_0)})^2$$ Équation I-6

Où r est la distance entre les deux atomes, r_0 est la distance à l'équilibrer, De est l'intensité de l'interaction et α fixe la largeur du puits de potentiel.

POTENTIEL LEONARD-JONES 12-6

Ce modèle (Eq. I-7) est utilisé pour décrire les interactions de Van der Waals :

$$U_{vdw}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$ Équation I-7

Où ε est l'intensité de l'interaction, σ est la distance d'équilibre, r est la distance entre les deux atomes considérés.

POTENTIEL ELECTROSTATIQUE

L'interaction électrostatique est décrite par le potentiel de Coulomb (Eq. I-8) :

$$U(r) = \frac{1}{4\pi\varepsilon_0}\frac{Q_1 Q_2}{r}$$ Équation I-8

Où ε_0 la permittivité électrique du vide, r est la distance entre les deux atomes en interaction, Q_1 et Q_2 sont des charges électriques de ces deux atomes.

On peut noter ici que ce type d'interaction n'est pas utilisé que pour décrire des charges électriques réelles, par exemple celles portés par les ions. Dans la plupart des champs de force atomistiques, des charges partielles, fixées au niveau des atomes, permettent de tenir compte du caractère électronégatif ou électropositif de tel ou tel atome ou groupe fonctionnel. Notamment, et c'est particulièrement important ici, dans les modèles simples, les liaisons hydrogènes sont prises en compte sous la forme de charges électriques partielles portées par les atomes concernés. Il est clair que cette description est très approchée et ne prends pas en compte correctement le caractère directionnel des liaisons hydrogènes.

I.1.4 INTEGRATEUR VERLET-VITESSE

L'intégration numérique des équations du mouvement implique une discrétisation du temps. C'est pourquoi des erreurs cumulatives sont générées. Il est donc indispensable de choisir un algorithme d'intégration minimisant ces erreurs.

L'algorithme de Verlet [Verlet 1967], peut être déduit de deux développements (Eq. I-9, Eq. I-10) de Taylor au 3e ordre, par exemple, pour la coordonnée *x(t)* en fonction du temps, aux instants t-Δt et $t+\Delta t$:

$$x(t+\Delta t) = x(t) + \Delta t \cdot v_x(t) + \frac{1}{2} a_x(t) \cdot \Delta t^2 + \frac{1}{6} b_x(t) \cdot \Delta t^3 + O_{erreur}(\Delta t^4)$$ Équation I-9

$$x(t-\Delta t) = x(t) - \Delta t \cdot v_x(t) + \frac{1}{2} a_x(t) \cdot \Delta t^2 - \frac{1}{6} b_x(t) \cdot \Delta t^3 + O_{erreur}(\Delta t^4)$$ Équation I-10

Avec *v*, *a* sont respectivement la vitesse et l'accélération, *b* est la dérivée troisième de *x* par rapport au temps, et Δt le pas de temps. En sommant les deux expressions précédentes, nous obtenons :

$$x(t+\Delta t) = 2 \cdot x(t) - x(t-\Delta t) + a_x(t) \cdot \Delta t^2 + O_{erreur}(\Delta t^4)$$ Équation I-11

La nouvelle coordonnée d'une particule est calculée à partir de la position au moment *t*, la position au moment t-Δt et l'accélération au moment *t*. Dans cette version originale de l'algorithme de Verlet, la vitesse n'est pas directement générée, elle doit être calculée par ailleurs selon l'équation :

$$v_x(t) = \frac{x(t+\Delta t) - x(t-\Delta t)}{2\Delta t} + O_{erreur}(\Delta t^2)$$ Équation I-12

Il est à noter que l'erreur numérique associée à l'équation (Eq.I-12) est du 2e ordre au lieu du 4e, ce qui peut impliquer des difficultés lorsqu'il est important de disposer de la valeur des vitesses. D'autre part, l'initialisation de cet algorithme est problématique, dans la mesure où on ne peut pas imposer directement une vitesse. Enfin, cette approche impose de conserver en mémoire vive les coordonnées à trois instants successifs.

L'algorithme dit « Verlet-vitesse » permet des résoudre ces difficultés sans impacter la précision de l'intégration des équations du mouvement.

L'algorithme de Verlet-Vitesse manipule directement la vitesse, contrairement à la version originale. Les équations de I-13 jusqu'à I-15 composent le schéma explicite de l'algorithme Verlet-Vitesse :

$$x(t+\Delta t) = x(t) + v_x(t) \cdot \Delta t + \frac{1}{2} \cdot a_x(t) \cdot \Delta t^2$$ Équation I-13

$$v_x\left(t + \frac{1}{2} \cdot \Delta t\right) = v_x(t) + \frac{1}{2} \cdot a_x(t) \cdot \Delta t$$ Équation I-14

$$v_x(t+\Delta t) = v_x\left(t+\frac{1}{2}\cdot\Delta t\right)+\frac{1}{2}\cdot a_x(t+\Delta t)\cdot\Delta t$$ Équation I-15

On peut noter que le calcul de la vitesse à l'instant $t+\frac{1}{2}\Delta t$ n'est pas utile en soi lorsque la force ne dépend pas des vitesses. Dans ce cas, les équations I-13 et I-15 peuvent être simplifiées en une équation unique :

$$v_x(t+\Delta t) = v_x(t)+\frac{1}{2}\cdot[a_x(t)+a_x(t+\Delta t)]\Delta t$$ Équation I-16

I.1.5 Ensembles statistiques, thermostats

Dans beaucoup d'applications, il est indispensable d'imposer des conditions thermodynamiques au système. Il faut donc forcer la trajectoire à être compatible avec l'ensemble statistique choisi. Un ensemble est défini par les grandeurs qui y sont conservées : le nombre d'atomes N, le volume V, la pression P et la température T. On peut citer trois ensembles thermodynamiques très courants dans les simulations de dynamique moléculaire : l'ensemble microcanonique *NVE* ; l'ensemble canonique *NVT* ; l'ensemble isotherme et isobare *NPT*. Les simulations MD effectuées dans le cadre de cette thèse sont basées sur l'ensemble canonique *NVT*.

L'application correcte (en termes de pas de temps) d'un algorithme de Verlet ou Verlet-vitesse à un nombre fixe de particules dans une boîte de simulation de volume fixe permet de remplir les conditions de l'ensemble *NVE*. En revanche, imposer les conditions d'autres ensembles nécessite l'utilisation de méthodes plus sophistiquées : thermostats pour fixer la température, barostats pour fixer la pression, etc… Seuls des thermostats ont été utilisés dans le cadre de cette thèse.

La température est un paramètre qui mesure l'énergie cinétique instantanée d'un système moléculaire. Celle-ci peut donc être définie à l'aide des vitesses atomiques :

$$T = \frac{\sum_{i=1}^{N} m_i {v_i}^2}{3Nk_B}$$ Équation I-17

Où N est le nombre de particules, m_i et v_i sont respectivement la masse et la vitesse de la particule i, et k_B la constante de Boltzmann. Le thermostat est un dispositif servant à maintenir la température constante dans le système moléculaire. Quelques thermostats usuels sont présentés ici.

I.1.5.1 Thermostat de Berendsen

Le thermostat de Berendsen [Berendsen *et al.* 1984] fixe l'énergie cinétique du système moléculaire par renormalisation directe des vitesses. Il est extrêmement efficace pour thermaliser les systèmes hors équilibre à la température que l'on souhaite :

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}$$ Équation I-18

$$\lambda^2 = 1 + \frac{\Delta t}{\tau}\{\frac{T_0}{T} - 1\}$$ Équation I-19

Où λ est le facteur de modification des vitesses des particules du système moléculaire, τ la constante de la puissance du thermostat, T_0 est la température du thermostat, T est la température instantanée du système, Δt est le pas de temps.

Cependant, une fois le système équilibré, ce couplage ne permet de préserver aucun ensemble canonique car ce thermostat supprime la fluctuation de la température.

I.1.5.2 Thermostat d'Andersen

Le thermostat d'Andersen [Andersen 1980] est le thermostat le plus simple capable de produire des ensembles canoniques corrects. Il s'agit de sélectionner de manière aléatoire dans chaque pas de temps un certain nombre de particules et de modifier leurs vitesses selon la distribution de Maxwell-Boltzmann à température T (Eq. I-20) :

$$P(p) = (\frac{1}{2\pi m k_B T})^{3/2} \cdot \exp(-p^2/2mk_BT)$$ Équation I-20

Avec P la distribution de probabilité de la vitesse, p la quantité de mouvement. Le thermostat Andersen n'impose pas que la température soit modifiée à chaque pas de temps, mais plutôt d'intervenir tous les $1/v$ pas de temps, où v est la fréquence de couplage avec le réservoir de chaleur.

Le thermostat d'Andersen peut générer des ensembles canoniques corrects. Cependant, les autocorrélations des vitesses sont détruites, ce qui interdit l'usage de ce thermostat pour calculer des coefficients de diffusion.

I.1.5.3 Thermostat Nosé-Hoover

Le thermostat Nosé-Hoover, introduit par Nosé [Nosé 1984[a], Nosé 1984[b]], puis amélioré par Hoover [Hoover 1985], est un thermostat déterministe produisant une statistique correcte des ensembles canoniques.

En pratique, les équations du mouvement sont modifiées (Eq. I-21 Eq. I-22, Eq. I-23) par un terme semblable à un terme de dissipation visqueuse de viscosité variable :

$$\frac{d\vec{r}_i}{dt} = \frac{\vec{p}_i}{m_i}$$ Équation I-21

$$\frac{d\vec{p}_i}{dt} = \vec{F}_i - \varepsilon \cdot \vec{p}_i$$ Équation I-22

$$\dot{\varepsilon} = \frac{1}{Q}\left[\sum_{i=1}^{N} \frac{{p_i}^2}{m_i} - D_f k_B T_o\right]$$ Équation I-23

Où $\vec{p}_i$ est la quantité de mouvement de l'atome i, m_i la masse de l'atome i, D_f est le nombre de degrés de liberté, T_0 la température du thermostat et Q un paramètre ajustable, appelé masse virtuelle. Notons que le paramètre ε est identique pour toutes les particules d'un système. Lorsque la température instantanée du système est plus grande que T_0, le paramètre ε augmente et prends une valeur positive. Le terme additionnel est alors équivalent à une dissipation visqueuse, et l'énergie du système diminue. Lorsque la température instantanée du système est moins grande que T_0, ε diminue, et lorsque la valeur de cette variable devient négative, l'énergie du système augmente.

Le paramètre Q fixe la fréquence d'oscillation de la température. Si la valeur de Q est trop grande, les fluctuations de température sont très grandes, en revanche, si la valeur de Q est trop faible, la température est correctement fixée, mais l'échantillonnage des micro-états est insuffisant et la statistique de l'ensemble canonique n'est plus correcte.

I.1.5.4 Thermostat Nosé-Hoover Chaîne

Dans certains cas, le thermostat Nosé-Hoover peut présenter des défauts d'ergodicité ou de convergence. Pour éviter de ce genre de problèmes, il est possible de chaîner plusieurs thermostats de Nosé-Hoover (Fig. I-5).

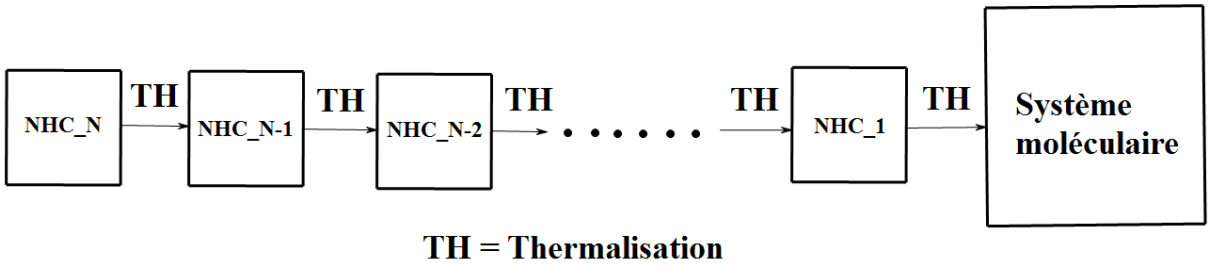

Figure I-5 Schéma du thermostat Nosé-Hoover Chainé.

Le thermostat Nosé-Hoover Chaîné [Martyna *et al.* 1992] est une chaîne de thermostats connectés en série dont un thermostat thermalise le suivant, ce dernier thermalise encore le suivant jusqu'au système moléculaire. Pour une longueur de chaîne M, on a :

$$\dot{\vec{p}_i} = \vec{F}_i - \varepsilon_1 \cdot \vec{p}_i$$ Équation I-24

$$\dot{\varepsilon}_1 = \frac{1}{Q_1}\left[\sum_{i=1}^{N}\frac{\mathrm{p_i}^2}{m_i} - \mathrm{N}k_B T_o\right] - \varepsilon_1\varepsilon_2$$ Équation I-25

$$\dot{\varepsilon}_k = \frac{1}{Q_k}[Q_{k-1}\varepsilon_{k-1}^2 - k_B T_o] - \varepsilon_\mathrm{k}\varepsilon_{\mathrm{k}+1}\,, \qquad \forall \mathrm{k} \in \{1, \cdots, \mathrm{M}-1\}$$ Équation I-26

$$\dot{\varepsilon}_M = \frac{1}{Q_M}[Q_{M-1}\varepsilon_{M-1}^2 - k_B T_o]$$ Équation I-27

I.1.6 RECHERCHE DE VOISINS

En dynamique moléculaire, le calcul de tous les termes de forces intermoléculaires est bien évidemment le plus consommateur de temps de calcul. D'autre part, sa complexité est avant tout due à la reconstruction régulière de la liste de voisins. Initialement, pour repérer les voisins d'une particule quelconque, il faut calculer les distances entre cette particule et toutes les autres particules. Ensuite, les particules situées à une distance inférieure à une valeur fixée sont sélectionnées. L'ordre de grandeur de la complexité de calcul pour chaque reconstruction de liste de voisins est N^2 (N est le nombre de particules dans le système). Cette opération numérique est extrêmement lourde lors que N est grand. La réduction significative de la durée de la simulation MD demande de :

1. Minimiser la fréquence de mise à jour de la liste de voisins.
2. Concevoir une structure efficace de tableau décrivant tous les voisins.
3. Minimiser la quantité de voisins à considérer pour chaque mise à jour.

Les deux premiers points peuvent être assurés par la liste de Verlet [Verlet 1967], le troisième point peut être obtenu par une approche « Link-Cell ».

I.1.6.1 Liste de Verlet

La liste de Verlet est la liste de voisins pour toutes les particules du système. On considère que les interactions intermoléculaires sont négligeables au-delà d'une distance appelée « *cutoff* » r_{cutoff}. Les voisins d'un atome sont donc, parmi tous les autres atomes, ceux qui sont situés à une distance inférieure au *cutoff*.

Dans le but de réduire la fréquence de reconstruction du voisinage, il faut noter que le changement de voisins d'une particule peut être négligé dans un pas de temps puisque le déplacement pour chaque particule est faible. Le nombre de pas de temps entre deux mises à jour de la liste de voisins doit être soumis à la condition que le déplacement maximal d'une particule quelconque du système moléculaire entier ne dépasse pas la moitié d'une distance

fixée δ_{skin}, appelée « épaisseur de peau ». En pratique, une reconstruction de voisinage ne s'appuiera pas sur le seul *cutoff*, mais sur la somme du *cutoff* et de l'épaisseur de peau :

$$r_{cutskin} = r_{cutoff} + \delta_{skin} \qquad \text{Équation I-28}$$

Dans le rayon r_{cutoff}, les particules sont comptées comme des voisines de la particule centrale. L'interaction entre la particule cible et ses voisines situées dans l'anneau entre r_{cutoff} et $r_{cutskin}$ est multipliée par un facteur variant selon la distance, de façon à avoir une transition régulière :

$$facteur = \left[1 - \frac{d - r_{cutoff}}{r_{cutskin} - r_{cutoff}}\right] \qquad \text{Équation I-29}$$

Avec *d* est la distance entre la particule centrale et une de ses voisines dans l'anneau. Ce facteur sert à réduire progressivement l'intensité de l'interaction, en sachant qu'il est impérativement nécessaire d'éviter l'effet de simple troncature qui cause la discontinuité du potentiel. Cette simple troncature du potentiel impacte la conservation d'énergie du système moléculaire pendant la simulation. La recherche de voisin ainsi que les paramètres *cutoff* et *cutskin* sont exposés dans (Fig. I-6).

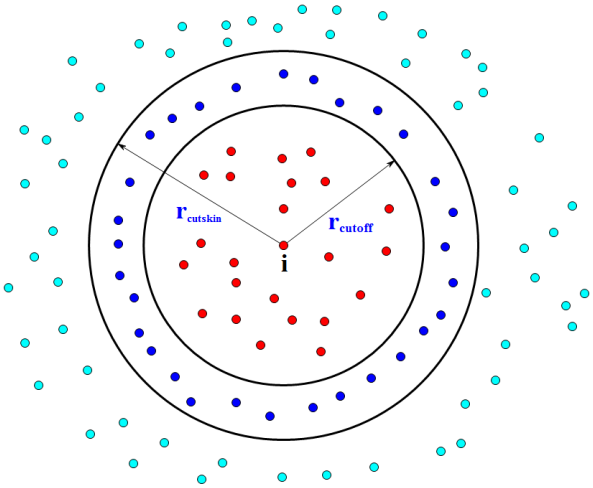

Figure I-6 Illustration de la recherche de liste de voisin pour la particule *i*, les particules rouges sont comptées directement comme voisines de la particule *i* ; les particules bleues sont comptées également comme voisines de la particule *i*, mais l'intensité de l'interaction sera multipliée par un facteur selon la distance entre elles et *i* ; les particules vertes ne sont pas comptées comme voisines de la particule *i*.

La liste de Verlet (Fig. I-7) est composée d'un tableau monodimensionnel indiquant pour chaque particule le nombre de ses voisines ainsi que leurs indices.

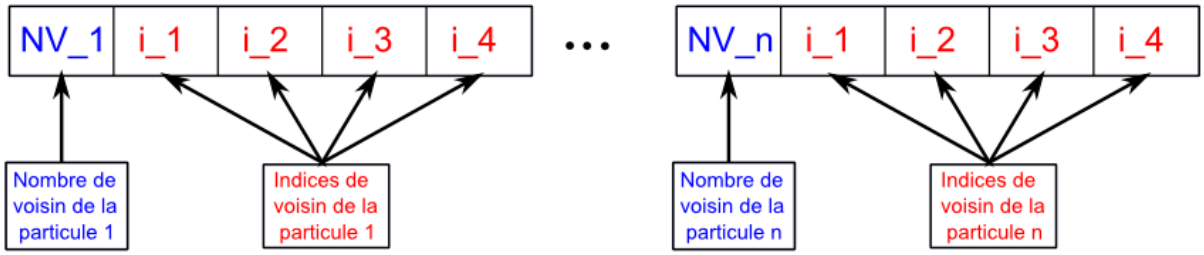

Figure I-7 La liste de Verlet est composée d'un tableau comportant autant de segments qu'il y a de particules dans le système. La case de début de chaque segment est le nombre de voisins de la particule ciblée, les autres cases du segment contiennent les numéros de voisines de la particule ciblée.

La liste de Verlet est formée par une série de segments dont le nombre correspond au nombre de particules. Chaque segment contient les informations de la particule correspondante, le premier élément d'un segment indique le nombre de voisins suivi de leurs indices. La taille d'un tel tableau n'est pas connue avant la recherche de voisins.

I.1.6.2 Link-Cell

La liste de Verlet permet de diminuer la quantité de calcul liée aux forces intermoléculaires, mais le volume de calcul nécessaire pour construire la liste est toujours de l'ordre de N^2. Afin de réduire le temps de calcul de la liste de Verlet jusqu'à la grandeur N, nous utilisons une approche de type « Link-Cell » [Allen and Tildesley 1989, Mattson and Rice 1999, Yao *et al.* 2004, Heinz and Hünenberer 2004, Gonnet 2012] pour la recherche de voisins. Il s'agit de découper le domaine en sous-cellules de nombre $M_x \times M_y \times M_z$ (Fig. I-8). En supposant que L_x, L_y, L_z sont les longueurs du domaine sur les trois directions, la taille d'une sous-cellule est S_x, S_y, S_z, sachant que la taille de sous-cellules doit impérativement être supérieure à la plus grande portée $r_{cutskin}$, nous avons donc :

$$S_x = \frac{L_x}{M_x}, S_y = \frac{L_y}{M_y}, S_z = \frac{L_z}{M_z} \qquad \text{Équation I-30}$$

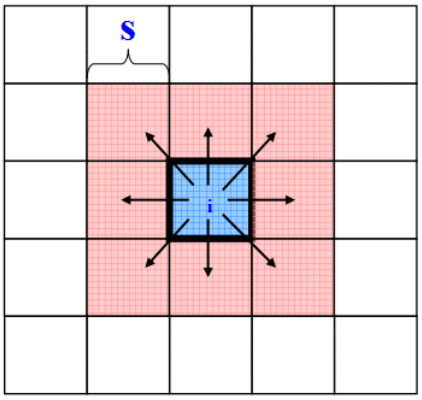

Figure I-8 Méthode Link-Cell : pour chercher les voisins de la particule i, au lieu de balayer le système moléculaire en entier, il suffit de chercher les voisins dans la case où se situe la particule *i* et toutes les cases de son voisinage. Donc, pour un système bidimensionnel, 9 cases sont à balayer ; pour un système tridimensionnel, 27 cases sont à balayer.

Ce découpage est en fait un regroupement de particules par leurs positions. Il classe les particules par sous-cellule. Après cette opération, au lieu de balayer le système moléculaire en entier pour chercher les voisins d'une particule, il suffit seulement de balayer la sous-cellule où la particule se trouve, ainsi que les 26 sous-cellules environnantes.

I.1.7 CONDITIONS AUX LIMITES PERIODIQUES

Les conditions aux limites périodiques (« PBC », pour « Periodic Boundary Conditions ») permettent de limiter les effets de bord [Sadus 2002]. Concrètement, ces conditions aux limites permettent de mettre en œuvre un système sans bords. Un système sans bords ne doit évidemment pas être confondu avec un système infini.

La figure I-9 illustre le cas bidimensionnel. Le système (au centre) est reproduit et translaté dans les deux directions de l'espace.

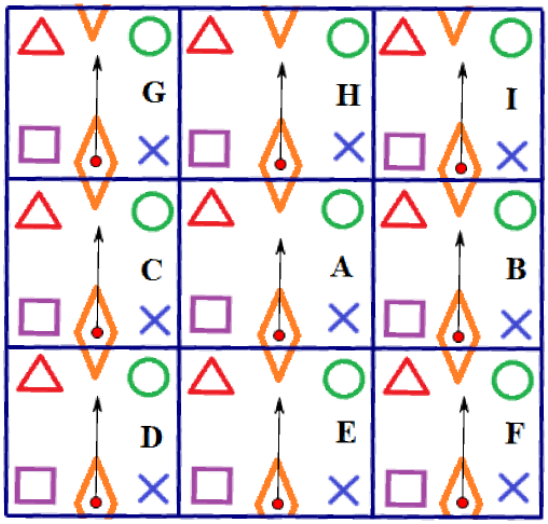

Figure I-9 Conditions aux limites périodiques : toutes les cases sont identiques ; la case centrale est la boîte de simulation.

Si on impose que la taille du domaine doit impérativement être supérieure à 2 fois r_{cutoff}, [Allen and Tildesley 1989, Haile 1992], une particule ne peut pas interagir avec elle-même ou avec l'une de ses images.

En pratique, il n'est évidemment pas nécessaire de reproduire numériquement tous les cases et calculer les mouvements des molécules qui s'y trouvent. Il suffit que la boîte de simulation soit légèrement plus grande que le domaine réel (Fig. I-10).

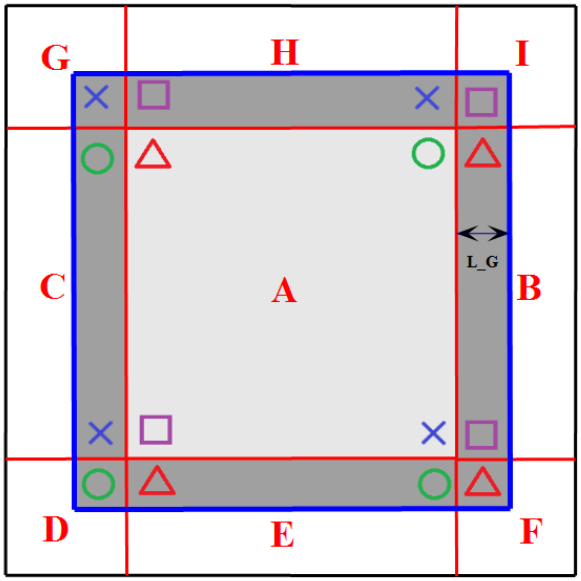

Figure I-10 Mise en place des PBC : exemple bidimensionnel. Les particules fantômes sont générées par translation dans toutes les directions.

Au lieu d'engendrer des images complètes du domaine, il s'agit simplement de reproduire et translater, pour chaque direction de l'espace, une partie du domaine suffisamment étendue pour s'assurer que le voisinage des particules proches du bord du domaine réel soit complet. La largeur des domaines fantômes L_G doit être supérieure à la plus grande portée d'interaction dans le système.

I.2 Modélisation Gros Grain

I.2.1 COARSE-GRAINING

Les modèles atomistiques de systèmes réalistes comportent un grand nombre de degrés de liberté, ce qui limite fortement l'échelle des systèmes qu'il est possible de simuler dans ce cadre. En pratique, il est extrêmement difficile, voire impossible d'exécuter des simulations atomistiques de systèmes supramoléculaires [Schulz *et al.* 2009]. Pour autant, prédire et comprendre la conformation de systèmes macromoléculaires est un enjeu crucial dans l'étude de la relation structure-propriété de nombreux matériaux. Une solution consiste alors à utiliser une méthode d'approximation : les approches gros grain.

L'utilisation d'approches gros grain repose sur l'hypothèse qu'il est possible de se focaliser sur les comportements mésoscopiques et de ne pas tenir compte, du moins explicitement, des conformations locales des atomes, autrement dit le modèle à construire ne nécessite pas une haute résolution atomistique du système moléculaire. Dans ce cas, on vise un compromis concernant la précision des détails moléculaires afin de réduire effectivement le nombre de degrés de liberté du système [Müller-Plathe 2002, Noid *et al.* 2008]. Les premiers modèles CG ont été introduits dans les années 1980 [Bird *et al.* 1987], dans le domaine de la physique des polymères. Cette approche est efficace pour la simulation de molécules de haut poids moléculaire comme les lipides, protéines et ADN [Tozzini 2005, Izvekov and Voth 2005, Knotts *et al.* 2007, Shelley *et al.* 2001, Shen and Gnanakaran 2009].

La modélisation gros grain consiste à considérer des groupes d'atomes qui interagiront les uns avec les autres comme les atomes dans les approches atomistiques. Cette approximation permet de réduire le nombre de particules du système moléculaire, et donc de réduire drastiquement le nombre de degrés de liberté. Par ailleurs, les liaisons entre grains étant souvent moins rigides que les liaisons entre atomes, le passage à l'échelle gros grain peut impliquer, dans certains cas, une augmentation significative du pas de temps maximal, ce qui réduit encore le temps de calcul.

Le modèle gros grain est généralement dérivé d'un modèle atomistique. La première démarche de construction d'un modèle gros grain consiste à définir les grains, c'est-à-dire à choisir quels atomes seront groupés. Ce choix comporte une part d'arbitraire et constitue un compromis entre accélération du calcul et perte d'information. On peut noter qu'un bon compromis dépend alors fortement du type de système considéré et des propriétés qu'on cherche à calculer et prédire. Un grain peut être un monomère [Srinivas *et al.* 2011], plusieurs monomères [Padding and Briels 2002], quelques atomes [Li *et al.* 2005, Tschöp *et al.* 1998], ou un ensemble plus complexe [Kremer 2003, Wang and Voth 2010]. En effet, le nombre d'atomes à grouper dans un grain peut être très grand, par exemple, dans le modèle de Curcó *et al.* [Curcó *et al.* 2007], chaque bloc de protéine, contenant 11 acides aminés, a été modélisé via un nombre de grains compris entre 2 et 4. Dans le cas particulier des polymères modélisés à partir de leurs monomères, Les grains peuvent être déployés aux centres géométriques [Reith *et al.* 2001, Reith *et al.* 2003] des monomères correspondants ou aux barycentres [Abrams and Kremer 2001], ce qui facilite le calcul des potentiels gros grains effectifs.

Le passage de l'échelle atomistique aux approches CG connaît un développement rapide dans la littérature, notamment dans le cas des bio-macromolécules. On peut citer des applications aux chaînes lipidiques [Izvekov and Voth 2005], aux peptides [Nguyen *et al.* 2004[a], Nguyen *et al.* 2004[b], Nguyen *et al.* 2004[c]], aux protéines [Basdevant *et al.* 2007, Fernández and Colubri, 2002, Buchete *et al.* 2004]. La modélisation CG est appliquée à divers types de macromolécules [Peter and Kremer, 2009, Izvekov and Voth 2005, Maupetit *et al.* 2007, Carbone *et al.* 2008, Marrink *et al.* 2004, Monticelli *et al.* 2008, Marrink *et al.* 2007].

I.2.2 CHAMPS DE FORCES GROS GRAIN

La plus grande difficulté de la modélisation CG réside dans la définition des interactions entre les grains, autrement dit, la définition du champ de forces CG. Les champs de force CG sont chimiquement dépendants, cela signifie qu'il n'existe pas de champ de force CG universel, et donc qu'il est indispensable de dériver les champs de forces CG de considérations atomistiques ou statistiques. La diversité des approches CG existant dans la littérature tient pour l'essentiel à la diversité des solutions apportées à cette question épineuse.

Les champs de force CG reposent en général sur des modèles de potentiel d'interaction très similaires à ceux des champs de force atomistiques, comportant des interactions intramoléculaires et des interactions intermoléculaires. Parmi toutes les approches mises en œuvre récemment, les deux plus fréquentes sont l'inversion de Boltzmann, à partir des distributions de probabilité issues des modèles atomistiques [Li *et al.* 2005] ou des données expérimentales, d'une part, et la méthode du Force-Matching [Izvokov and Voth 2005], d'autre part. L'inversion de Boltzmann consiste à déduire le potentiel à partir de la densité de probabilité de la grandeur statistique. Le Force-Matching consiste à calculer directement la force entre les particules ciblées.

I.2.2.1 Inversion de Boltzmann

L'inversion de Boltzmann [Reith *et al.* 2001, Li *et al.* 2005] permet de déduire un potentiel d'interaction d'une densité de probabilité. En ensemble canonique, la probabilité d'un état du système moléculaire est proportionnelle au facteur de Boltzmann :

$$P \propto \exp(-\frac{U}{k_B T}) \qquad \text{Équation I-31}$$

Où P est la densité probabilité et U l'énergie potentielle du système. Si, inversement, on connaît la probabilité, il suffit d'inverser le facteur pour obtenir le potentiel à partir de la distribution statistique :

Donc, pour les interactions intramoléculaires, on obtient :

$$U(r) = -k_B T \cdot \ln[P(r)/4\pi r^2] \qquad \text{Équation I-32}$$

$$U(\theta) = -k_B T \cdot \ln[P(\theta)/\sin\theta] \qquad \text{Équation I-33}$$

$$U(\varphi) = -k_B T \cdot \ln[P(\varphi)] \qquad \text{Équation I-34}$$

Où U est le potentiel intramoléculaire, r est la longueur de la liaison, θ est l'angle entre deux liaisons successives et φ est l'angle dièdre entre trois liaisons successives.

Pour les interactions intermoléculaires de symétrie sphérique :

$$U(r) = -k_B T \cdot \ln[P(r)/4\pi r^2] \qquad \text{Équation I-35}$$

Où r est la distance entre deux types de particules.

Dans le cas de la définition d'un potentiel gros grain à partir d'un modèle atomistique, la densité de probabilité est déduite de trajectoires de dynamique moléculaire atomistique. La calibration d'un modèle de potentiel gros grain consiste alors à ajuster les paramètres du modèle de potentiel pour que la distribution déduite du potentiel gros grain soit en cohérence avec la distribution provenant du modèle atomistique.

I.2.2.2 Force-Matching

Le Force-Matching [Ercolessi and Adams 1994] est une méthode directe de calcul des interactions CG à partir du modèle atomistique. L'objectif est de calculer une force effective exercée sur un grain à partir des forces exercées sur les atomes qui le constituent (Fig. I-11).

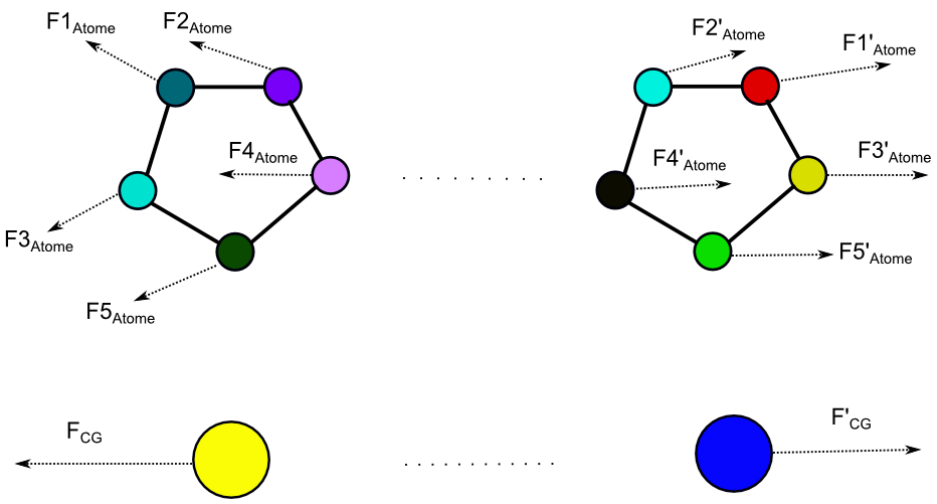

Figure I-11 Principe du Force-Matching. Les forces exercées à l'échelle gros grain (en bas) sont les sommes des forces exercées à l'échelle atomistique (en haut).

La force instantanée exercée sur un grain A par un grain B est la somme des forces exercées par les atomes du grain B sur les atomes du grain A :

$$\vec{F}_{B\to A} = \sum_{i\in A}\sum_{j'\in B} \vec{f}_{j'\to i} \qquad \text{Équation I-36}$$

Où *i* désigne un atome appartenant à A, *j'* désigne un atome appartenant à B, et $\vec{f}_{j'\to i}$ désigne la force exercée par l'atome *j'* sur l'atome *i*. Un lissage temporel de cette force brute peut être indispensable. On choisit ensuite un modèle de potentiel et on ajuste ses paramètres pour caler la force qui en dérive sur la force calculée à partir de la trajectoire atomistique.

I.3 Composition des cellules du bois

I.3.1 XYLOLOGIE DU BOIS

Le bois, nommé également xylème secondaire, est un tissu végétal à parois lignifiées issu du fonctionnement de l'assise génératrice liberoligneuse (cambium). Le bois est un matériau naturel rigide, résistant et isolant, obtenu à partir du tronc et des branches des arbres (Fig. I-12). Le bois assure trois fonctions principales pour l'arbre : la fonction de conduction de la sève brute des racines de l'arbre à la cime ; la fonction de soutien mécanique de la tige ; la fonction de stockage, de restitution, d'élaboration et de transformation chimique des substances.

Figure I-12 Photo à gauche : Forêt expérimentale de bois de l'INRA d'Eguisheim, Alsace, France. Photo à droite : Les troncs coupés des arbres. Photos pris le 08/10/2010 par l'auteur.

La formation du bois est conduite par le cambium qui produit chaque année une couche de bois, appelée cerne annuel, à partir de laquelle on peut déterminer l'âge cambial de l'arbre. Les utilisations du bois sont très diverses : énergie, construction, ameublement, fabrication d'outils et d'objets artisanaux, papier. Cette diversité d'utilisation s'explique en partie par la diversité de ses propriétés physiques et chimiques

Les propriétés du bois résultent de sa structure de matériau poreux qui lui confère un caractère anisotrope et hétérogène. Trois sections perpendiculaires entre elles se distinguent : la section transversale perpendiculaire au tronc, la section longitudinale radiale, parallèle aux rayons, et la section longitudinale tangentielle, perpendiculaire aux rayons (Fig. I-13). Les propriétés du bois sont différentes d'une direction à l'autre.

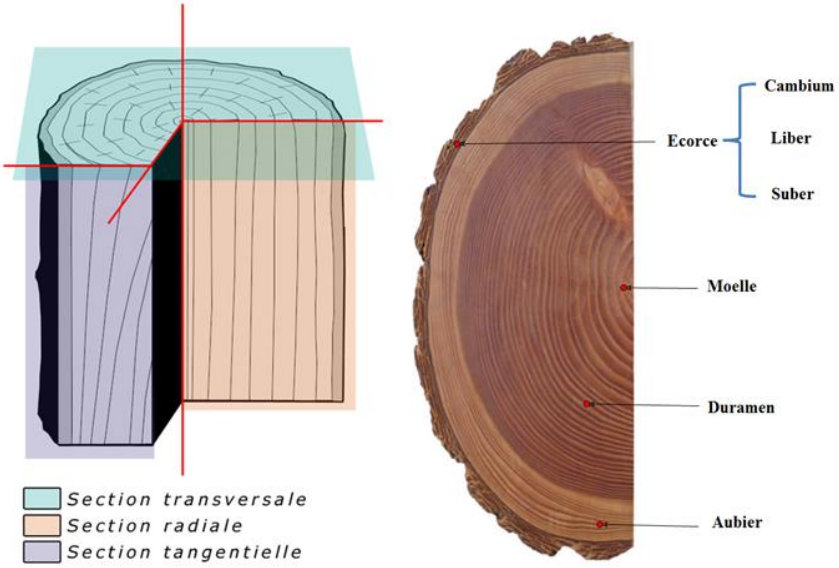

Figure I-13 Image à gauche : Anatomie du bois selon les sections transversales, radiale, tangentielle. Les trois sections sont respectivement marquées en couleur vert, violet et rose. Image à droite : La coupe transversale du tronc d'un chêne.

La moelle (ou cœur) est la partie la plus centrale de l'arbre, elle ne contient que le xylème primaire et joue un rôle important chez les jeunes pousses de plante. Elle disparait avec le temps pour ne laisser qu'un canal de faible section. La moelle est un ensemble de tissus spongieux qui évoluent lorsque la plante vieillit.

Le duramen (ou bois parfait) est la partie principale du tronc, il est formé par des cernes plus anciens. Le duramen n'est pas fonctionnel et ne contient que des cellules mortes et lignifiées. La moelle et le duramen ne sont pas indispensables pour un arbre vivant : chez certains vieux arbres encore vivant, les troncs sont souvent vides suite à l'attaque de nombreux agents de dégradation.

L'aubier (ou bois imparfait) est formé de couches concentriques de cellules récentes (cernes) non encore lignifiées. Ces couches, dans lesquelles circulent les matières nutritives, se duraminisent après une période allant de 4 à 20 ans.

L'écorce est la partie la plus externe du tronc, il revêt l'arbre et constitue une couche protectrice imprégnée d'une substance imperméable. L'écorce contient trois parties : le liber, le phelloderme et le suber. Le liber est constitué d'un ensemble de vaisseaux qui permet le

retour de la sève élaborée. Il forme la partie interne vivante de l'écorce. Le suber et le phelloderme sont mis en place par le phellogène (assise génératrice subéro-phellodermique) et constituent l'écorce externe.

I.3.2 STRUCTURE DE LA CELLULE DU BOIS

De nombreuses propriétés physiques et mécaniques du bois peuvent être expliquées à partir de la structure fine des parois cellulaire ligneuse. Celle-ci comprend trois couches: la couche intercellulaire, la paroi primaire, la paroi secondaire (Fig. I-14).

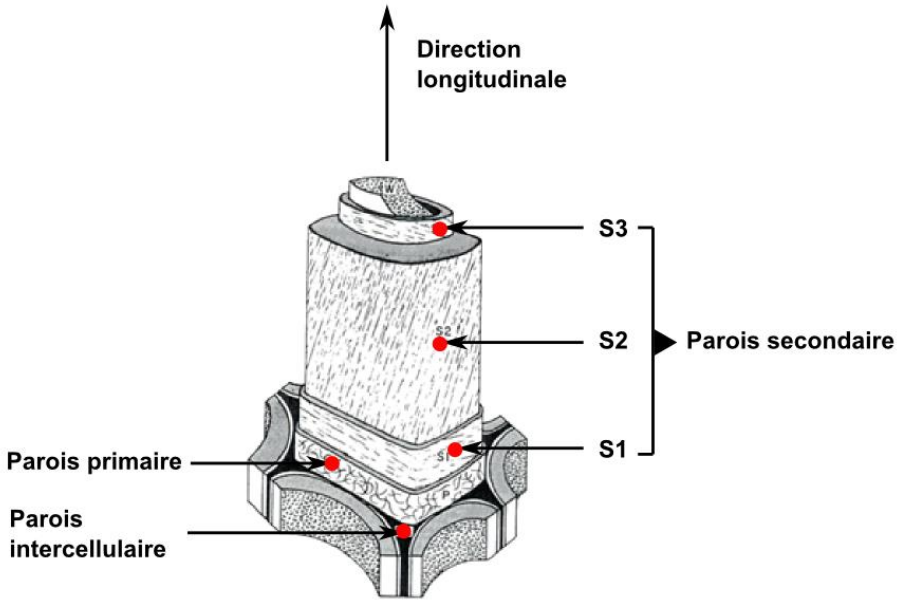

Figure I-14 Représentation de la paroi d'une cellule de bois.

La couche intercellulaire est la couche la plus externe de la cellule. C'est une fine membrane de séparation entre les cellules formée après la division du noyau des cellules. La couche intercellulaire contient des substances pectiques qui peuvent fixer le calcium et se remplit de lignine pendant la lignification. Sa forme est adaptée à l'élongation de cellule qui a lieu notamment pendant la différenciation cellulaire. Son épaisseur varie entre 0.5 et 1.5 μm.

Les parois primaire et secondaire contiennent une armature formée de microfibrilles de cellulose inclus dans une matrice d' hémicelluloses et de lignine déposée à la fin de la différenciation cellulaire.

La paroi primaire de la cellule s'appuie sur la couche intercellulaire. Elle apparaît dès la fin de la différenciation cellulaire. C'est une paroi très fine et élastique, d'une épaisseur d'environ 0.1 μm [Keller 1991]. Elle est composée de plusieurs couches de microfibrilles de cellulose enchevêtrées dans lesquelles se dépose de la lignine, des substances pectiques et des hémicelluloses au cours de la différenciation cellulaire.

La paroi secondaire, d'une épaisseur de 1~10 *µm* [Gindl *et al.* 2004], n'est pas aussi extensible que la couche intercellulaire, elle est produite et déposée sur la paroi primaire par la cellule de bois lorsque la différenciation est terminée. C'est une couche relativement épaisse et rigide car elle contient une forte densité de cellulose. Mécaniquement, elle est la plus résistante de la cellule. La paroi secondaire comporte trois couches : S1, S2 et S3.

Dans la couche S1, les microfibrilles de cellulose sont disposées en hélices d'orientations variables et alternées. L'angle d'orientation des microfibrilles varie de 60º à 80º [Core *et al.* 1979] par rapport à l'axe de la cellule. Cette variation d'angle permet de former un réseau croisé de celluloses dans cette couche qui assure une forte résistance mécanique à la traction transversale.

La couche S2, qui est la plus volumineuse parmi les trois couches, impose son comportement à l'ensemble de la paroi. Cette couche est formée par l'organisation dense de lamelles de microfibrilles de cellulose présentant un angle par rapport à l'axe de la cellule de 5º à 30º [Core *et al.* 1979].

La couche S3 est en fait une paroi tertiaire très similaire à la couche S1 bien que son épaisseur soit plus mince. Elle est constituée par des lamelles de celluloses moins parallèles que la couche S2. L'orientation générale des microfibrilles par rapport à l'axe de la cellulose constitue un angle variant de 60º à 90º [Core *et al.* 1979].

I.3.3 CONSTITUANTS CHIMIQUES DE LA CELLULE DU BOIS

Le bois se compose élémentairement de 50% de carbone, 42% d'oxygène, 6% d'hydrogène, 1% d'azote et 1% d'éléments divers. Les constituants structuraux du bois sont la cellulose, les hémicelluloses et la lignine. Les propriétés physiques du bois sont assurées par ces trois composants chimiques principaux. Il existe en plus de ces trois constituants principaux des substances mineures extractibles de nature organique, ainsi que des cendres.

I.3.3.1 La cellulose

Le bois contient 40% à 60% de cellulose selon l'essence. La cellulose est la molécule organique de polysaccharide la plus abondante sur terre. Elle est la composante principale des microfibrilles qui peuvent être considérées comme des cordes fabriquées de fils de très longues molécules de cellulose. C'est un polymère linéaire d'unité de D-glucose cyclique de deux configurations natives $I\alpha$ et $I\beta$. La structure de la cellulose $I\alpha$ est *P1* triclinique avec uniquement une cellobiose dans l'unité répétitive de la chaîne dont les paramètres sont : $a = 0{,}6717$ *nm*, $b = 0{,}59962$ *nm*, $c = 1{,}040$ *nm*, $\alpha = 118{,}08^o$, $\beta = 114{,}80^o$, $\gamma = 80{,}37^o$. La structure de la cellulose $I\beta$ est $P2_1$ monoclinique avec deux cellobioses dans l'unité répétitive de la chaîne dont les paramètres sont : $a = 0{,}7784$ *nm*, $b = 0{,}8201$ *nm*, $c = 1{,}0380$ *nm*, $\alpha = \beta = 90^o$, $\gamma = 96{,}5^o$ Voir (Fig. II-4).

La cellulose est construite par la répétition de l'unité de dimère cellobiose (Fig. I-15) qui est formée par deux monomères de glucose cyclique β, dont l'une est retournée de 180º par rapport à l'autre. La longueur et la largeur de la cellulose dans la nature sont respectivement 100 nm et 20 nm en moyenne, mais sur certaines espèces, la longueur peut atteindre 1000 nm et la largeur 20 nm [Elazzouzi-Hafraoui *et al.* 2007].

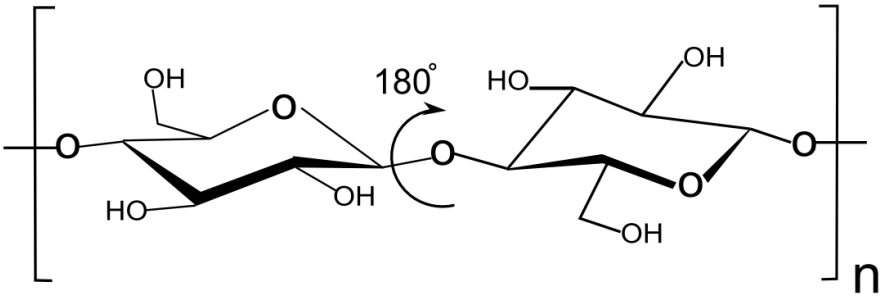

Figure I-15 La structure chimique de cellobiose et la rotation de dimère autour de la liaison d'oxygène β1-4.

Les macromolécules de cellulose associées forment des microfibrilles aux états cristalline et amorphe (Fig. I-16), qui sont elles-mêmes associées en couches pour former les parois des fibres végétales. Le monomère de la cellulose est le cellobiose. Dans le bois, le degré de polymérisation « DP » est compris entre 5000 et 10000. La structure de la cellulose permet de former des macromolécules grâce aux liaisons hydrogène qui s'établissent entre les groupements hydroxyles de la cellulose et d'autres molécules. Le taux de cristallisation de la cellulose dans le bois est entre 60% et 70% [Trouy-Triboulot and Triboulot 2012]. La figure I-16 expose la coexistence de cellulose cristalline et amorphe dans les parois végétales.

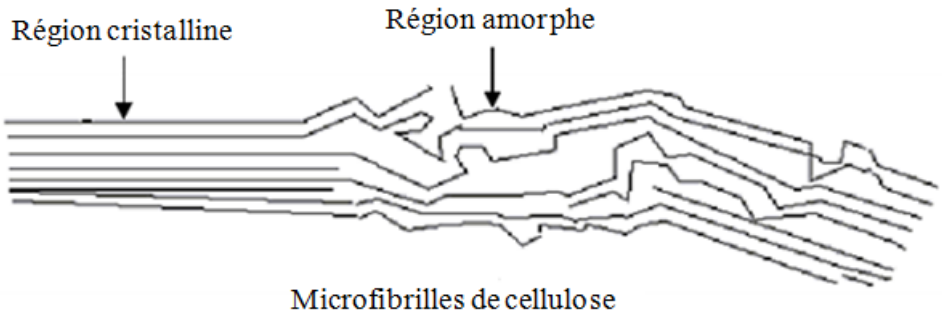

Figure I-16 Régions cristalline et amorphe de cellulose [Satge 2002].

I.3.3.2 Les hémicelluloses

Comme la cellulose les hémicelluloses sont des polysaccharides. La conformation des chaînes moléculaires des hémicelluloses a un caractère amorphe, aléatoire et présente peu de rigidité. Les hémicelluloses sont composés de différents types de monomères de sucres. Les hémicelluloses possèdent des chaînes de longueurs plus courtes que celles de la cellulose, ainsi que des chaînes ramifiés. Il existe deux types de molécules dans la catégorie des hémicelluloses : les pentosanes et les hexosanes. La distinction entre ces deux types de chaînes moléculaires est liée au type du monomère élémentaire considéré (Fig.I-17). Les hexosanes, représentées principalement par des chaînes des glucomannanes, sont construites à partir de monomères à 6 carbones de glucose ou de mannose. Les pentosanes sont constituées de monomères à 5 carbones de xylose, ce qui permet de former les chaînes de xylane.

Autrement dit, les hémicelluloses présentées dans le bois sont principalement des xylanes et des glucomannanes. La composition des hémicelluloses est différente pour des bois feuillus et pour des bois résineux. Les hémicelluloses des bois feuillus sont majoritairement composées de xylane et les hémicelluloses des bois résineux sont composées principalement de glucomannanes.

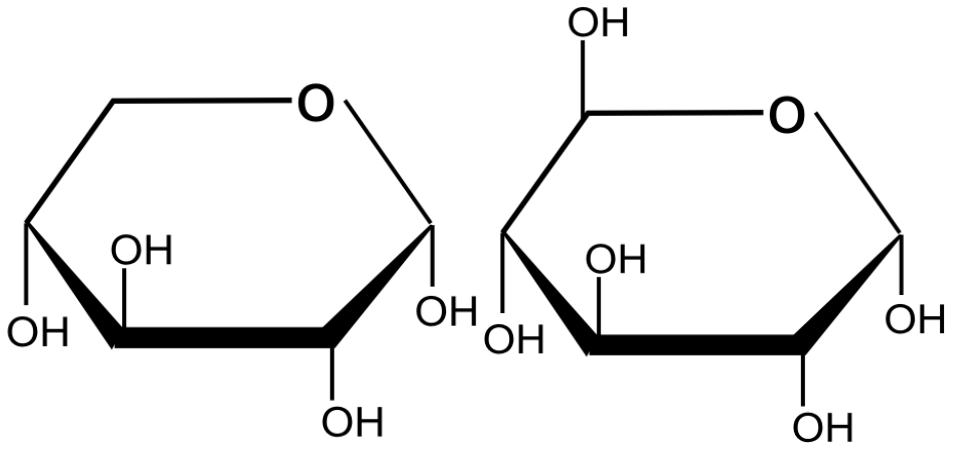

Figure I-17 La figure à gauche : la structure chimique du monomère de xylane. La figure à droite : la structure chimique du monomère de xyloglucane.

I.3.3.3 La lignine

Avec la cellulose et les hémicelluloses, la lignine est l'un des principaux composants chimiques du bois. Elle est le deuxième biopolymère renouvelable le plus abondant sur la Terre juste après la cellulose. La lignine est chimiquement très différente de la cellulose et des hémicelluloses. La fonctionnalité principale de la lignine est d'apporter à la paroi cellulaire la rigidité mécanique nécessaire lui conférant son imperméabilité à l'eau et sa résistance à décomposition. La lignine ne peut être définie avec précision car sa grande variabilité et la formation de lignine dépend fortement de l'environnement physico-chimique. La lignine est bio synthétisée dans le bois à partir de trois alcools phénoliques : les alcools paracoumarylique, coniférylique et sinapylique (Fig.I-18).

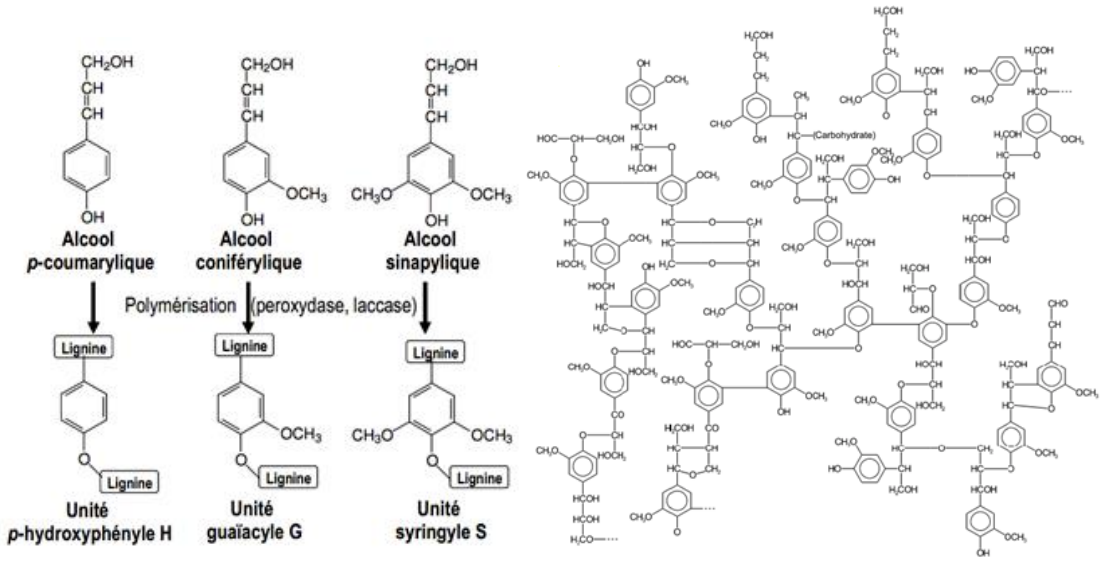

Figure I-18 Figure à gauche : représentation chimique des alcools p-coumarylique, coniférylique et sinapylique. Figure à droite : structure possible d'une lignine [Lapierre 2010].

La proportion de ces trois types de monomère varie de manière importante en fonction du végétal, de l'espèce, de l'organe et du tissu. Il n'est donc pas possible de définir précisément la structure chimique de la lignine, car presque toutes les lignines présentent des variations chimiques dues à l'association des monolignols, dérivés indirects d'un acide aminé servant de source pour la biosynthèse de lignines, par différentes liaisons chimiques sans caractère ordonné ni répétitif pour former après réduction, méthylation et déshydrogénation, un polymère amorphe et hydrophobe.

I.3.3.4 Les extractibles et les composants minéraux

Les extractibles comprennent de nombreux composants pouvant être extraits du bois par des solvants divers. Généralement, les bois tropicaux contiennent plus d'extractibles que le bois venant des zones tempérées. Les éléments extraits sont divers : des tannins, des terpènes, des lignanes, des stilbènes, des flavonoïdes et d'autres composants aromatiques. On retrouve également des lipides, des cires, et des acides gras [Trouy-Triboulot and Triboulot 2012].

Les composants minéraux sont normalement analysés à partir des cendres après combustion totale du bois à 600-800 ºC. Les principaux composés minéraux sont le potassium, le calcium, le magnésium et la silice.

I.3.3.5 Modèles de la structure des parois végétales

Avec les composants chimiques présentés dans cette section I.3.3, quelques modèles expérimentales de la paroi végétale peuvent être citées. Dans le modèle de [Terashima *et al.* 2009], la nanostructure des assemblages dans la couche primaire et la couche secondaire de la paroi de cellules a été présentée. Le modèle général [Terashima *et al.* 1993] de la paroi plus lignifiée a été également présenté, ainsi que les modèles améliorés [Terashima *et al.* 1995, Terashima *et al.* 1996, Terashima *et al.* 1996]. Ces modèles de la sous-structure, existent dans les couches primaire et secondaire de cellules, constituent importantes références pour la construction de modèles numériques du réseau moléculaire de la paroi.

I.4 Modélisation des assemblages de polymères des parois végétales

I.4.1 MODELES ATOMISTIQUES

L'application des approches atomistiques à la cellulose est une thématique déjà bien établie, et dont on peut ici donner quelques exemples de résultats. Celles-ci ont permis de prédire des structures stables de la cellulose cristalline [Viëtor *et al.* 2000] et de cristaux de cellulose en présence d'eau [Matthews *et al.* 2006, Yui and Hayashi 2007]. Les méthodes de mécanique moléculaire ont permis de prédire les propriétés mécaniques de modèles de cristaux de cellulose [Eichhorn *et al.* 2005, Tanaka and Iwata 2006]. La température de transition vitreuse d'un modèle de cellulose amorphe a été prédite [Mazeau and Heux 2003] et les micro-hétérogénéités structurales du réseau de liaisons hydrogène de la cellulose cristalline Iβ [Mazeau 2005] ainsi que la sensibilité de ce réseau à la température [Bergenstråhle *et al.* 2007] ont fait l'objet de simulations de dynamique moléculaire.

La lignine a fait l'objet d'investigations numériques moins nombreuses que la cellulose, probablement du fait de l'arrivée tardive de champs de force dédiés [Petridis and Smith 2009]. On peut citer l'exploration conformationnelle d'oligomères [Besombes *et al.* 2003, Besombes and Mazeau 2004] et d'agrégats [Petridis *et al.* 2011].

Enfin, on peut citer quelques études de la conformation des hémicelluloses isolées [Levy *et al.* 1991, Mazeau *et al.* 2005], appariées [Miyoshi *et al.* 2006], en solution [Almond and Sheehan 2003, Umemura and Yuguchi 2005], ou en phase amorphe [Mazeau *et al.* 2005].

En revanche, si la modélisation et la simulation numérique des assemblages de polymères pariétaux n'est pas un sujet entièrement neuf [Houtman and Atalla 1995], celle-ci ne s'est réellement développée que récemment, notamment dans le cas des assemblages cellulose-lignine [Besombes and Mazeau 2005[a], Besombes and Mazeau 2005[b]]. Le seul cas des assemblages cellulose-hémicellulose permet de souligner l'importance de cette voie d'exploration.

L'interaction entre les hémicelluloses et la cellulose est déterminante pour les propriétés, notamment mécaniques, des parois végétales [Somerville 2004]. Malgré son importance, l'interaction hémicelluloses-cellulose, reste toujours mal comprise [Roberts and Awad El-Karim 1983, Schönberg *et al.* 2001]. L'adsorption d'hémicellulose sur les microfibrilles de cellulose a été expérimentalement observée en solution [Kabel *et al.* 2007, Köhnke *et al.* 2010, Köhnke *et al.* 2011, Tokoh *et al.* 2002, Winter *et al.* 2005]. L'adsorption de xylane sur des microfibrilles de cellulose a été observée par AFM [Linder *et al.* 2003, Henriksson and Gateholm 2001] et par microscopie électronique [Mora *et al.* 1986]. La possibilité de former des agrégats de xylane sur la surface de monocristaux de cellulose est désormais bien connue. En revanche, la conformation du xylane adsorbé à la surface de la cellulose fait encore l'objet de débats. D'une part, certaines constatations suggèrent que les hémicelluloses sont parallèles à l'axe de la microfibrille de cellulose dans la paroi végétale [Stevanic and salmén 2009, Olsson *et al.* 2011]. D'autre part, d'autres mesures suggèrent que le xylane, dans la paroi végétale, interagit peu avec la cellulose [Åkerholm and Salmén 2001, Salmén and Olsson 1998] Afin de lever les ambiguïtés issues des investigations expérimentales, les outils de simulation numérique à l'échelle atomistique ont été mis à contribution [Zhang *et al.* 2011, Hanus and Mazeau 2006, Mazeau and Charlier 2012]. Des conformations parallèles, antiparallèles ou obliques du xylane ont ainsi été prédites.

Une limite majeure de ces approches réside dans leur poids numérique. En effet, d'une part, elles ne permettent pas de produire des trajectoires impliquant de nombreuses molécules d'hémicellulose [Charlier and Mazeau 2012], et d'autre part, le degré de polymérisation de la cellulose y est limité : sa valeur est de 14 dans les simulations de Zhang *et al*. [Zhang *et al.* 2011], alors que sa valeur réelle dépasse 40. C'est pourquoi le développement de modèles CG des assemblages de polymères pariétaux est indispensable.

I.4.2 MODELES GROS GRAIN

La modélisation des polymères pariétaux et de leurs assemblages est à priori un terrain d'application privilégié des approches CG, du fait de la taille des systèmes impliqués. La cellulose a fait l'objet de développements de modèles CG. La diversité des choix de la conversion au CG et de définition des potentiels d'interaction explique la diversité des articles publiés sur ce seul sujet malgré son caractère relativement récent.

On peut citer la modélisation CG de la cellulose amorphe introduite par Queyroy *et al.* [Queyroy *et al.* 2004]. Dans ce modèle, chaque grain est positionné sur l'atome d'oxygène de liaison β-O-4 qui relie deux monomères successifs. Un type de grain spécifique est défini pour la description des moitiés de monomères en bouts de chaîne. Un défaut de ce type de modèle CG est que l'information relative à l'orientation des monomères est perdue, ou doit être reconstruite à posteriori. Lorsqu'il s'agit d'accélérer des phases de relaxation lente, de procéder à un back-mapping, et de poursuivre la simulation à l'aide d'un modèle atomistique [Queyroy *et al.* 2004], ce défaut n'est pas fortement préjudiciable. En revanche, lorsque le modèle CG est utilisé tel quel pour produire les trajectoires à analyser, ce n'est plus

nécessairement le cas, en fonction des applications visées. Une solution pour remédier à ce problème consiste à décrire les monomères par plusieurs grains, au prix d'une moindre réduction du nombre de degrés de liberté par rapport aux modèles atomistiques. Le modèle proposé par Bu *et al*. [Bu *et al.* 2009] entre dans cette catégorie : un monomère de cellulose est décrit par trois grains. Cette approche a notamment permis de prédire une torsion spontanée d'un monocristal de cellulose. On peut également citer les simulations de Bellesia *et al.* [Bellesia *et al.* 2012] par dynamique de Langevin, mettant en œuvre un modèle CG de cellulose comportant 5 grains par monomère.

Il existe une alternative à la mise au point spécifique d'un champ de force CG dédié à la cellulose : il s'agit d'utiliser un champ de force CG un peu plus générique, en l'occurrence le champ de forces MARTINI [Marrink *et al.* 2004, Marrink *et al.* 2007] récemment étendu aux carbohydrates [López *et al.* 2009]. Cette approche a récemment donné lieu à plusieurs publications [Wu *et al.* 2012, Wohlert and Berglund 2011].

Un défaut des modèles CG de la cellulose est qu'ils ne permettent pas de former un monocristal de cellulose stable. Il est alors indispensable d'imposer un potentiel de confinement pour empêcher la dissociation des monocristaux CG de se dissocier, ce qui introduit un biais dans les simulations. A ce jour, il n'existe qu'un modèle CG de cellulose cristalline [Srinivas *et al.* 2011], dans lequel un monomère est décrit par un seul grain, le champ de force étant calibré à partir de distributions de probabilité issues de calculs atomistiques. Ce modèle permet de générer des cristaux de cellulose numériquement stables jusqu'à $1\mu s$ de simulation.

Peu de simulations gros grain impliquant l'interaction des hémicelluloses avec la cellulose ont été réalisées. On peut citer l'approche Monte Carlo de Morris *et al.* [Morris *et al.* 2004], basée sur des potentiels d'interaction génériques, ainsi que deux modèles de paroi végétale très simplifiés mais de grande échelle, s'appuyant sur des potentiels partiellement arbitraires [Manyuhina *et al.* 2007] ou calibrés sur des mesures de coefficients d'élasticité [Adler and Buehler 2013]. C'est dans cet espace thématique peu exploré que se situe la présente thèse.

II. MODÈLE GROS GRAIN

Résume du chapitre II

Dans ce chapitre, un modèle gros grain du xylane, en interaction avec la cellulose cristalline, est présenté.

Dans un premier temps, les matériels utilisés pour la mise au point de ce modèle gros grain sont présentés. En particulier, les modèles et trajectoires atomistiques, fournis par Karim Mazeau, du CERMAV (Grenoble), et desquels sont déduits les potentiels gros grain, sont détaillés.

Ensuite, la stratégie de passage de l'échelle atomistique à l'échelle gros grain est présentée.

Les potentiels gros grain d'interaction intramoléculaires du xylane, déduits par inversion de Boltzmann, sont ensuite exposés et validés.

Suit la présentation d'une démarche d'optimisation du potentiel d'interaction intermoléculaire xylane-xylane, initialisée par un paramétrage initial obtenu par Force-Matching.

Enfin, suivant une approche similaire, la déduction des potentiels d'interaction intermoléculaire xylane-cellulose cristalline est décrite.

II.1 Matériels

II.1.1 MODELES ATOMISTIQUES

Nous présentons dans cette section les modèles atomistiques utilisés pour générer les trajectoires desquelles ont été déduits les potentiels gros grain dans le cadre de la thèse. Elles sont fournies par Karim Mazeau du CERMAV de l'Université Joseph Fourrier Grenoble 1.

II.1.1.1 Trajectoires atomistiques des chaînes de Xylane

Le champ de force utilisé pour calculer la conformation des xylanes est Molecular Mechanic 3 ou MM3 [Allinger *et al.* 1989, Allinger *et al.* 1990]. Ce champ de force a été utilisé pour le calcul de la conformation élémentaire des dimères de xylose et les hélices des chaînes de xylanes.

Le modèle atomistique de xylane est composé par l'unité répétitive de dimère D-xylopyranose avec la liaison *1-4β*. (Fig. II-1) Seule la configuration $^4C_1$ a été considérée pour la construction des cycles. Chaque monomère a été construit à partir des paramètres standards issus de la cristallographie au niveau de la longueur de liaisons, angles interne des liaisons.

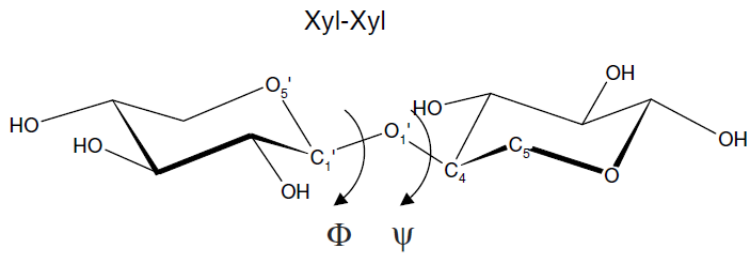

Figure II-1 Structure chimique du dimère de xylane avec la définition de rotations internes par les angles de torsion *Φ* et *Ψ*.

La forme géométrique d'un disaccharide est majoritairement dominée par la liaison glycosidique [French *et al.* 1993]. Afin de construire la configuration d'une chaîne de xylane, il est nécessaire d'établir la conformation autour de la liaison glycosidique entre deux monomères. L'orientation relative de deux unités saccharides peut être exprimée à l'aide de deux paramètres portant respectivement sur les deux angles de torsion autour de la liaison glycosidique : $\Phi = O_5'\text{-}C_1'\text{-}O_1'\text{-}C_4$ et $\Psi = C_5\text{-}C_4\text{-}O_1'\text{-}C_1'$.

Les hélices de la chaîne de xylane ont été générées par le programme POLYS, abréviation de « Polysaccharide » [Engelsen *et al.* 1996]. POLYS nécessite en entrée les géométries de différentes unités de monosaccharide et les angles de torsion Φ et Ψ. Le monomère provient de l'étude préliminaire du disaccharide et a été implanté dans le MONOBANK [Pérez and Delage 1991].

La génération des conformations de chaînes aléatoires contenant un grand nombre de monomères a été effectuée en utilisant le programme METROPOL [Boutherin *et al.* 1997].

La chaine de xylane est formée à partir d'un disaccharide initial, en ajoutant un monomère suivant à l'aide de valeurs Φ et Ψ aléatoires. Le nouveau résidu, décrit par Φ et Ψ, sera accepté si l'énergie de la nouvelle conformation est inférieure ou égale à celle de la conformation précédente. Sinon, un test de « Metropolis » intervient : on génère tout d'abord un nombre aléatoire entre 0 et 1, ensuite, on compare ce nombre avec le facteur de Boltzmann $e^{-\Delta E/kBT}$, où ΔE est l'écart d'énergie, k_B est la constante de Boltzmann et T est la température. Si le nombre généré est inférieur au facteur, la nouvelle conformation sera acceptée, sinon, Φ et Ψ se voient affecté de nouvelles valeurs aléatoires, la génération de la chaine recommencera jusqu'à ce que la nouvelle conformation soit acceptée.

Ensuite, la structure générée fait l'objet d'une thermalisation par dynamique moléculaire. Suit la génération de quatre trajectoires dans l'ensemble *NVT* sous la température 300 K. Les modèles atomistiques concernant le xylane [Mazeau *et al.* 2005] contiennent 3 extractions de trajectoire de 500 pas de temps chacune sous forme de fichier PDB « Protein Data Bank ». Chaque trajectoire ne comporte qu'une chaine de xylane. Les degrés de polymérisation (DP) dans ces trois trajectoires (Fig. II-2) sont respectivement 5, 10, 25. La représentation visuelle des modèles atomistiques sont illustrée dans la figure II-2.

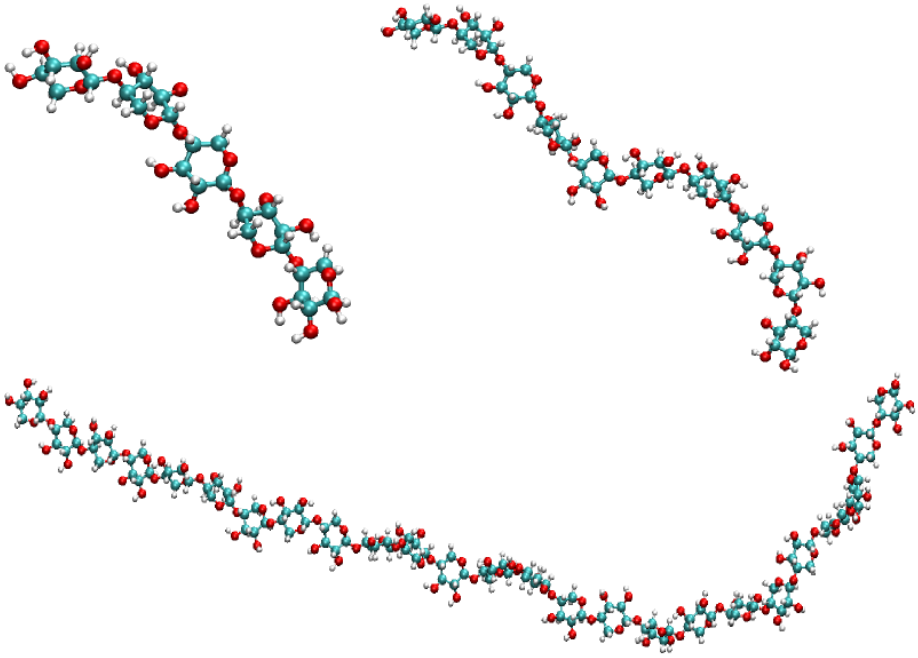

Figure II-2 Modèles atomistiques des chaînes moléculaires de xylane de quatre longueurs. La figure en haut à gauche DP = 5. La figure en haut à droite DP = 10. La figure en bas DP = 25.

II.1.1.2 Trajectoire atomistique Xylane-Xylane

Pour modéliser l'interaction intermoléculaire xylane-xylane, un extrait de trajectoire de 500 pas a été utilisé. Ce modèle atomistique est composé de deux chaines de xylane de 5 xyloses chaque. Les deux chaînes, ont été comprimées de façon à les faire interagir. La simulation atomistique a ensuite été exécutée sous la température de 300 K. Le modèle généré de l'interaction entre deux chaines moléculaire du xylane est présenté dans la figure II-3.

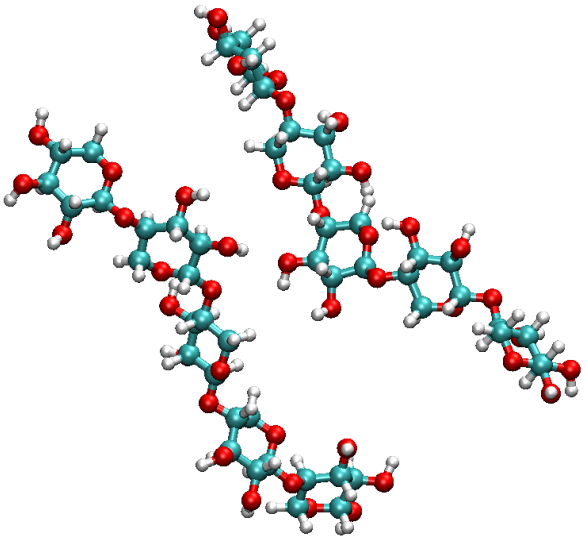

Figure II-3 Modèle atomistique de l'interaction de deux chaînes moléculaires de xylane de DP = 5.

II.1.1.3 Trajectoire atomistique Xylane-Cellulose

Pour modéliser l'interaction intermoléculaire xylane-cellulose, un extrait de trajectoire de 500 pas a été utilisé. Ce modèle atomistique est composé d'une chaine de xylane de 5 xyloses, et de 3 couches de cellulose de 8 chaines chacune, chaque chaine comportant 8 glucoses.

La surface (110) est l'une des deux surfaces dominantes, (110) et (1-10), dans les microfibrilles de cellulose [Mazeau *et al.* 2012, Van Daele *et al.* 1992, Hanley *et al.* 1992, Helbert *et al.* 1998[a], Helbert *et al.* 1998[b], Revol 1982, Sassi and Chanzy 1995, Sassi *et al.* 2000, Sugiyama *et al.* 1985]. De plus, il a été constaté par Résonnance magnétique nucléaire (RMN, ou NMR en anglais) que les chaînes du xylane sont préférentiellement en interaction avec la surface (110), beaucoup plus qu'avec la surface (1-10) [Bergenstråhle *et al.* 2008, Larsson *et al.* 2004].

Dans le cadre de la thèse, seule une trajectoire atomistique de l'absorption d'une chaîne de xylane de DP = 5 sur la surface (110) de la cellulose a été utilisée pour la modélisation de potentiel gros grain de xylane-cellulose, les deux autres surfaces (1-10) et (100) sont également prises en compte, les intensités des potentiels d'interaction xylane-cellulose étant ajustées à l'aide du ratio entre l'énergie d'adsorption du xylane sur ces surfaces et cette même énergie dans le cas de la surface (110), ratio obtenu, dans chaque cas, par dynamique moléculaire atomistique.

Le cristal de cellulose a été construit à partir des paramètres de la structure Iβ cristal, illustré dans la figure II-4, déterminé par diffraction des Rayon X et diffusion de Neutrons [Nishiyama *et al.* 2002], dont les paramètres de cristal sont $a = 0{,}77848$ *nm*, $b = 0{,}82018$ *nm*, $c = 1{,}38010$ *nm*, $\alpha = \beta = 90^o$, $\gamma = 96{,}5^o$, les données correspondantes à la structure géométrique sont illustrées dans la figure II-4. La cellule a été dupliquée dans l'espace pour former 8 couches de celluloses de 8 chaines chacune.

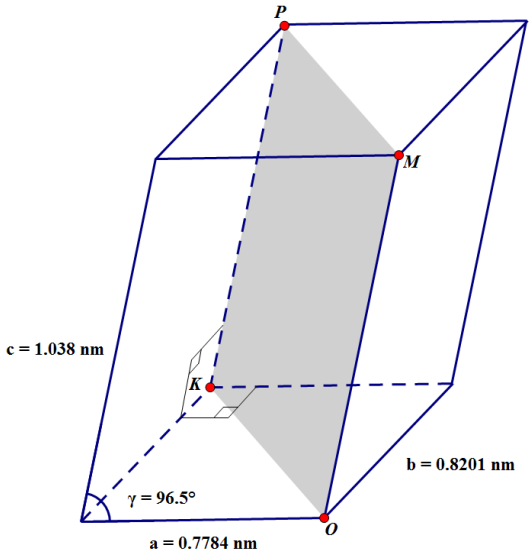

Figure II-4 La structure géométrique de la cellulose de configuration Iβ, données structurales obtenues par diffraction des rayons X et diffusion de Neutrons [Nishiyama *et al.* 2002].

Les 8 couches ont été équilibrées dans l'ensemble *NPT*, sous une température de 300K et une pression de 1 atm [Mazeau 2005]. Ensuite, le paramètre de la cellule de direction *b* a été manuellement augmenté afin de créer une membrane de cellulose exposant deux surfaces équivalentes et parallèles à (110). Le modèle atomistique final ne contient que 3 couches de cellulose qui sont déployées dans une boîte de simulation avec PBC dans toutes les directions après le retrait de 5 couches de cellulose. Ce modèle représente une surface accessible des

microfibrilles natives de cellulose. Seuls les groupes fonctionnels de la couche externe de la cellulose sont actifs. Les coordonnées du reste des atomes de la cellulose sont fixées.

Une chaîne de xylane de 5 monomères a été générée selon la description résumée dans la section II.1.1.1, elle est ensuite mise au-dessus de la surface (110) de la cellulose avec une orientation aléatoire. Une trajectoire de ce système xylane-cellulose a été ensuite générée [Mazeau and Charlier 2012] dans l'ensemble canonique *NVT*. La représentation visuelle du modèle xylane-cellulose est présentée dans la figure II-5.

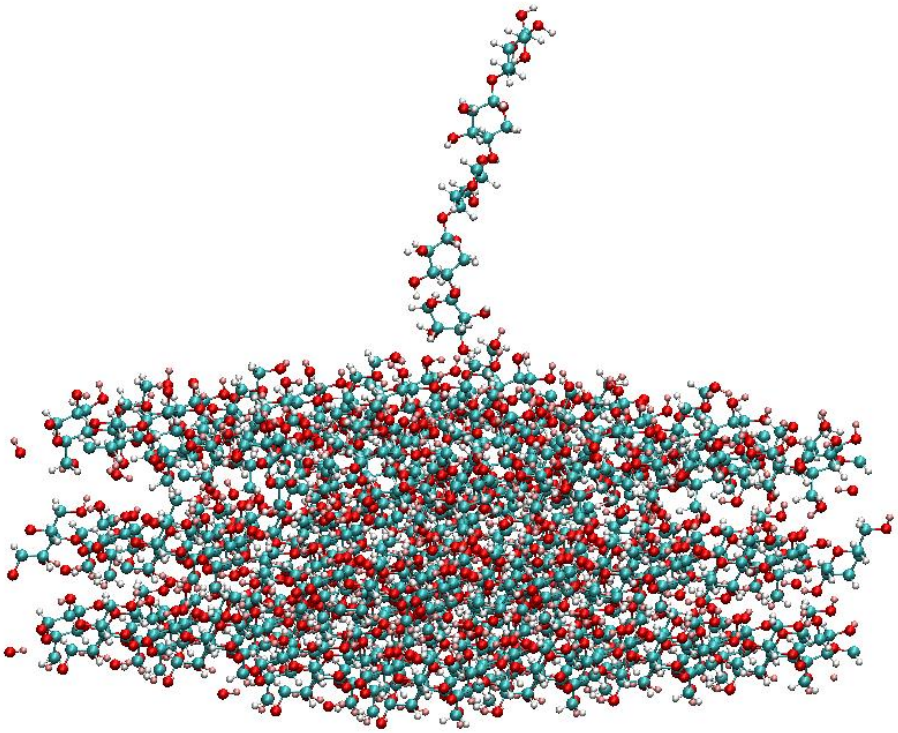

Figure II-5 Modèle atomistique de l'interaction d'une chaîne de xylane de DP=5 avec trois couches de cellulose chacune constituée de 8 chaînes de 8 monomères chacune. La surface exposée est la surface 110 de la cellulose cristalline.

II.1.1.4 Champs de force atomistiques.

Le champ de forces UFF [Rappé *et al.* 1992], abréviation des « Universal Force Field », a été utilisé pour les simulations de dynamique moléculaire atomistiques. L'UFF a été utilisé auparavant [Hanus and Mazeau 2006, Henao and Mazeau 2008] pour la description des propriétés des polysaccharides complexes. La distribution des charges électrostatiques provient de la méthode d'équilibration des charges pour la dynamique moléculaire [Rappé and Goddard 1991].

Pour les interactions intermoléculaires, deux types de modèles du potentiel sont pris en compte : la force Van der Waals et la force électrostatique.

Pour les forces de Van der Waals, on utilise le potentiel Leonard-Jones 12-6. Les paramètres fournis sont donnés pour des paires d'atomes de même nature dans le tableau II-1 :

	σ_i (nm)	ε_i (kJ/mol)
H	**0,2571**	**0,046024**
O	**0.3118**	**0,062760**
C	**0.3431**	**0,010983**

Table II-1 Paramètres de « Universal Forces Fields » [Rappé *et al.* 1992].

Les paramètres pour des paires d'atomes de natures différentes s'en déduisent simplement :

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \qquad \text{Équation II-1}$$

$$\varepsilon_{ij} = (\varepsilon_i \cdot \varepsilon_j)^{1/2} \qquad \text{Équation II-2}$$

Où les indices *i* et *j* désignent ici des natures d'atomes.

Le calcul des interactions électrostatiques est rendu possible par la donnée des charges partielles portées par les atomes :

Les paramètres sont détaillés par la suite.

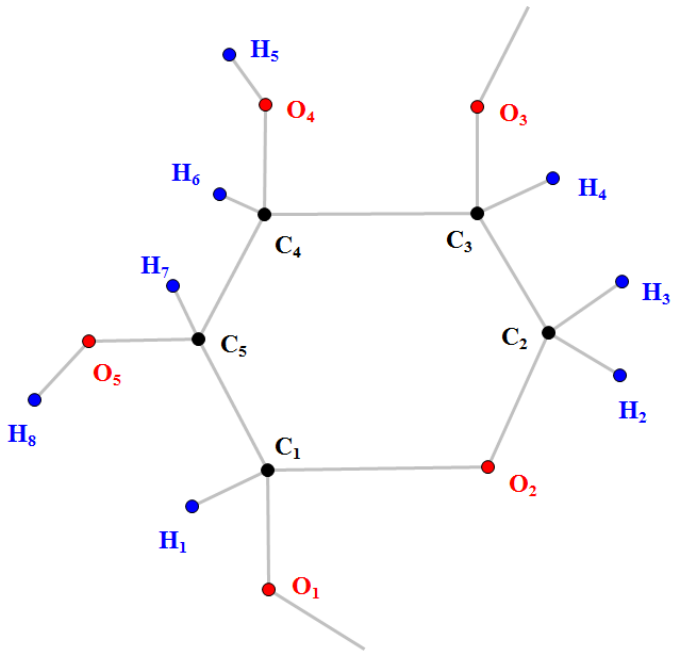

Figure II-6 Numérotation des atomes d'un monomère de xylose.

Atome	C_1	C_2	C_3	C_4	C_5	O_1	O_2	O_3	O_4
Charges (e)	**0.3715**	**-0.0582**	**0.0177**	**0.0874**	**0.1994**	**-0.6269**	**-0.6040**	**-0.6269**	**-0.7311**
Atome	**O_5**	**H_1**	**H_2**	**H_3**	**H_4**	**H_5**	**H_6**	**H_7**	**H_8**
Charges (e)	**-0.6860**	**0.1695**	**0.1625**	**0.1937**	**0.2264**	**0.4320**	**0.2064**	**0.2237**	**0.4127**

Table II-2 Charges électriques de la xylose, données fournies par le Karim Mazeau de CERMAV, Grenoble, France.

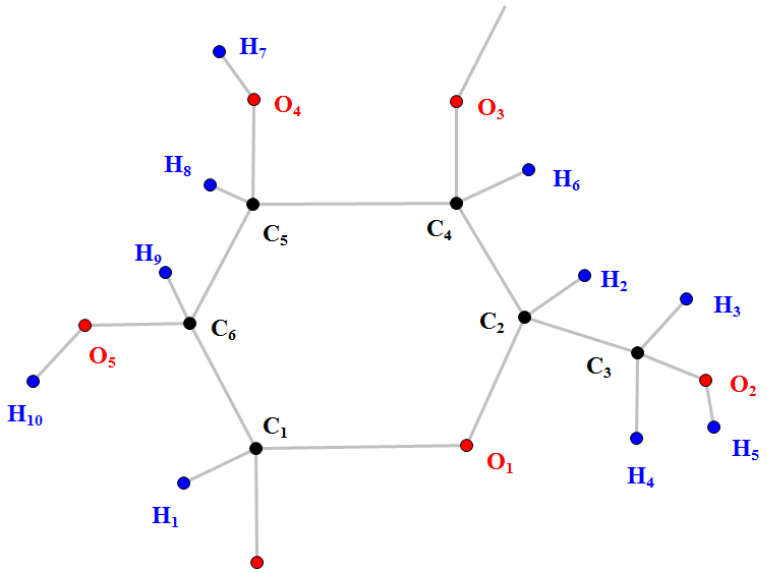

Figure II-7 Numérotation des atomes du monomère de cellulose.

Atome	C1	C2	C3	C4	C5	C6	O1
Charge (e)	**0.3197**	**0.1246**	**-0.0046**	**0.1013**	**0.1067**	**0.1432**	**-0.6870**
Atome	**O2**	**O3**	**O4**	**O5**	**H1**	**H2**	**H3**
Charge (e)	**-0.7156**	**-0.5749**	**-0.7693**	**-0.7318**	**0.2133**	**0.2569**	**0.1591**
Atome	**H4**	**H5**	**H6**	**H7**	**H8**	**H9**	**H10**
Charge (e)	**0.2099**	**0.4523**	**0.1721**	**0.4264**	**0.1973**	**0.2397**	**0.4199**

Table II-3 Charges électriques du monomère de cellulose, données fournies par le Karim Mazeau du CERMAV.

II.1.2 LOGICIEL DE DYNAMIQUE MOLECULAIRE

Un code de dynamique moléculaire gros grain a été développé dans le cadre de la thèse. Ses fonctionnalités sont listées ci-dessous:

- Générations aléatoires des structures moléculaires des hémicelluloses.

- Générations de microfibrilles.
- Générations des structures composites hémicellulose-cellulose. Un algorithme a été développé pour éviter les superpositions des particules pendant cette phase.
- Génération des trajectoires CG.

Ce code s'appuie sur l'intégrateur de Verlet-Vitesse [Verlet 1967] pour l'intégration des équations du mouvement. La thermodynamique du système moléculaire est contrôlée par le thermostat Nosé-Hoover Chaîne [Martyna *et al.* 1992]. Les thermostats Nosé-Hoover [Nosé 1984[a], Nosé 1984[b], Hoover 1985, Evans and Holian 1985], Berendsen [Lemak and Balabaev 1994], Andersen [Andersen 1980] sont également disponibles. L'approche « Link-Cell » [Grest *et al.* 1989] est utilisé pour accélérer la construction de la liste de Verlet et le calcul est parallélisé par OpenMP Les conditions aux limites périodiques sont également implémentées.

Le calcul des simulations à l'échelle CG présentées dans cette thèse a été réalisé par un ordinateur fixe.

II.1.3 Logiciel de visualisation VMD

Pour la visualisation des trajectoires, le logiciel de simulation MD génère pour chaque simulation un fichier en format PDB qui peut être lu par le logiciel « Visual Molecular Dynamics » ou VMD. Il s'agit d'un logiciel de visualisation de structure moléculaire. Il a été initialement développé en 1992 par Mike Krogh, Bill Humphrey et Rick Kufrin. La première version est sortie en 1995, et actuellement VMD est maintenu par la collaboration de « Beckman Institute » et le groupe « Theoretical et Computational Biophysics ».

II.2 Définition des grains

II.2.1 STRATEGIE DE CONVERSION CG

La stratégie de conversion de l'échelle atomistique à l'échelle gros grain consiste à remplacer un monomère par un grain situé au barycentre du monomère considéré. (Fig. II-8)

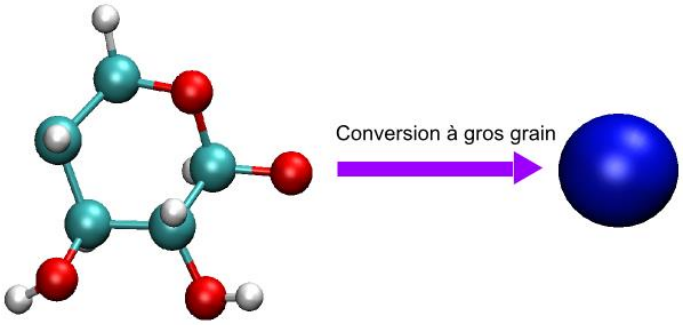

Figure II-8 Conversion d'un monomère de xylane de l'échelle atomistique à l'échelle CG.

Par exemple, pour une chaîne de xylane de DP = 5, sa version CG est une chaîne de 5 grains situant aux barycentres de monomères correspondants (Fig. II-9).

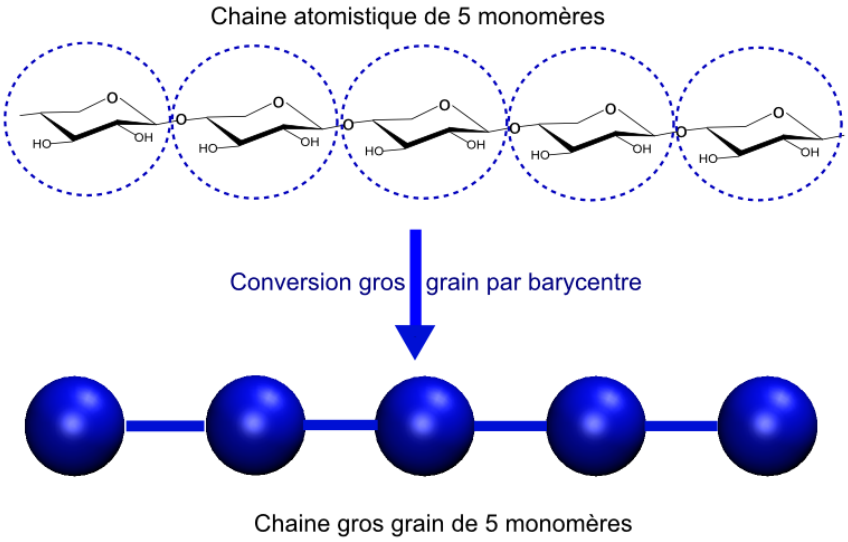

Figure II-9 Conversion d'une chaîne atomistique de xylane à l'échelle CG.

La coordonnée du barycentre du monomère peut être définie comme suit :

$$B(x,y,z) = \frac{\sum_{i=1}^{N} m_i b_i(x,y,z)}{\sum_{i=1}^{N} m_i} \qquad \text{Équation II-3}$$

Où N est le nombre total d'atomes dans un monomère, m_i est la masse molaire du $i^{\text{ème}}$ atome dont b_i est la coordonnée. Quel que soit le type de monomère, nous appliquons toujours ce type de conversion.

II.2.2 Gain et Compromis

Le changement d'échelle atomistique à l'échelle gros grain simplifie considérablement les molécules. En outre, le nombre total de particules, et donc de degrés de liberté du système moléculaire en entier s'en trouvent fortement réduit.

Dans notre cas, un monomère de xylane est composé de 17 atomes et un monomère de cellulose est composé de 21 atomes. Après conversion au gros grain, pour un système composé de xylane, le nombre de degrés de liberté est 17 fois plus faible que celui du modèle atomistique ; de même pour le système gros grain de cellulose, le nombre de degrés de liberté est 21 fois plus faible que dans le cas du modèle atomistique.

Même si le modèle gros grain permet de réduire les temps de calcul et de modéliser ainsi les systèmes moléculaires de grande taille, les informations détaillées des modèles atomistiques sont perdues. En effet, la procédure de conversion ne conserve que l'information relative au barycentre de monomère et la perte de l'information atomistique ne permet pas d'accéder à tous les détails internes d'un monomère. Dans le cas du modèle CG choisi, on perd également une information sur l'orientation de chaque monomère, puisque les

monomères sont assimilés à des points matériels. Néanmoins, on fera l'hypothèse que ces détails atomistiques ont une influence très limitée sur les propriétés que nous déduirons des calculs au niveau macroscopique. On peut, par ailleurs, nuancer la perte d'information subie en considérant que les effets d'un certain nombre de détails à petite échelle, par exemple l'orientation des monomères ou les liaisons hydrogène, etc., sont pris en compte de manière « effective » à travers le paramétrage du champ de force CG. Il sera donc essentiel de valider le champ de force finalement obtenu en s'assurant qu'il reproduit bien le comportement prédit par le modèle atomistique à l'échelle qui nous intéresse, en l'occurrence, celle du polymère. Il n'est donc pas contradictoire de sacrifier les détails atomistiques pour assurer l'accessibilité des propriétés macroscopiques d'un grand système moléculaire.

II.2.3 MODELES GROS GRAIN

Nous illustrons dans cette section la représentation des modèles gros grain qui ont été dérivés des modèles atomistiques pour les trajectoires :

- Da la dynamique d'une chaine du xylane de différentes longueurs ;
- De la dynamique des deux chaines de xylane en interaction ;
- De la dynamique d'une chaine de xylane interagissant avec trois couches de cellulose.

Les trajectoires CG sont obtenues par la stratégie de conversion présentée dans la section II.1.1.

II.2.3.1 Xylane

La conversion des chaines de xylane à l'échelle CG ne pose pas de difficulté (Fig. II-10). Cette figure est la version CG des trajectoires présentes dans la figure II-2.

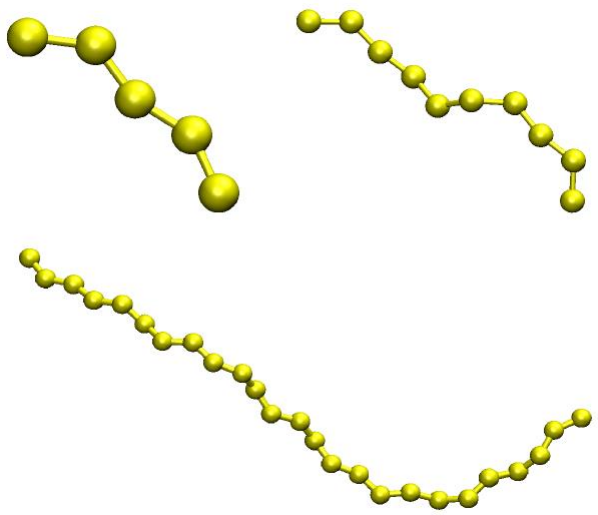

Figure II-10 Modèles CG des chaînes moléculaires du xylane de différents DP. Figure en haut à gauche DP = 5. Figure en haut à droite : DP = 10. Figure en bas : DP = 25.

II.2.3.2 Xylane-xylane

La conversion du système constitué d'une paire d'oligomères de xylane de DP = 5 à l'échelle CG ne pose pas de difficulté (Fig. II-11).

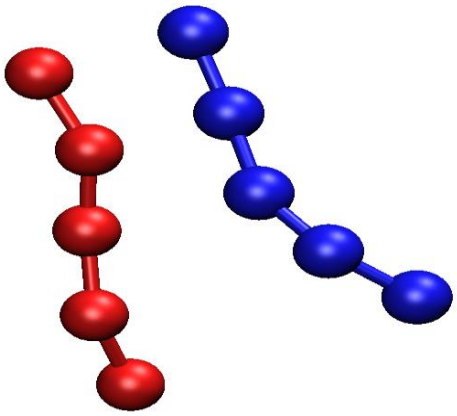

Figure II-11 Modèle CG de l'interaction de deux chaînes moléculaires du xylane de DP = 5.

II.2.3.3 Xylane-cellulose

La conversion gros grain du système cellulose-xylane nécessite quelques précautions supplémentaires. Le problème réside dans la perte d'intégrité des monomères situés aux deux bouts d'une chaîne de cellulose. Autrement dit, à travers les conditions aux limites périodiques, la boîte de simulation découpe en deux parties les monomères des bouts de chaîne. Ce découpage ne permet pas de convertir facilement les monomères entiers au bout de la chaîne lors de la conversion CG. Pour remédier à ce problème, la partie découpée d'un monomère est translatée à l'autre bout de la chaîne afin de former un monomère intact (Fig. II-12, Fig. II-13).

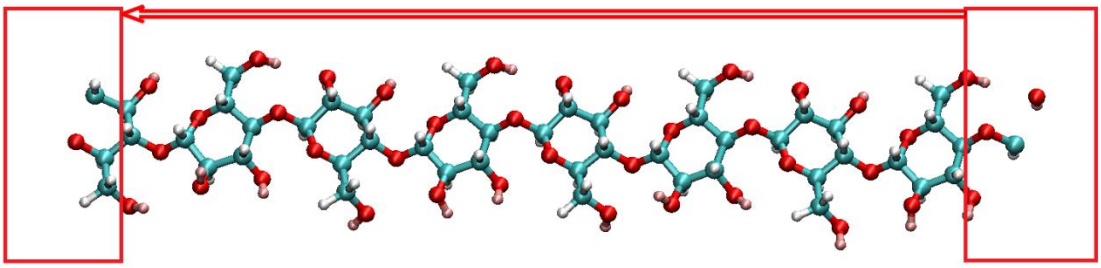

Figure II-12 Translations d'une partie de monomère atomistique d'un bout de la chaîne à l'autre bout pour la reconstruction d'un monomère intact.

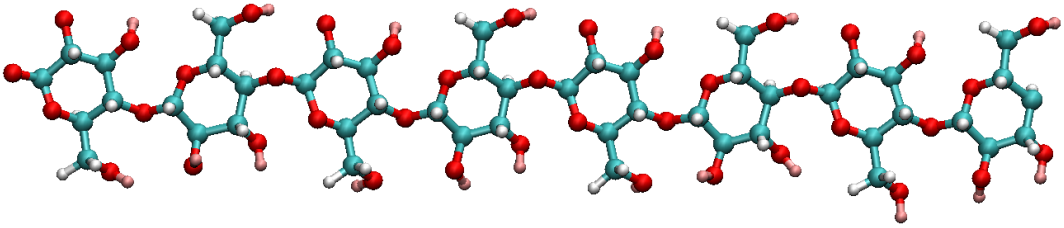

Figure II-13 Une chaîne intacte de cellulose après la reconstruction de monomères coupés par translation.

Après la reconstruction des monomères pour toutes les chaînes dans le système moléculaire xylane-cellulose, la trajectoire est prête à la conversion en gros grain (Fig. II-14).

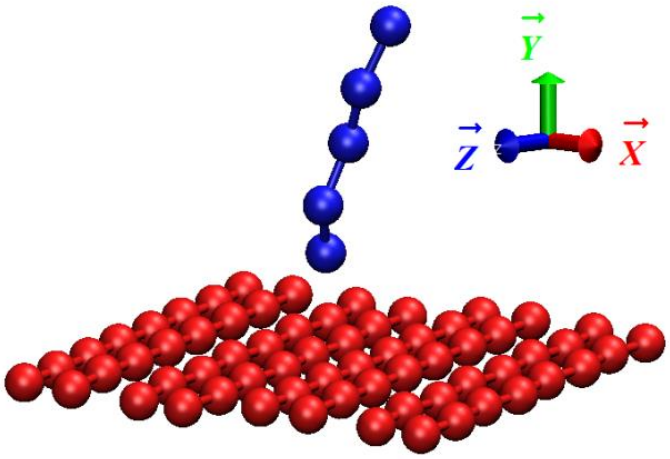

Figure II-14 Conversion CG du modèle de système cellulose-xylane.

Les chaînes de cœur sont par ailleurs éliminées, dans la mesure où elles sont inutiles pour la déduction du potentiel CG intermoléculaire cellulose-xylane.

II.2.4 FIXATION DE LA CELLULOSE

Dans la trajectoire atomistique du système xylane-cellulose, la couche active est la surface 110, qui interagit avec le xylane, les autres chaînes de cellulose étant figées. Par conséquent, le déplacement des barycentres des monomères de la surface 110 de la cellulose est très limité par rapport à leurs positions moyennes. Il est donc possible de poser l'hypothèse que la surface active peut-être définie comme figée à l'échelle gros grain. Aucune interaction cellulose-cellulose ne sera donc explicitement considérée, et la cellulose cristalline interviendra sous la forme d'un assemblage de centres d'interactions extérieurement imposés au xylane. Pour la justification d'une telle hypothèse, nous avons défini un critère, exprimé dans la formule suivante :

$$Norm_{dis} = \frac{D_i}{D_s} \qquad \text{Équation II-4}$$

Où D_i est le déplacement moyen du $i^{ème}$ grain autour sa position moyenne, D_s est une des périodes géométriques du modèle xylane-cellulose, définie à 0,519 nm : la distance moyenne entre deux chaines voisines.

Cette norme définit la grandeur du déplacement de grains par rapport à l'échelle de description du système moléculaire. Dans la figure II-15, chaque cercle représente un grain de

la surface *110* de la cellulose, les couleurs sont liées à la valeur de la norme, l'échelle de référence étant donnée à droite de la figure. La valeur maximale de la norme est 0,024, elle est assez petite pour pouvoir négliger la mobilité des grains appartenant à la surface *110* de la cellulose par rapport à celle des grains de xylane.

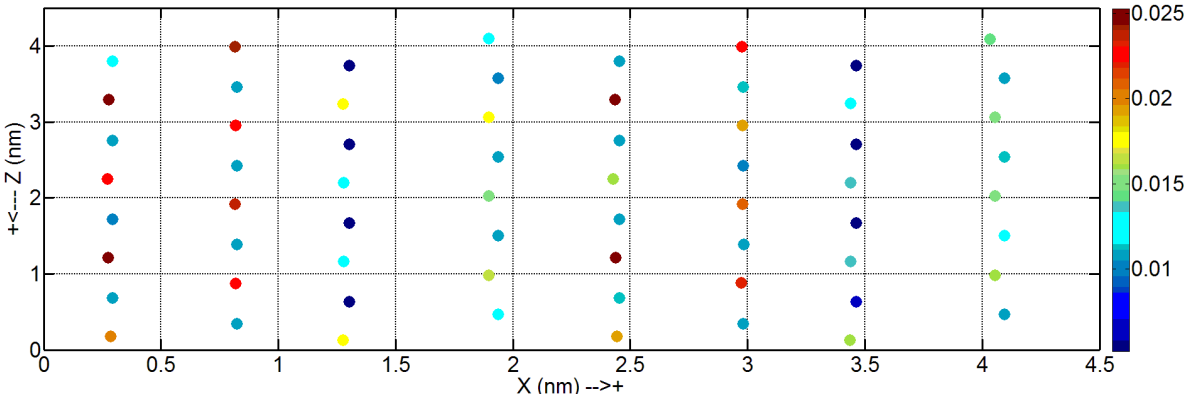

Figure II-15 Norme de déplacement du barycentre de chaque monomère de la surface *110* de la cellulose.

Par ailleurs, on va considérer que l'interaction du xylane avec les couches profondes de la cellulose est intégrée dans le potentiel modélisé de l'interaction entre la surface *110* et le xylane.

II.3 Potentiels intramoléculaires du xylane

II.3.1 PROCEDURES DE MODELISATION

Après le changement d'échelle d'atomistique au gros grain, une étape importante est la dérivation des potentiels gros grain intramoléculaires des trajectoires atomistiques. Les étapes à mettre en œuvre sont les suivantes :

1. Développement de l'algorithme de calcul :
   - ✓ de la longueur de liaison entre deux grains voisins,
   - ✓ de l'angle de trois grains successifs,
   - ✓ de l'angle dièdre parmi quatre grains successifs ;
2. Application de l'algorithme développé précédemment sur la trajectoire gros grain issue de la trajectoire atomistique afin d'obtenir les distributions brutes des liaisons, des angles, et des angles dièdres.
3. Traitement des types de potentiels par l'inversion de Boltzmann en choisissant pour chaque terme d'interaction un modèle de potentiel empirique.
4. Pour chaque type de potentiel, on ajuste les paramètres du modèle de potentiel de façon à ce que la distribution de probabilité prédite par le potentiel coïncide avec la distribution de probabilité issue de la conversion CG des trajectoires atomistiques.

Ces procédures de la modélisation de potentiels s'appuient sur les trajectoires du xylane de 5, 10 et 25 xyloses. Les modèles de potentiels seront par la suite soumis à une validation.

II.3.2 MODELES DU XYLANE DE 5, 10 ET 25 XYLOSES

Idéalement, dans une chaîne moléculaire sans branches dont les grains alimentent les échantillons pour la modélisation de potentiels, chaque grain participe à neuf forces dont deux

forces de liaisons, trois forces d'angle et quatre forces d'angle dièdre. Pour que cela soit possible, un grain doit avoir six grains voisins dont trois de chaque côté afin de s'assurer qu'il est entièrement couvert par les neuf forces. Ce critère « *Neuf Forces* » assure que les grains échantillonnés ne subissent pas trop d'effet de bord. La figure II-16 illustre ce critère.

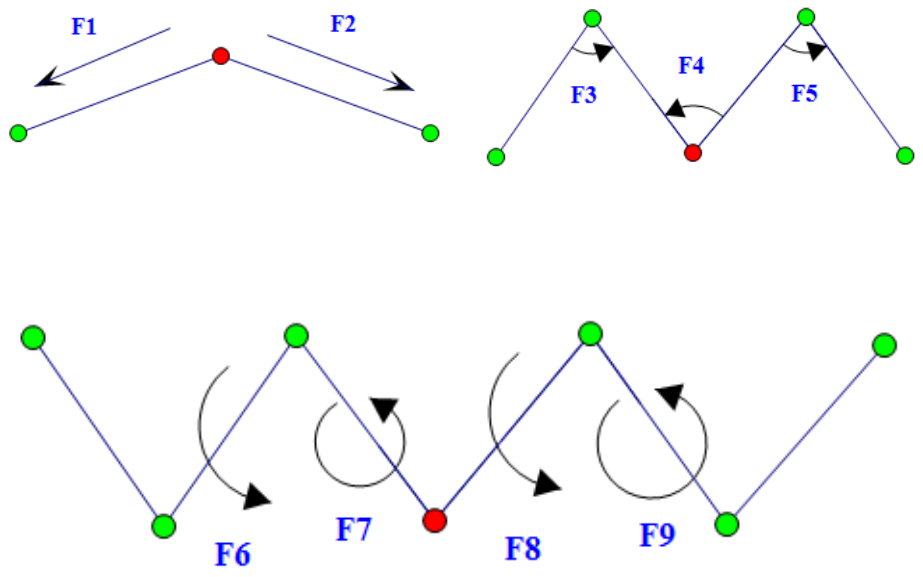

Figure II-16 Critère « neuf forces ». Figure en haut à gauche : les deux forces d'élongation appliquées sur le grain rouge. Figure en haut à droite : les trois forces de flexion appliquées sur le grain rouge. Figure en bas : les quatre forces de torsion appliquées sur le grain rouge.

Les grains non couverts par ces neuf forces ne seront pas utilisés pour alimenter les échantillons de la modélisation de potentiels en raison des effets de bord qu'ils subissent. On se fixe donc deux conditions : la chaîne doit être suffisamment longue et les échantillons statistiques ne proviennent que des grains du centre des chaines qui ont assez de voisins afin de s'assurer d'être couverts par les neuf forces.

Même si aucun grain de la trajectoire du xylane de 5 xyloses n'est qualifié par le critère de « *Neuf Forces* », nous présentons ses distributions statistiques pour illustrer l'existence des effets de bord.

Dans les figures qui suivent, les grains susceptibles d'alimenter l'échantillonnage seront entourés d'un carré rouge. Une exception pour le xylane de 5 xyloses : les trois grains entourés par le carré rouge ne forment aucun angle dièdre, par conséquent, l'échantillonnage de l'angle dièdre de cette chaîne a été appliqué sur la chaîne entière. Pour les trajectoires d'échantillon de DP = 10 et DP = 25, seuls les grains couverts par les « *Neuf Forces* » ont été pris en compte.

II.3.2.1 Modèle du xylane de 5 xyloses

La figure II-17 illustre l'extraction des distributions du modèle DP = 5, les paramètres extraits sont exposés dans le tableau II-1.

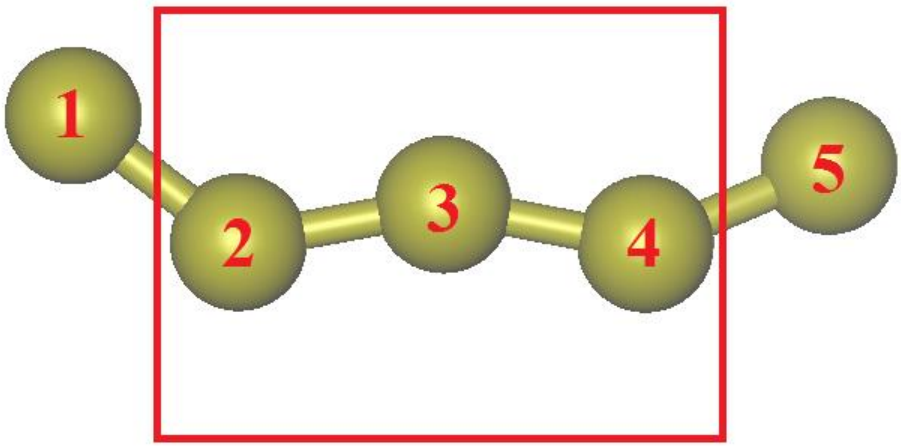

Figure II-17 Modèle CG de la chaîne moléculaire du xylane de DP = 5.

	Echantillons	Quantité d'échantillons
Liaisons	**[23], [34]**	**2*500**
Angles	**∠234**	**1*500**
Angles dièdres	**∠1234, ∠2345**	**2*500**

Tableau II-1 Numérotation des grains d'échantillonnage dans la trajectoire du xylane de DP = 5.

Les figures II-18, II-19 et II-20 montrent respectivement la distribution de la longueur de liaison, de l'angle et de l'angle dièdre. Il est facile de constater que dans ces trois distributions extraites de la trajectoire du xylane DP = 5, les distributions sont fortement bruitées. Toutes les perturbations présentes dans les données statistiques de DP = 5 sont dues, d'une part, à la faible taille des échantillons, et à l'effet de bord important du fait de la très courte longueur de la chaîne.

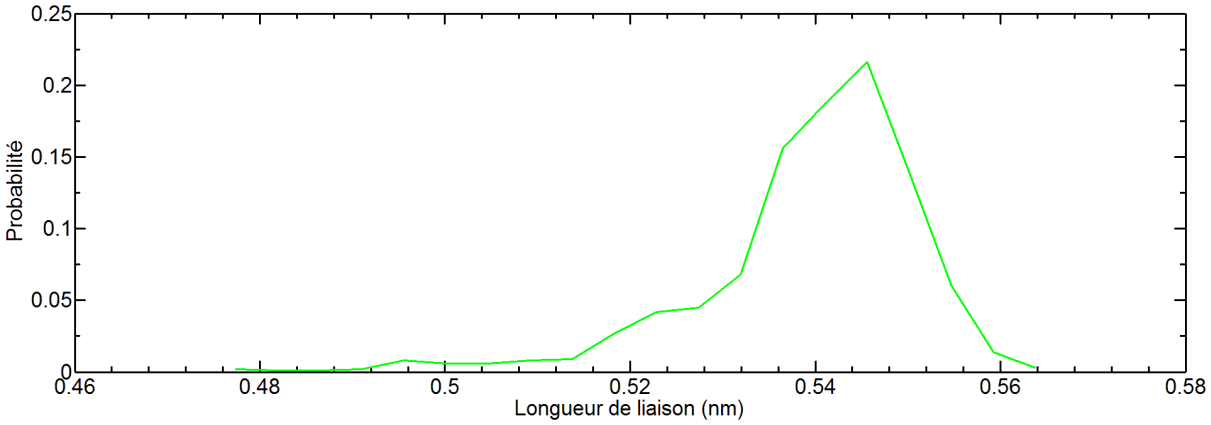

Figure II-18 Distribution de la longueur de liaison entre deux grains successifs le long de la chaîne moléculaire de xylane de DP = 5.

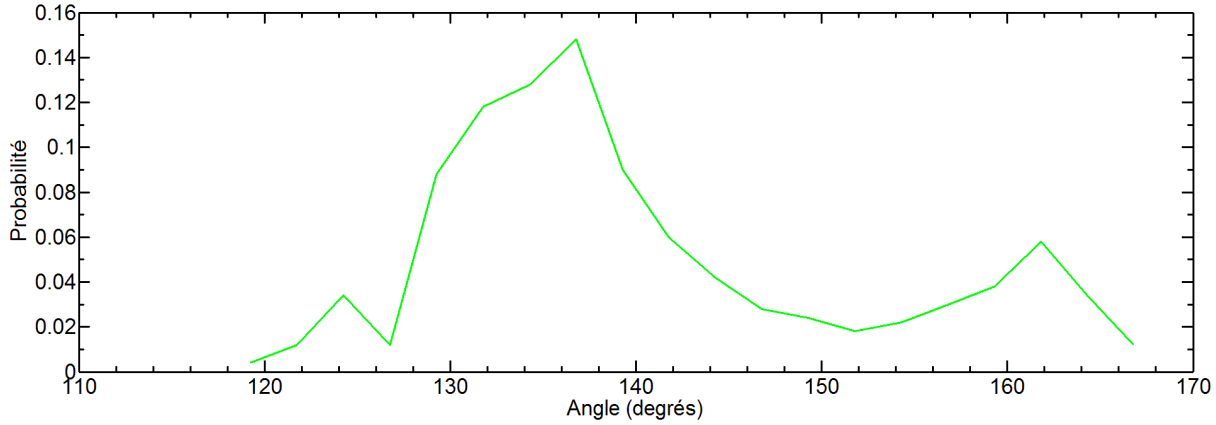

Figure II-19 Distribution de l'angle entre trois grains successifs le long de la chaîne moléculaire de xylane de DP = 5.

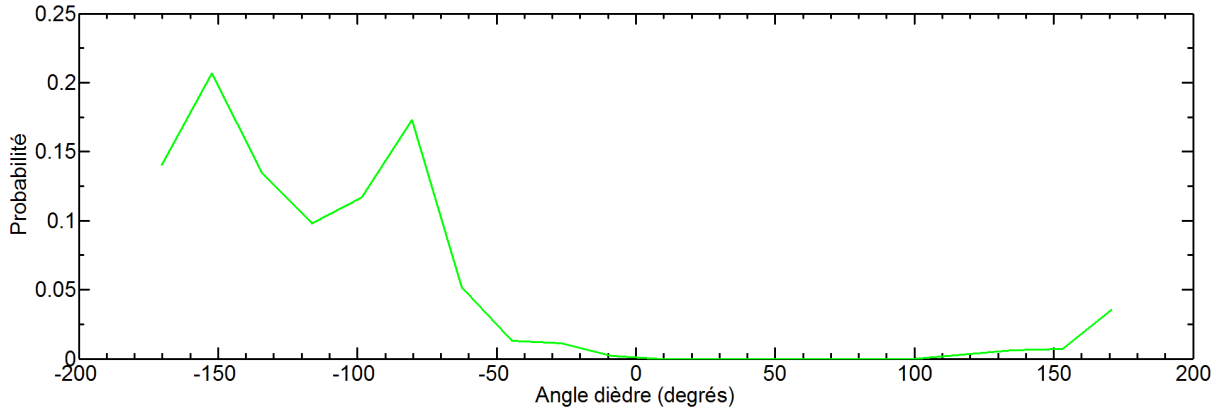

Figure II-20 Distribution de l'angle de torsion (angle dièdre) entre quatre grains successifs le long de la chaîne moléculaire de xylane de DP = 5.

II.3.2.2 Modèle du xylane de 10 xyloses

Nous exposons dans cette section l'échantillonnage du modèle du xylane de DP = 10 (Fig. II-21). Les paramètres extraits sont détaillés dans le tableau II-2.

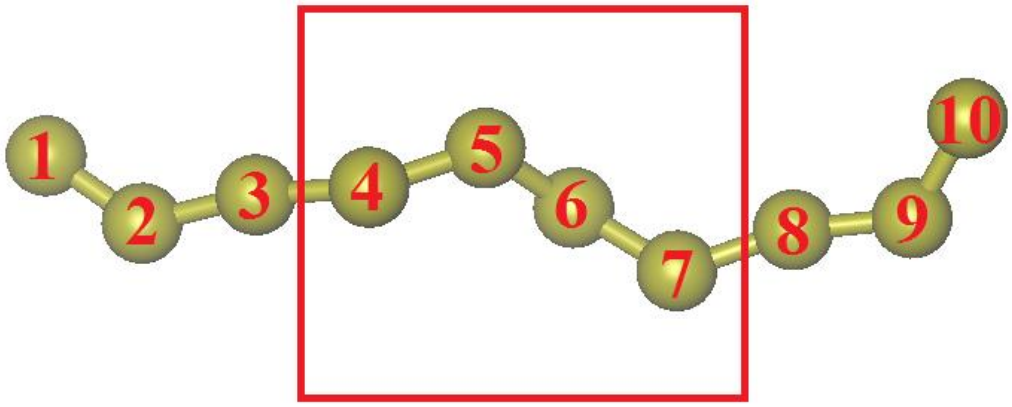

Figure II-21 Modèle CG de la chaîne moléculaire du xylane de DP = 10.

	Echantillons	Quantité d'échantillons
Liaisons	**[45], [56], [67]**	**3*500**
Angles	**∠456, ∠567**	**2*500**
Angles dièdres	**∠4567**	**1*500**

Tableau II-2 Numérotation des grains d'échantillonnage dans la trajectoire du xylane de DP = 10.

Vu les distributions extraites, qui sont présentées dans les figure II-22, II-23 et II-24, du modèle du xylane de DP = 10, la qualité des distributions est visiblement meilleure que celle de DP = 5, du fait de l'augmentation relative de la taille d'échantillon et de l'évitement de l'effet de bord.

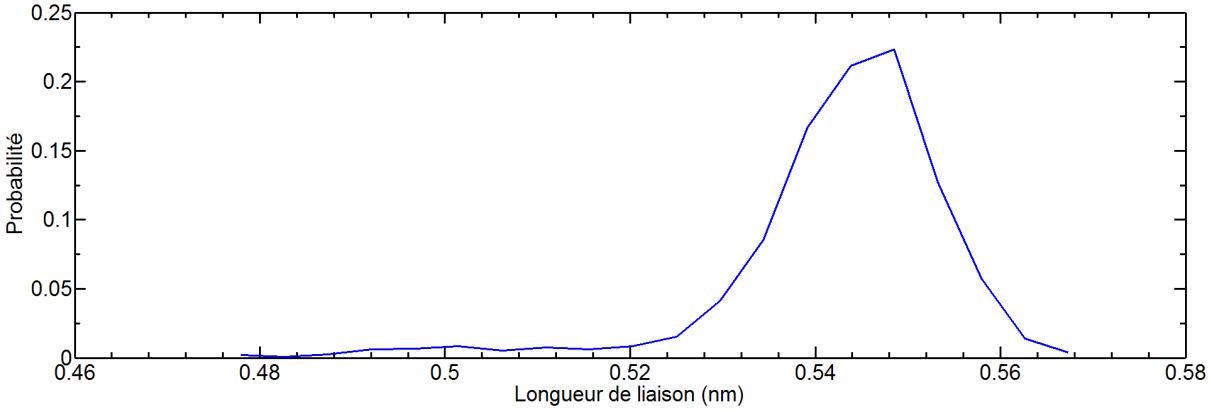

Figure II-22 Distribution de la longueur de liaison entre deux grains successifs le long de la chaîne moléculaire de xylane de DP = 10.

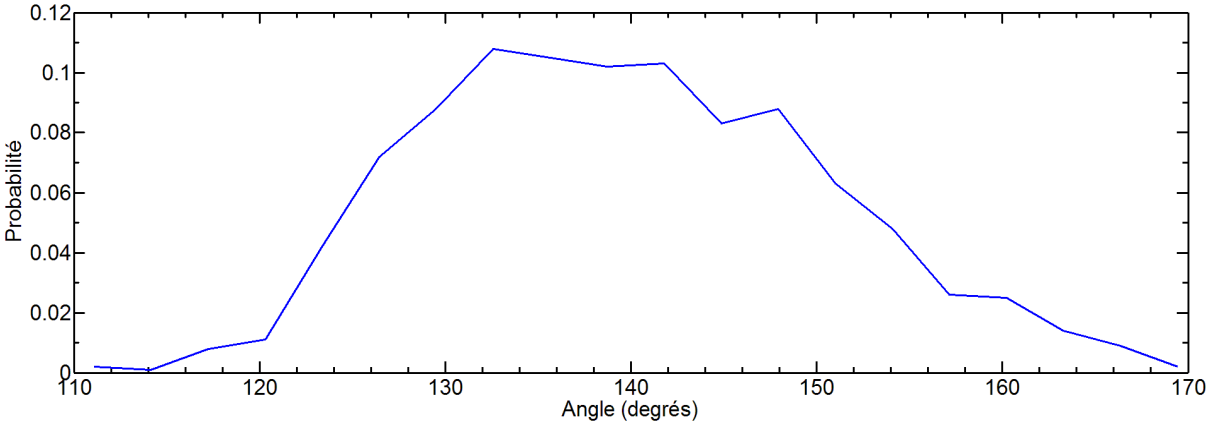

Figure II-23 Distribution de l'angle entre trois grains successifs le long de la chaîne moléculaire de xylane de DP = 10.

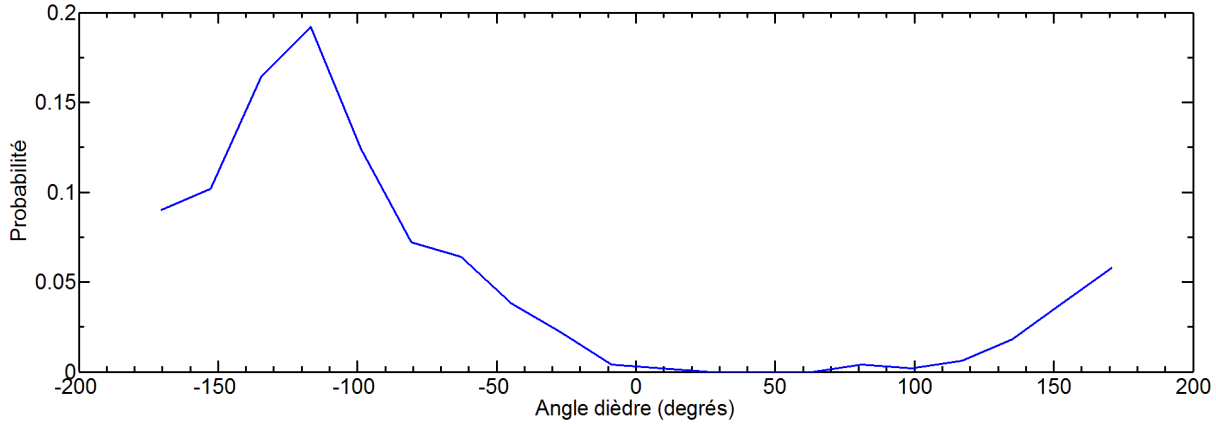

Figure II-24 Distribution de l'angle de torsion (angle dièdre) entre quatre grains successifs le long de la chaîne moléculaire de xylane de DP = 10.

II.3.2.3 Modèle du xylane de 25 xyloses

L'échantillonnage des termes intramoléculaires pour le modèle du xylane de DP 25 est illustré dans la figure II-25 ainsi les paramètres concernés sont listés dans le tableau II-3.

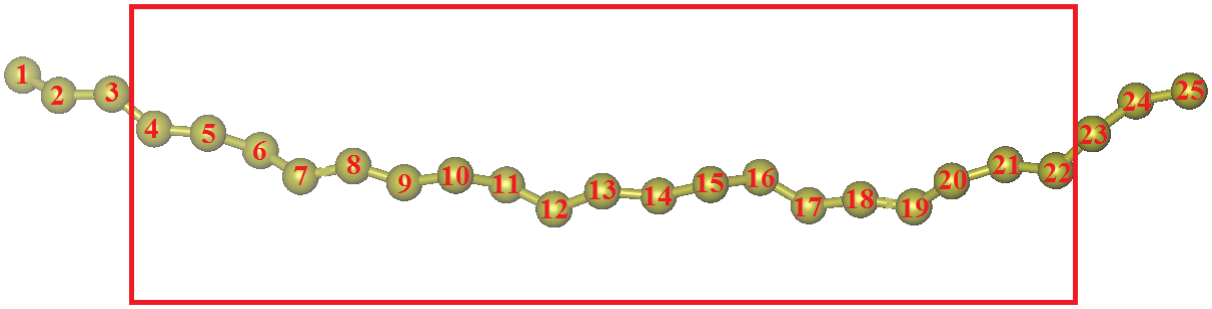

Figure II-25 Modèle CG de la chaîne moléculaire de xylane de DP = 25.

	Echantillons	Quantité d'échantillons
Liaisons	**[4,5], [5,6], [6,7], [7,8], [8,9], [9,10], [10,11], [11,12], [12,13], [13,14], [14,15], [15,16], [16,17], [17,18], [18,19], [19,20] , [20,21] , [21,22]**	**18*500**
Angles	**∠4.5.6, ∠5.6.7, ∠6.7.8, ∠7.8.9, ∠8.9.10, ∠9.10.11, ∠10.11.12, ∠11.12.13, ∠12.13.14, ∠13.14.15, ∠14.15.16, ∠15.16.17, ∠16.17.18, ∠17.18.19, ∠18.19.20, ∠19.20.21, ∠20.21.22**	**17*500**
Angles dièdres	**∠4.5.6.7, ∠5.6.7.8, ∠6.7.8.9, ∠7.8.9.10, ∠8.9.10.11, ∠9.10.11.12, ∠10.11.12.13, ∠11.12.13.14, ∠12.13.14.15, ∠13.14.15.16, ∠14.15.16.17, ∠15.16.17.18, ∠16.17.18.19, ∠17.18.19.20, ∠18.19.20.21, ∠19.20.21.22**	**16*500**

Tableau II-3 Numérotation des grains d'échantillonnage dans la trajectoire du xylane de DP = 25.

Il est aisé de constater que dans les figures II-26, II-27 et II-28, la trajectoire de DP = 25 fournit un assez grand nombre d'échantillons pour les trois termes d'interactions intramoléculaires : longueur de liaison, angle et angle dièdre. Les courbes de distributions ne présentent pas de bruits importants. Cette meilleure qualité provient évidement de la plus

grande taille de l'échantillon total. On peut également noter que les distributions sont assez proches de celles obtenues pour DP = 10. On a donc visiblement convergence.

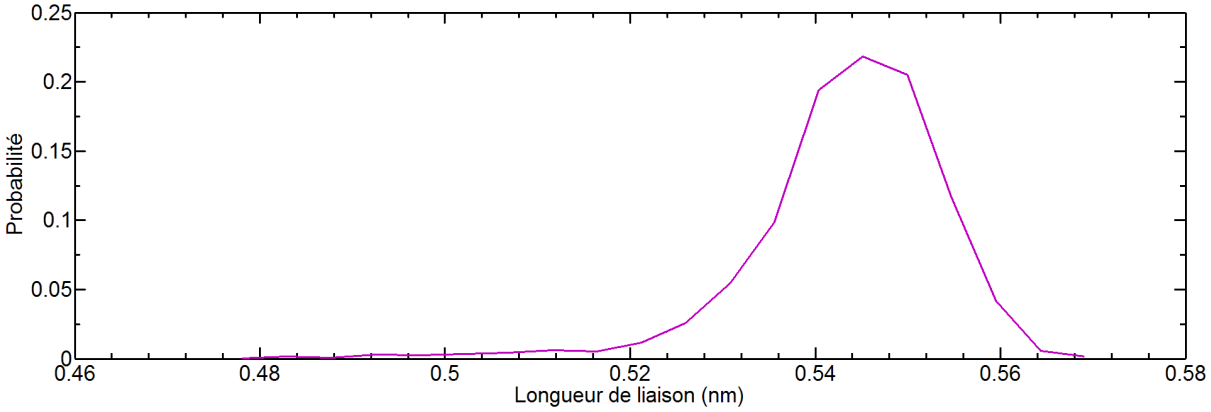

Figure II-26 Distribution de la longueur de liaison entre deux grains successifs le long de la chaîne moléculaire de xylane de DP = 25.

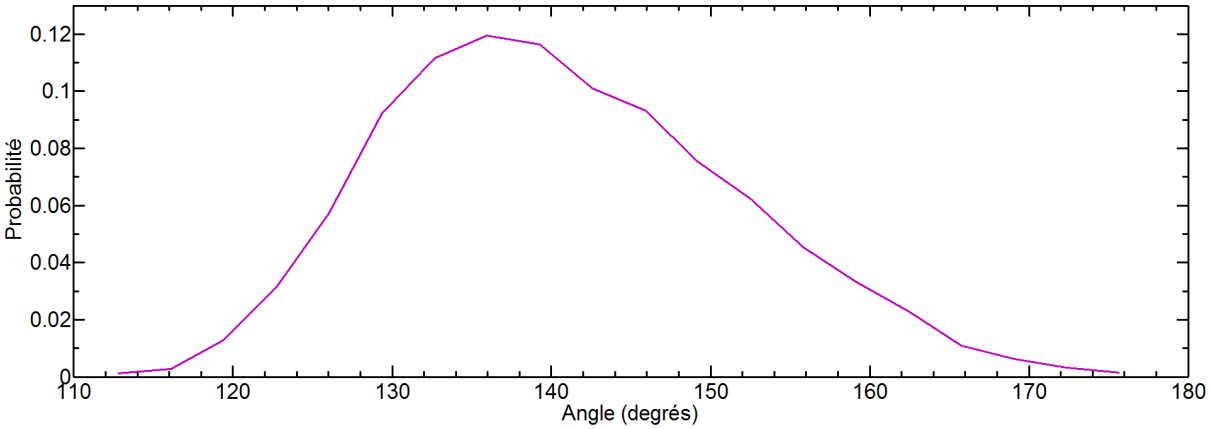

Figure II-27 Distribution de l'angle entre trois grains successifs le long de la chaîne moléculaire de xylane de DP = 25.

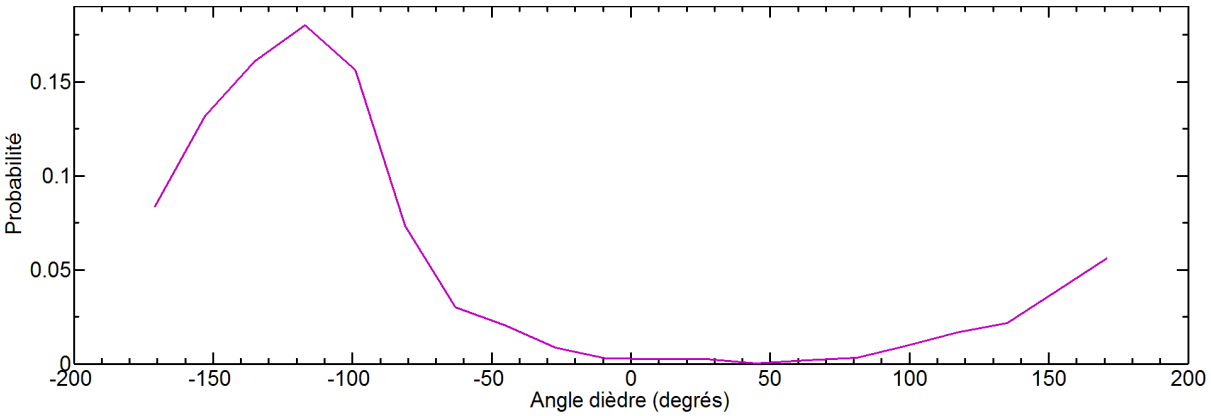

Figure II-28 Distribution de l'angle de torsion (angle dièdre) entre quatre grains successifs le long de la chaîne moléculaire de xylane de DP = 25.

II.3.2.4 Comparaison des modèles du xylane

Afin de donner une appréciation de la qualité d'échantillons de DP = 25, nous représentons les distributions de DP = 5, DP = 10 et DP = 25 dans les mêmes graphes pour une comparaison directe (Fig. II-29, Fig. II-30, Fig. II-31).

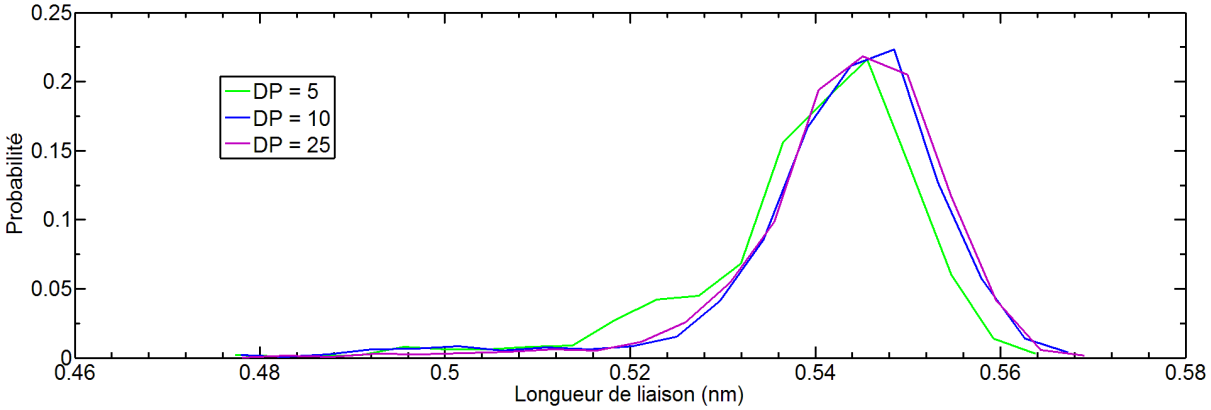

Figure II-29 Comparaison des trois distributions de longueur de liaison provenant des modèles moléculaires de xylane de DP = 5, 10, 25.

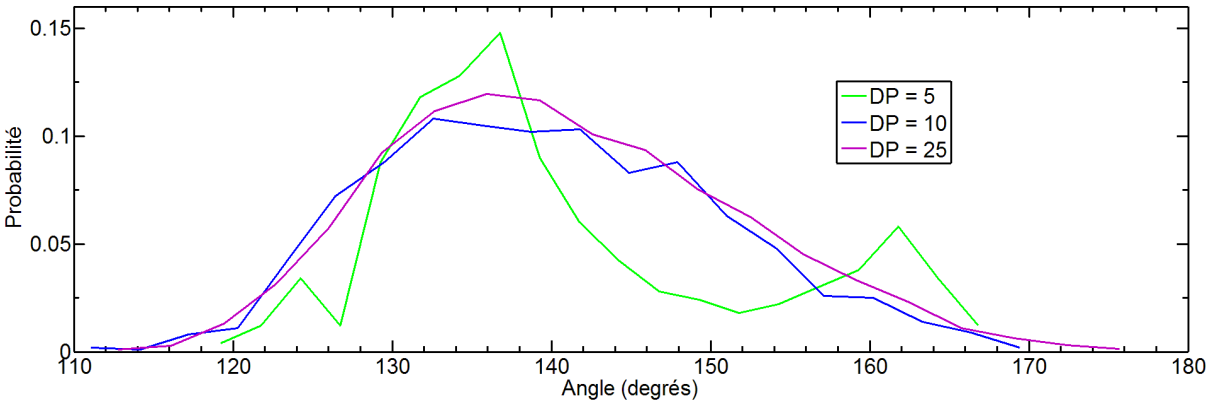

Figure II-30 Comparaison des trois distributions de l'angle provenant des modèles moléculaires de xylane de DP = 5, 10, 25.

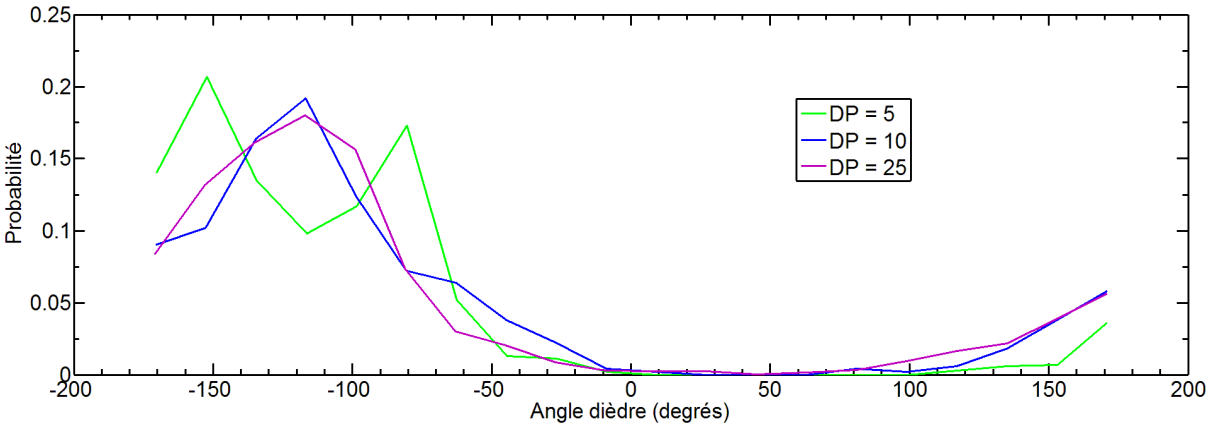

Figure II-31 Comparaison des trois distributions de l'angle dièdre provenant des modèles moléculaires de xylane de DP = 5, 10, 25.

Comme on peut le constater, les distributions de modèles DP = 10 et DP = 25 sont en fait très proches, mais le modèle DP = 25 fournit un échantillon plus, étendu, qui permet au modèle DP = 25 de fournir une distribution de meilleure qualité. Le modèle de DP = 25 a été adopté pour la modélisation de potentiels intramoléculaires.

II.3.3 CALIBRATION DE MODELES DE POTENTIELS INTRAMOLECULAIRES

La mise en œuvre de potentiels d'interaction gros grain peut s'appuyer, soit sur des potentiels tabulés, soit sur des expressions analytiques de modèles de potentiel classiques. C'est cette deuxième voie qui a été choisie dans le cadre de cette thèse. Dans ce cas, si on utilise une approche de type inversion de Boltzmann, le modèle de potentiel comporte un certain nombre de paramètres libres qu'il faut ajuster de façon à retrouver les distributions de probabilité correctes. Dans cette section, nous présentons la modélisation des trois termes d'interactions intramoléculaires à savoir : les potentiels d'élongation, de flexion et de torsion.

II.3.3.1 Potentiel d'élongation de liaison

Le modèle choisi pour l'élongation des liaisons est le potentiel harmonique. La minimisation de l'erreur entre la distribution issue de la trajectoire atomistique et celle déduite du modèle de potentiel aboutit au paramétrage suivant :

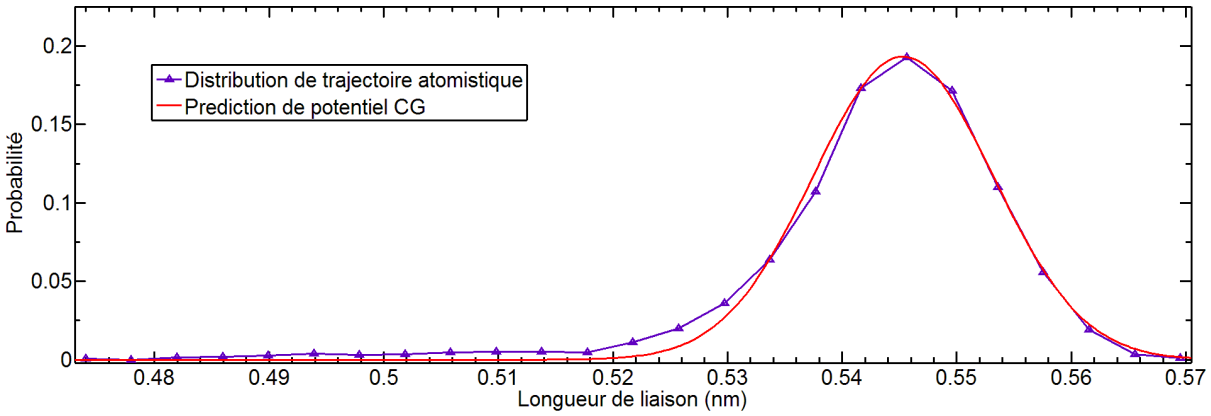

Figure II-32 Calibration du modèle harmonique pour le potentiel de liaison entre deux particules successives.

$$P(r) = Cr \cdot 4\pi r^2 \cdot e^{\frac{-k}{k_B T}\cdot(r-r_0)^2}$$ Équation II-5

$$\mathrm{k} = 20246{,}54\ \mathrm{kJ/nm^2 \cdot mol}\,;\mathrm{r_0} = 0{,}5451\ \mathrm{nm};\mathrm{T} = 300\ \mathrm{K}$$

Où *Cr* est le facteur de normalisation. On retrouve bien un ordre de grandeur correct pour la distance entre deux monomères successifs au sein d'une chaine.

II.3.3.2 Potentiel de déformation angulaire

Le modèle choisi pour la déformation angulaire est le potentiel « $\cos^2$ ». La minimisation de l'erreur entre la distribution issue de la trajectoire atomistique celle déduite du modèle de potentiel aboutit au paramétrage suivant :

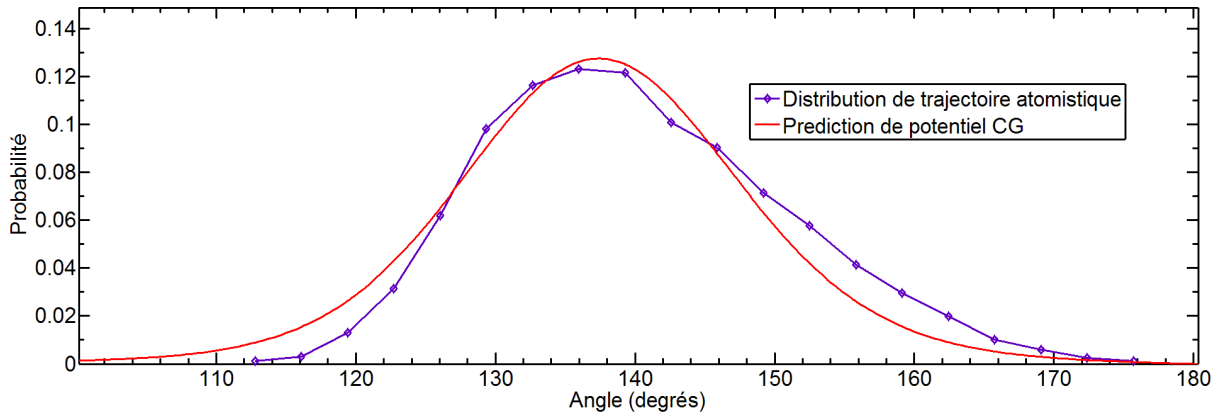

Figure II-33 Calibration du modèle pour le potentiel d'angle.

$$P(\theta) = C_\theta \cdot \sin\theta \cdot e^{\frac{-\beta}{k_B T}\cdot(\cos\theta - \cos\theta_0)^2}$$ Équation II-6

$$\beta = 74{,}75\,\mathrm{kJ/mol}\,;\,\theta_0 = 139{,}4°;\,\mathrm{T} = 300\ \mathrm{K}$$

Où C_θ est le facteur de normalisation.

II.3.3.3 Potentiel de torsion

Le type du potentiel pour l'angle dièdre est légèrement plus complexe : il prend la forme d'une série de Fourier.

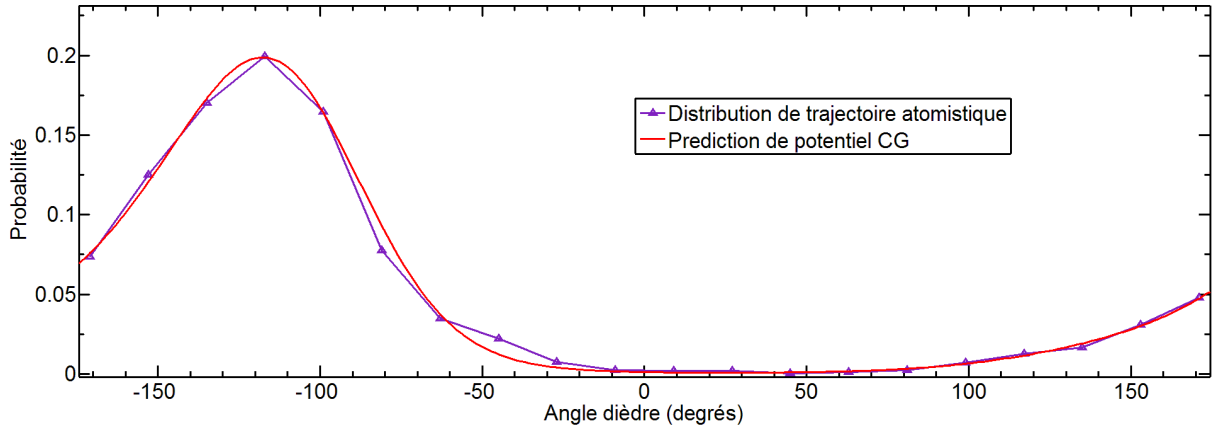

Figure II-34 Calibration du modèle pour le potentiel d'angle dièdre.

$$P(\varphi) = C \cdot exp(-\sum_{n=1}^{3} \frac{k_n}{k_B T} \cdot [1 - \cos(n\varphi - \varphi_n)])$$ Équation II-7

$$k_1 = -5{,}3179\ \mathrm{kJ/mol}\,;$$

$$k_2 = 0{,}2305\ \mathrm{kJ/mol}\,;\ k_3 = -0{,}4068\ \mathrm{kJ/mol}$$

$$\varphi_1 = 48{,}65^\circ;\ \varphi_2 = 170{,}74^\circ;\ \varphi_3 = 219{,}16^\circ;\mathrm{T} = 300\ \mathrm{K}$$

Où *C* est le facteur de normalisation. Il est à noter que la force de torsion n'a pas de sens lorsque l'un des deux angles, formés par *P1P2P3* et *P2P3P4* (Voir Fig. I-4), est égal à 180. Cette situation ne génère pas seulement un problème au niveau de la valeur de la force, mais également cause l'instabilité du calcul. Pour éviter ces problèmes, la force de torsion est progressivement diminuée jusqu'à son annulation lorsqu'un des angles s'approche de 180º. Dans nos simulations MD, la zone de diminution est de 170º à 178º, lorsqu'un des angles est supérieur à 178º, la force de torsion est fixée à 0. En supposant θ est l'angle *P1P2P3* ou *P2P3P4* (on choisit le plus défavorable, c'est-à-dire le plus grand des deux), le facteur de diminution dans l'intervalle de 170º à 178º est de type tangente hyperbolique :

$$Facteur = 1 - \tanh\{z\frac{\theta - \theta_{min}}{\theta_{max} - \theta_{min}}\}$$ Équation II-8

Où, *z* permet d'ajuster la vitesse de décroissance du facteur. Nous illustrons dans la figure II-35 un exemple de l'évolution du facteur pour la zone 170º ~178º.

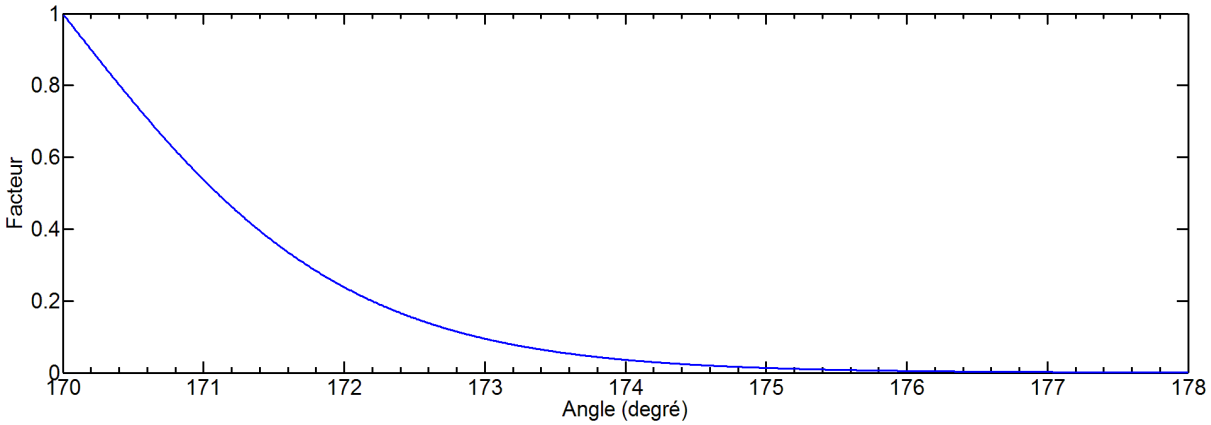

Figure II-35 Evolution du facteur lorsque l'angle *P2P3P4* s'approche de 180º.

II.3.4 VALIDATION DES POTENTIELS INTRAMOLECULAIRES

Les modèles précédemment calibrés doivent être validés. Il s'agit de s'assurer que les trajectoires générées à l'échelle gros grain à l'aide de ces modèles de potentiel ont une statistique correcte. En d'autres termes, il faut vérifier qu'on retrouve bien des distributions de probabilité très proches de celles déduites des trajectoires atomistiques.

Pour ce faire, il suffit de suivre les procédures ci-dessous.

- Implantation de potentiels modélisés dans le programme de MD qui a été développé dans le cadre de la thèse.
- Génération d'une chaîne de xylane de configuration aléatoire selon le potentiel modélisé. Cette chaîne sera utilisée en tant qu'état initial de la simulation. Pour les détails de la construction de la structure géométrique du xylane, voir III.1.

- Lancement de la simulation MD et génération d'une trajectoire en gros grain.
- Extraction directe de distributions de probabilité à partir de la trajectoire gros grain issue de l'étape précédente.
- Comparaison avec les échantillons de départ, issus de la trajectoire atomistique DP = 25.

On génère des distributions de distance entre grains successifs, d'angles, et d'angles dièdres. La trajectoire a été générée dans le programme MD sous les paramètres listés dans le tableau II-4.

Type	Ensemble	Temps total	Pas de temps	Température
Xylane de 25 grains	**NVT**	**8 ns**	**16 fs**	**300 K**

Tableau II-4 Les paramètres des simulations MD pour la validation de champs de force intramoléculaires.

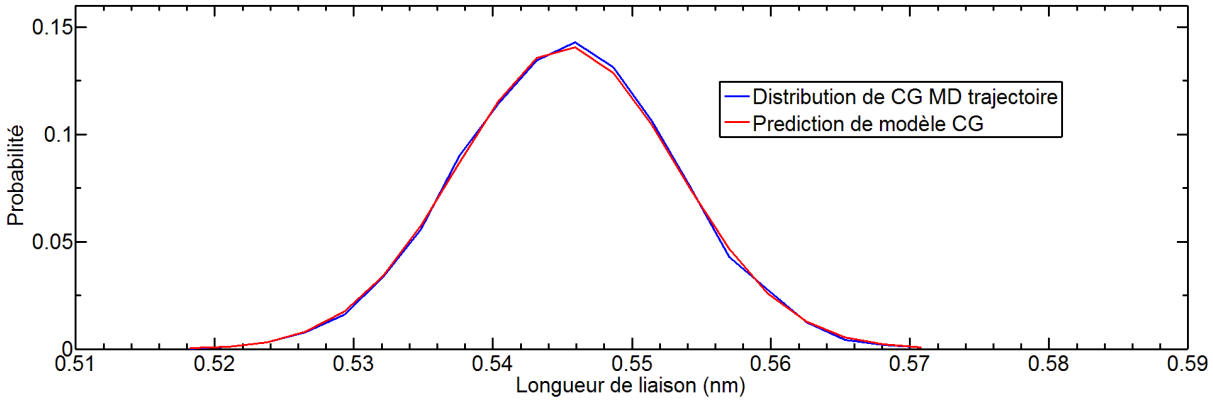

Figure II-36 Validation du potentiel de la liaison ; la courbe rouge est le modèle du potentiel modélisé, la courbe bleue est la distribution issue des simulations CG MD.

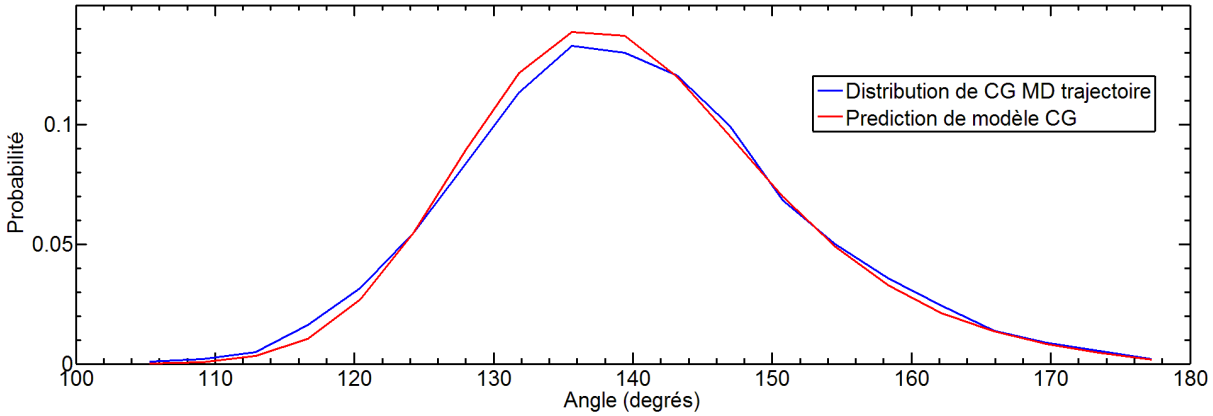

Figure II-37 Validation du potentiel de l'angle ; la courbe rouge est le modèle du potentiel modélisé, la courbe bleue est la distribution issue des simulations CG MD.

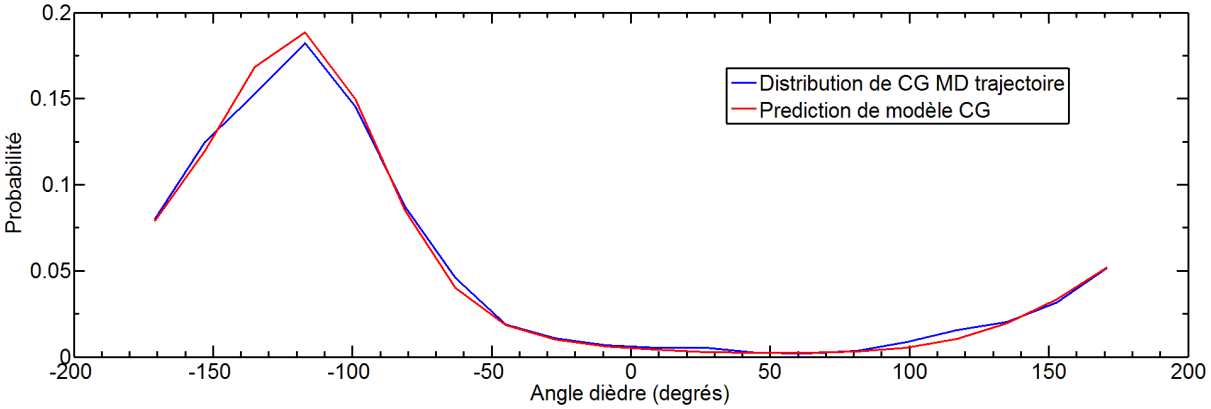

Figure II-38 Validation du potentiel de l'angle dièdre; la courbe rouge est le modèle du potentiel modélisé, la courbe bleue est la distribution issue des simulations CG MD.

Sur les figures ci-dessus (Fig. II-36, Fig. II-37, Fig. II-38), nous constatons un très bon accord pour les distributions de longueur de liaison, d'angle et d'angle dièdre. Cela permet de valider les paramètres de potentiels intramoléculaires.

II.4 Potentiel intermoléculaire xylane-xylane

Dans le cas des potentiels intramoléculaires, il est aisé de découpler les différentes interactions. La conséquence pratique est qu'on peut considérer que chaque distribution (longueur, angle, angle dièdre) peut être attribuée à un seul potentiel d'interaction. Cette relation univoque permet une déduction assez directe des potentiels. Par exemple, la distribution de la distance entre deux grains successifs d'une chaîne vient exclusivement du potentiel de liaison.

En revanche, pour la modélisation intermoléculaire, la distribution cible à reproduire est la conséquence de plusieurs paires de grains. Dans ce cas, même s'il est évidemment possible de déduire une distribution de distances grain-grains des trajectoires atomistiques, un potentiel qui en serait déduit ne pourrait être le potentiel d'interaction correct. Cette difficulté a mené à la mise en œuvre d'une méthode itérative, cette première distribution erronée n'étant alors que l'étape initiale d'un processus d'optimisation :

1. Déduire une distribution de distance grain-grain de la trajectoire atomistique et modéliser le potentiel à partir de cette distribution. Le paramétrage ainsi obtenu est évidemment faux et ne servira qu'à initialiser le processus itératif.
2. Balayer les valeurs des paramètres du potentiel intermoléculaire, un par un, autour la valeur initiale par les simulations CG MD. Chaque paramétrage du potentiel génère une distribution, le paramétrage qui génère la distribution la plus proche de la distribution cible est retenu.
3. Lancer plusieurs balayages des paramètres afin de minimiser l'erreur entre la distribution cible et la distribution générée. Autrement dit, répéter plusieurs fois l'étape 2 pour obtenir le paramétrage du potentiel qui génère la distribution la plus proche de la distribution cible.

La modélisation requiert donc de nombreuses simulations CG MD pour optimiser progressivement le paramétrage du potentiel à son optimum. Une inversion de Boltzmann est appliquée sur la distribution cible pour obtenir le paramétrage initial du potentiel xylane-xylane.

Il aurait été à priori possible d'utiliser une approche de type « Force-Matching », de façon à déduire directement la force recherchée. Dans le cas présent, cette démarche s'est avérée totalement infructueuse. La cause de cet échec est assez évidente. Le « Force-Matching » repose sur le calcul de forces instantanées, ce qui implique que la dérivation d'une force gros grain nécessite une prise de moyenne temporelle ou statistique de cette force. En d'autres termes, pour une même distance grain-grain donnant une même force grain-grain, il existe de nombreuses conformations possibles des monomères concernés à l'échelle atomistique donnant des valeurs de force différentes qu'il est nécessaire de moyenner. C'est pourquoi le Force-Matching est plus exigeant que l'inversion de Boltzmann en matière d'échantillon atomistique. Or, l'échantillon dont nous disposons est insuffisant pour répondre à ces critères. C'est pourquoi la dérivation du potentiel intermoléculaire xylane-xylane s'appuie sur une inversion de Boltzmann.

II.4.1 DERIVATION DE LA DISTRIBUTION CIBLE

II.4.1.1 Procédure générales

La première étape consiste à déterminer une distribution de référence, appelée « la distribution cible » par la suite. L'accord avec cette distribution est le critère de la modélisation du potentiel. En l'absence d'une distribution radiale extraite d'une phase amorphe de xylane simulée par dynamique moléculaire atomistique, seule à même de faire consensus, nous avons extrait trois distributions de distance à partir de la trajectoire xylane-xylane (Fig. II-39) :

1. « Toutes distances» : Il s'agit de la distribution de toutes les distances entre les grains n'appartenant pas à la même chaine.
2. « Distance entre milieux » : Elle est la distance entre les grains de milieu de chaîne. Le défaut de cet échantillonnage est évident : lorsque ces deux grains s'éloignent à cause d'un décalage des deux chaînes, l'interaction entre les grains choisis n'est plus significative.
3. « Plus courte distance » : Cette distribution concerne les deux grains définissant la plus courte distance entre deux chaînes.

Le modèle de Morse et le modèle de Lennard-Jones 12-6 ont été utilisés pour décrire l'interaction intermoléculaire xylane-xylane. Le potentiel de Morse a été choisi car le potentiel de Lennard-Jones 12-6 n'a pas permis d'obtenir des résultats satisfaisants.

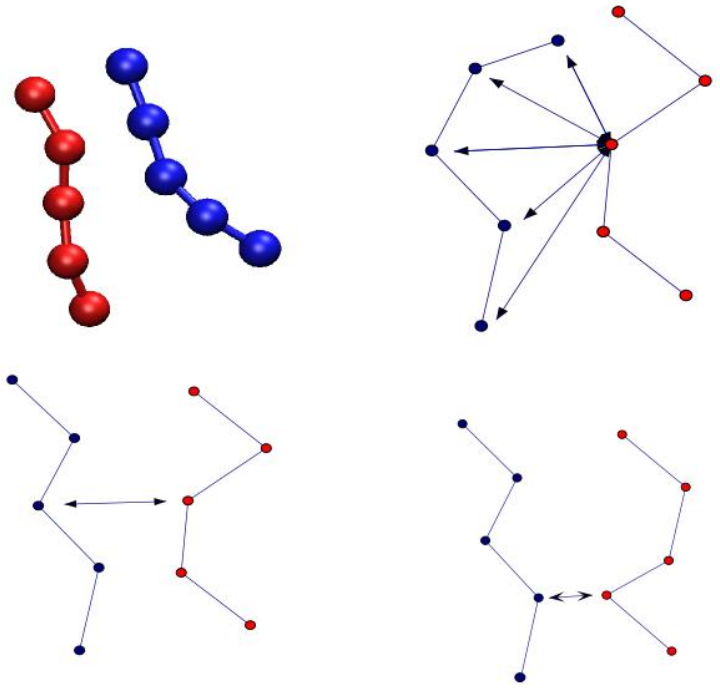

Figure II-39 Figure en haute à gauche : les deux chaînes de xylane CG. Figure en haut à droite : toutes les distances entre le grain du milieu de la chaîne rouge et tous les grains de la chaîne bleue. Figure en bas à gauche : la distance entre les deux grains du milieu provenant respectivement des deux chaînes du xylane. Figure en bas à droite : la plus courte distance entre deux chaînes définie par les deux grains le plus proche de deux chaînes.

II.4.1.2 Toutes distances

Comme on peut le constater sur la figure II-40, la distribution alimentée par toutes les distances est bruitée, en particulier à longue distance. L'ajustement d'un modèle de Morse à partir de cette distribution ne donne pas de résultats satisfaisants. Nous n'utiliserons donc pas cette distribution.

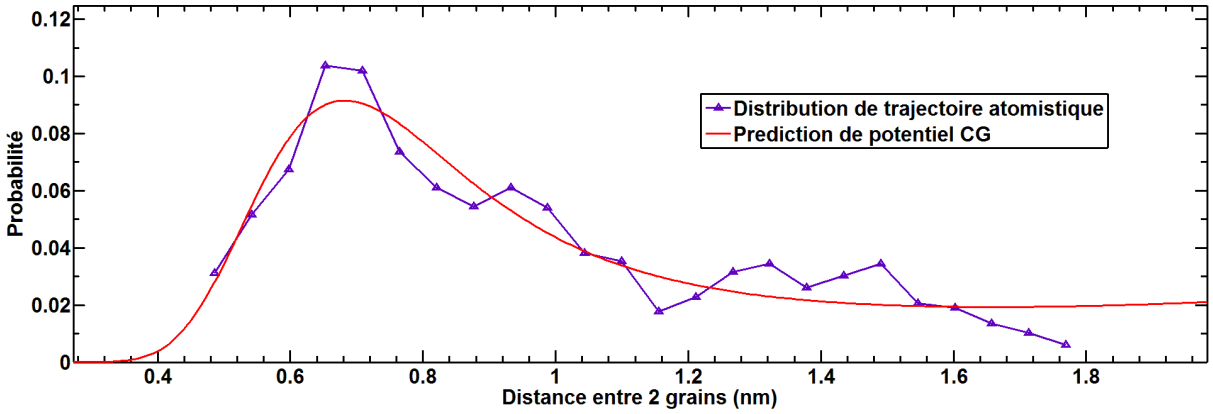

Figure II-40 Ajustement d'un modèle de potentiel de Morse sur la distribution « Toutes distances »

II.4.1.3 Distance entre milieux

La distribution « Distance entre milieux » est moins bruitée que la distribution « Toutes distances », elle est exposée dans la figure II-41.

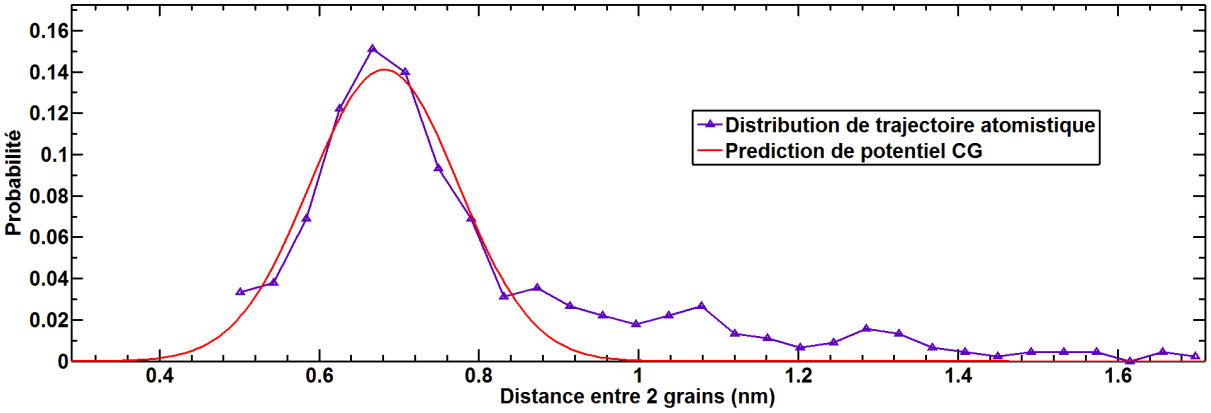

Figure II-41 Ajustement d'un modèle de potentiel de Morse sur la distribution « Distance entre milieux ».

Le potentiel de Morse permet un ajustement correct à courte distance, mais se trouve en défaut à longue distance. Cette distribution présente également un défaut plus fondamental. Lorsque les deux grains choisis pour l'alimenter sont éloignés l'un de l'autre, la distance qui sépare les deux chaînes est déterminée pas des interactions entre d'autres grains. Les valeurs prises par la distribution de probabilité à longue distance sont donc peu signifiantes.

II.4.1.4 Plus courte distance

La distribution de plus courte distance est bien moins bruitée que les deux précédentes et permet un ajustement acceptable d'un modèle de Morse qui est présenté dans la figure II-42.

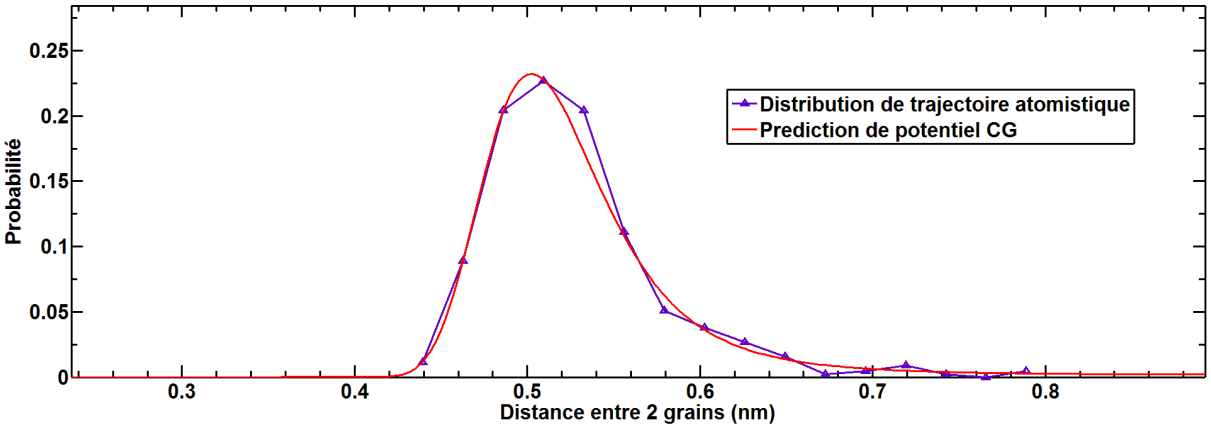

Figure II-42 Ajustement d'un modèle de potentiel de Morse sur la distribution « Plus courte distance ».

Il faut ici souligner le fait que cette distribution est davantage signifiante physiquement que les précédentes. D'une part, les grains choisis à chaque pas sont effectivement en interaction, et, d'autre part, cette distribution est représentative de l'interaction globale de deux chaines de xylane.

On adopte donc le modèle de Morse pour modéliser le potentiel d'interaction intermoléculaire xylane-xylane. Le paramétrage initial est donc :

$$U_{Morse}(r) = D_e \cdot (1 - e^{-\alpha(r-r_0)})^2 \qquad \text{Équation II-9}$$

$$D_e = 18{,}687\ \text{kJ/mol}\,;\ \alpha = 7{,}79\ \text{nm}^{-1};\ r_0 = 0{,}4985\ \text{nm}$$

Ces valeurs seront utilisées pour initialiser le processus d'optimisation du potentiel.

II.4.2 AJUSTEMENT A POSTERIORI DU POTENTIEL XYLANE-XYLANE

II.4.2.1 Définition d'une fonction erreur

Le paramétrage du modèle de potentiel xylane-xylane que nous avons obtenu dans la section II.4.1 permet d'initialiser le processus d'ajustement. Afin de sélectionner un paramétrage optimal du modèle de potentiel intermoléculaire entre deux grains, il est indispensable de définir un critère. Dans notre cas, la différence entre deux distributions est quantifiée par la fonction erreur suivante :

$$\rho = \sum \exp\left(-\frac{|r - r_m|}{L}\right) \cdot |g(r) - g_0(r)|$$ Équation II-10

Où, r est la distance entre deux grains, r_m et L sont respectivement la position du pic et la distance d'équilibre provenant de la distribution cible, $g(r)$ et $g_0(r)$ sont la distribution générée par CG MD et la distribution cible. Le facteur exponentiel permet de donner plus de poids aux erreurs commises dans l'intervalle le plus probable de r. On s'efforcera de minimiser la valeur de la fonction erreur.

II.4.2.2 Procédures d'ajustement à posteriori.

L'ajustement à posteriori du potentiel xylane-xylane consiste en une succession de rounds d'optimisation. A chaque round d'optimisation, chaque paramètre du modèle de potentiel est optimisé, et chaque optimisation de paramètre consiste en une série de simulations MD CG pour plusieurs valeurs du paramètre optimisé, les autres paramètres étant fixés. Pour chacune de ces séries de simulations, on choisit une valeur du paramètre qui minimise l'erreur.

1. Considérons le potentiel du xylane-xylane de type Morse pour l'interaction xylane-xylane avec les paramètres initiaux provenant de l'inversion de Boltzmann de la distribution « Plus courte distance » de la section II.4.1.
2. Les paramètres D_e et α sont temporairement fixés, le paramètre x_e est varié autour sa valeur initiale pour 20 simulations MD CG. Pour chaque simulation, on calcule la fonction erreur ρ entre la distribution généré $g(r)$ et la distribution ciblée $g_0(r)$. Les 20 simulations donnent un ensemble de 20 valeurs de la fonction erreur.
3. L'ensemble de ces valeurs est fitté par une fonction de type polynomial quadratique, la valeur r_0 correspondant au minimum de la fonction est considérée comme étant la valeur optimale de r_0 dans cette itération. Voir la figure II-43.

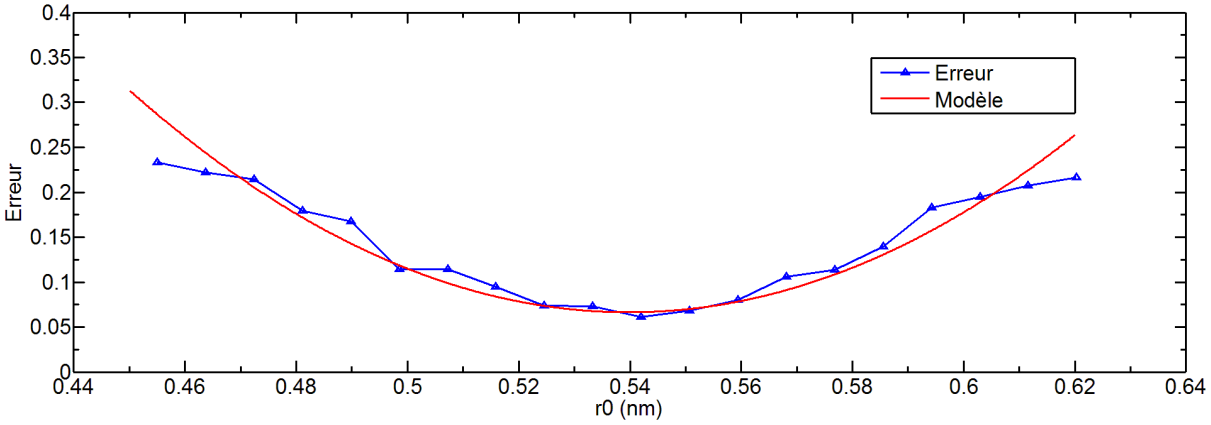

Figure II-43 Optimisation du paramètre r_0 dans une itération de paramètres. La courbe bleue: les valeurs de la fonction erreur pour chacune des simulations MG CG avec différentes valeurs de r_0. La courbe rouge : fit de la fonction erreur fonction de r_0. La valeur minimale de cette fonction correspond au r_0 optimisé.

4. Les paramètres α et D_e sont séquentiellement optimisés par la même méthode d'optimisation de r_0. Bien entendu chaque paramètre est optimisé en fixant les deux autres.

On répète toute l'itération, les étapes 1-4, jusqu'à ce que la fonction erreur aie une valeur suffisamment faible.

II.4.2.3 Détails de l'ajustement de paramètres a posteriori

Les simulations CG MD sont menées avec les paramètres ci-dessous :

Type	Ensemble	Temps total	Pas de temps	Température
Xylane-Xylane	**NVT**	**8 ns**	**16 fs**	**300 K**

Tableau II-5 Les paramètres des simulations CG MD de l'ajustement à posteriori.

Pour générer la configuration initiale du système moléculaire CG, deux chaînes du xylane de DP = 5 sont aléatoirement générées, elles sont ensuite poussées doucement pour s'approcher jusqu'à ce qu'elles interagissent. Pour les détails de la construction de la structure géométrique du xylane, voir III.1.

Sachant qu'une itération de paramètres pour le potentiel xylane-xylane contient 60 simulations, par souci de simplicité, nous ne montrons que l'évolution de la fonction erreur et les distributions en fonction du numéro d'itération. La distribution générée après chaque itération de paramètres est comparée avec la distribution cible par la fonction erreur, cela donne l'évolution de la fonction erreur ci-dessous :

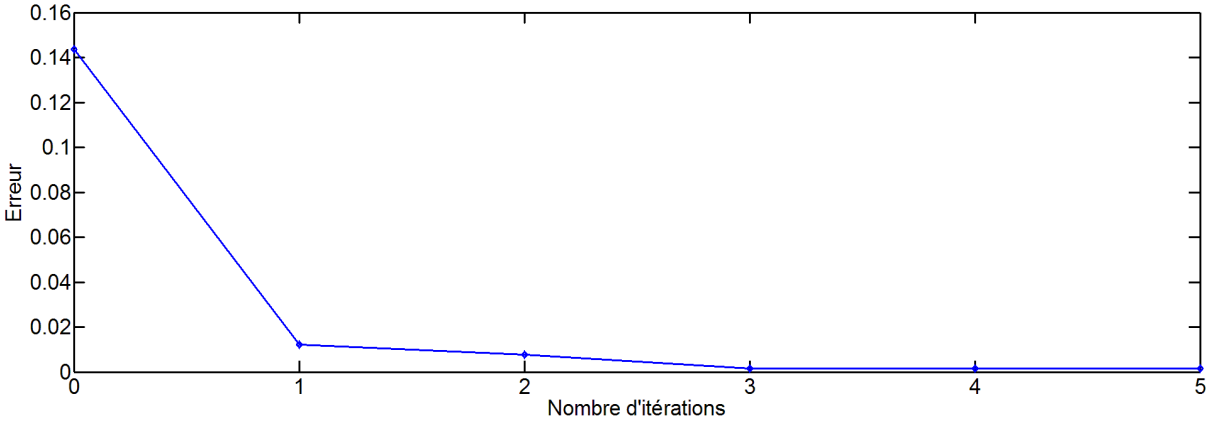

Figure II-44 Evolution de la fonction erreur en fonction du numéro d'itération.

Dans la figure II-44, l'itération 0 correspond à la configuration initiale dérivée de l'inversion de Boltzmann dans la section II.4.1. La valeur de la fonction erreur décroit progressivement de 0,1436 jusqu'à 0,0015. Au-delà de la quatrième itération, le terme d'erreur est faible, cela veut dire que la différence entre la distribution générée et la distribution cible est minimisée après 4 itérations de la procédure d'optimisation. Les distributions générées pour les 5 itérations sont présentées dans la figure II-45 ainsi que la distribution cible pour comparaison.

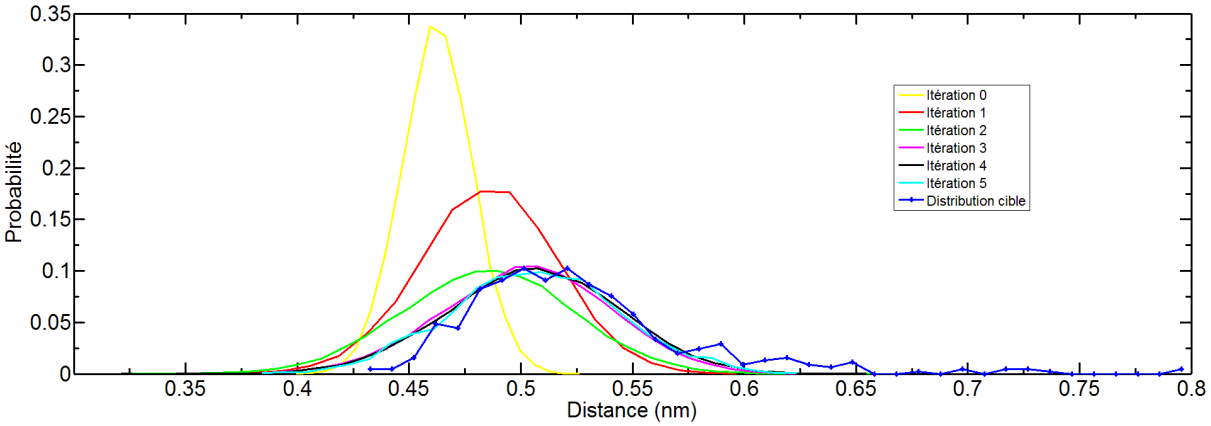

Figure II-45 Evolution de la distribution de la plus courte distance xylane-xylane au cours de l'optimisation des paramètres du potentiel CG. La courbe bleue avec points : la distribution originale de la plus courte distance entre deux chaînes du xylane de la trajectoire issue de la trajectoire atomistique.

L'ensemble de paramètres pour chaque itération est présenté ci-dessous :

	r_0 (*nm*)	α (nm^{-1})	D_e (*kJ/mol*)
Itération 0	**0,4985**	**7,790**	**18,687**
Itération 1	**0,5666**	**5,911**	**7,465**
Itération 2	**0,5633**	**3,494**	**6,392**
Itération 3	**0,5688**	**3,456**	**6,221**
Itération 4	**0,5764**	**3,421**	**6,141**
Itération 5	**0,5731**	**3,475**	**6,102**

Tableau II-6 Paramètres du potentiel inter-xylane pour chaque itération de la procédure d'optimisation.

On choisit finalement les paramètres de la dixième itération : r_0 = 0,5731 *nm*, α = 3,475 nm^{-1}, D_e = 6,102 *kJ/mol*.

II.5 Potentiel intermoléculaire xylane-cellulose

La modélisation du potentiel d'interaction intermoléculaire cellulose-xylane suivra une procédure similaire à celle qui a abouti au modèle de potentiel intermoléculaire xylane-xylane : un premier paramétrage sera obtenu par une méthode directe, puis celui-ci sera utilisé afin d'initialiser une optimisation itérative. Toute la procédure sera réalisée à l'aide d'une trajectoire atomistique : l'adsorption d'une chaine de 5 monomères de xylane sur un cristal de cellulose dont la surface exposée est de type 110.

II.5.1 Traitement de la trajectoire atomistique

La trajectoire CG issue de la trajectoire atomistique a dû être légèrement modifiée pour faciliter les calculs. En effet, dans les fichiers PDB décrivant la trajectoire, la cellulose « subit » les conditions aux limites périodiques alors que la chaine de xylane ne les « subit » pas. Cette différence de méthode de sauvegarde de trajectoire amène le xylane à sortir, en apparence, de la boîte de simulation (Fig. II-46).

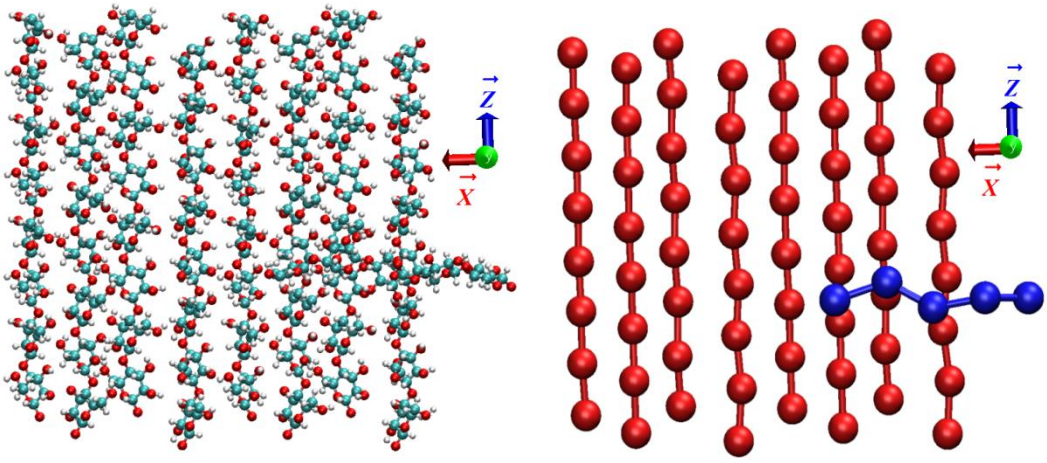

Figure II-46 La chaîne de xylane sort de la boîte de simulation. La figure à gauche : trajectoire atomistique ; la figure à droite : la version CG.

Afin d'éviter ce problème, il suffit d'étendre la surface de cellulose en appliquant les conditions aux limites périodiques dans le plan *XZ* (Fig. II-47).

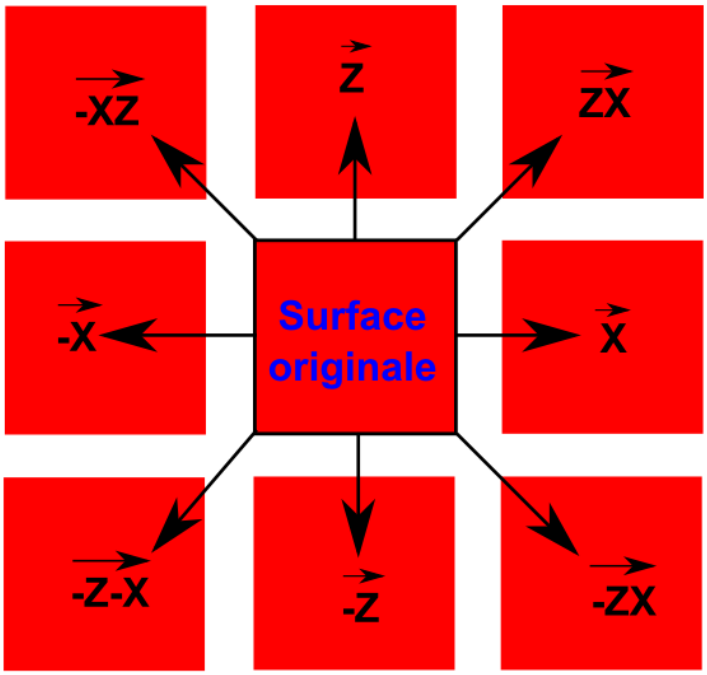

Figure II-47 Duplication du système moléculaire cellulose-xylane vers les directions *X* et *Z*.

II.5.2 CALIBRATION INITIALE DU MODELE DE POTENTIEL

Quelques essais ont été réalisés avec des paramétrages issus d'une inversion de Boltzmann. Ceux-ci s'étant montrés infructueux, le paramétrage initial du modèle de potentiel a été déduit d'une procédure de Force-Matching. Il s'agit de calculer les forces subies par le monomère central de la chaine de xylane de la part de chacun des monomères de la surface de cellulose. Le calcul de cette force est évidemment réalisé à partir de la trajectoire atomistique.

La force subie par le monomère central de xylane de la part d'un des monomères de cellulose est la somme des forces subies par tous les atomes du monomère de xylane de la part de tous les atomes du monomère de cellulose considéré. Les interactions prises en compte sont donc les interactions de Van der Waals et les interactions électrostatiques. Afin de réduire le volume de calculs à exécuter, on ne considèrera, à chaque itération extraite de la trajectoire atomistique, que les monomères de cellulose en-deçà d'un *cutoff* fixé à 1,4 nm (Fig. II-48).

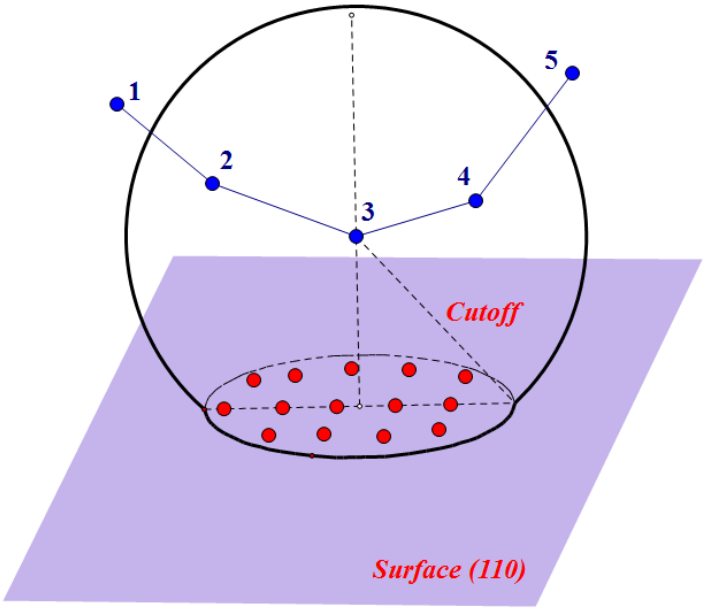

Figure II-48 Sélection des grains de cellulose interagissant avec le grain de xylane central, définition du « *cutoff* ».

On obtient ainsi, pour chaque itération extraite, une série de couples de monomères xylane-cellulose, ainsi qu'une force s'exerçant sur le monomère de xylane pour chacun de ces couples. La force est projetée sur le vecteur qui lie le barycentre des deux monomères concernés, puisque nous recherchons une interaction de symétrie sphérique. Une moyenne des forces calculées et projetées est effectuée par intervalles de 0,0289 nm (Fig. II-49).

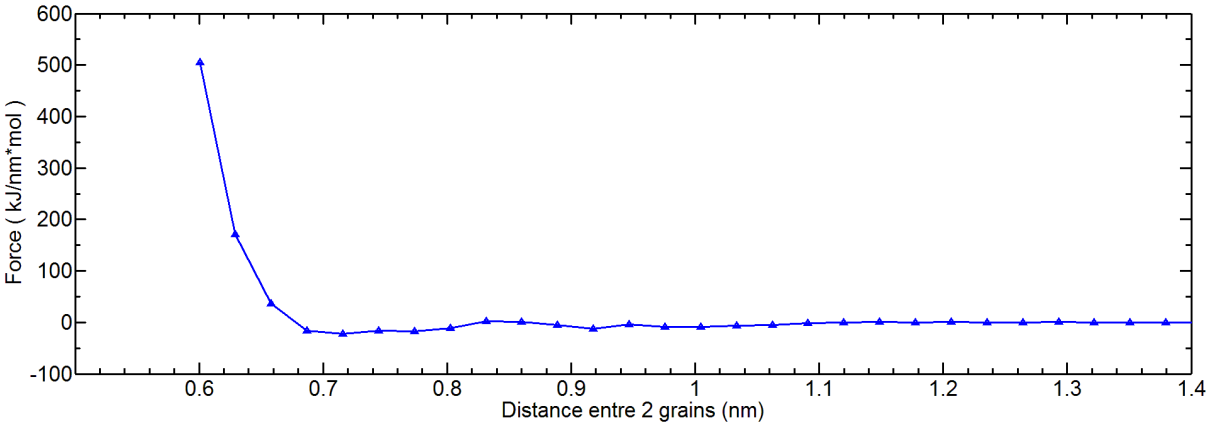

Figure II-49 Force reçue par le grain 3 en fonction de la distance.

La force ainsi obtenue a été modélisée via un potentiel de Lennard-Jones 12-6 et un potentiel de Morse. Cependant, le modèle de Morse n'a pas donné de résultats concluants, il

n'est donc pas utilisé comme modèle du potentiel xylane-xylane. Par conséquent, seul l'ajustement d'un modèle de Lennard-Jones 12-6 est présenté ici (Fig. II-50).

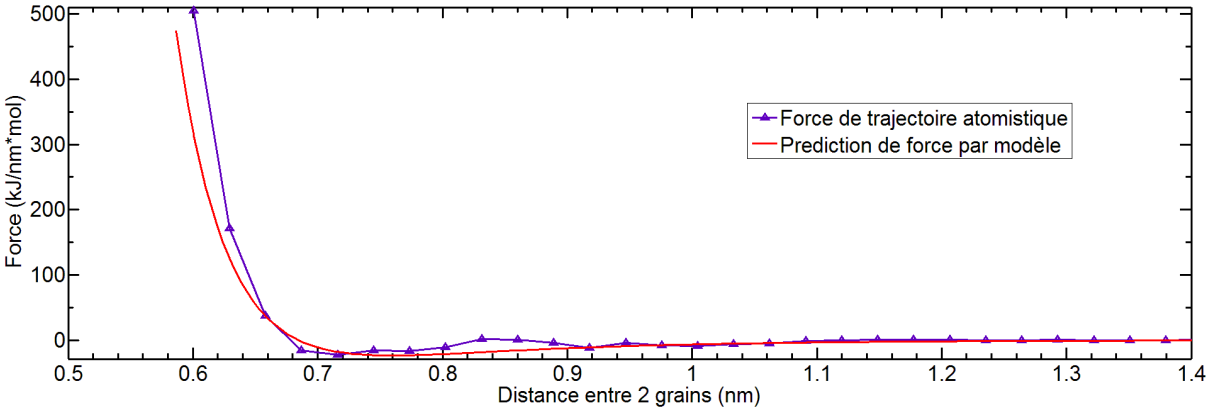

Figure II-50 Ajustement d'un modèle de Lennard-Jones 12-6 avec la force issue du Force-Matching.

On obtient les paramètres suivants :

$$\varepsilon = 32{,}77\ \mathrm{kJ/mol}\,;\sigma = 0{,}57\ \mathrm{nm} \qquad \text{Équation II-11}$$

II.5.3 OPTIMISATION DU POTENTIEL

II.5.3.1 La distribution cible

Comme dans le cas de l'interaction xylane-xylane, il est impératif d'établir un critère pour valider les paramétrages obtenus. Pour ce faire, nous imposons comme critère de référence pour la validation du potentiel la distribution de la distance entre le grain du milieu de la chaîne de xylane et la surface de la cellulose définie par la coordonnée *Y* moyenne des grains de la couche superficielle. (Fig. II-51). Cette distribution sera la distribution cible.

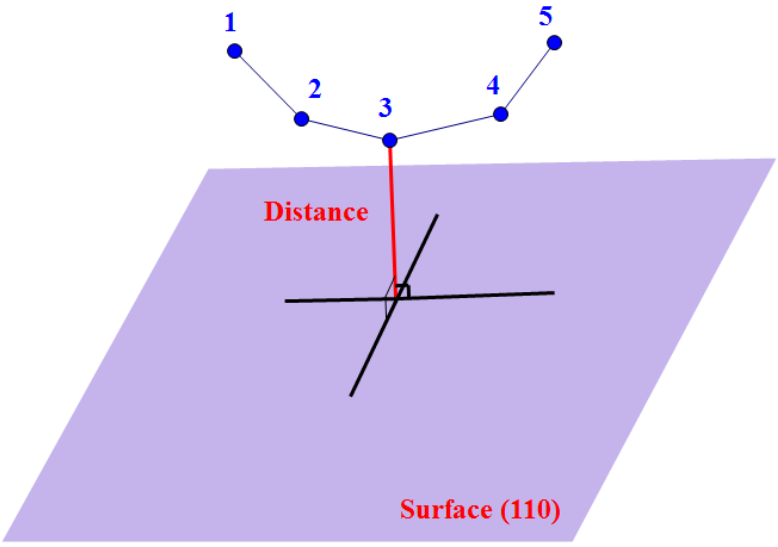

Figure II-51 Distance entre le grain 3 et la surface de la cellulose est le paramètre de validation.

En revanche, nous choisissons toujours le grain du milieu de la chaîne du xylane pour éviter l'effet de bord. La distribution cible, extraite de la trajectoire atomistique, est exposée dans la figure II-52.

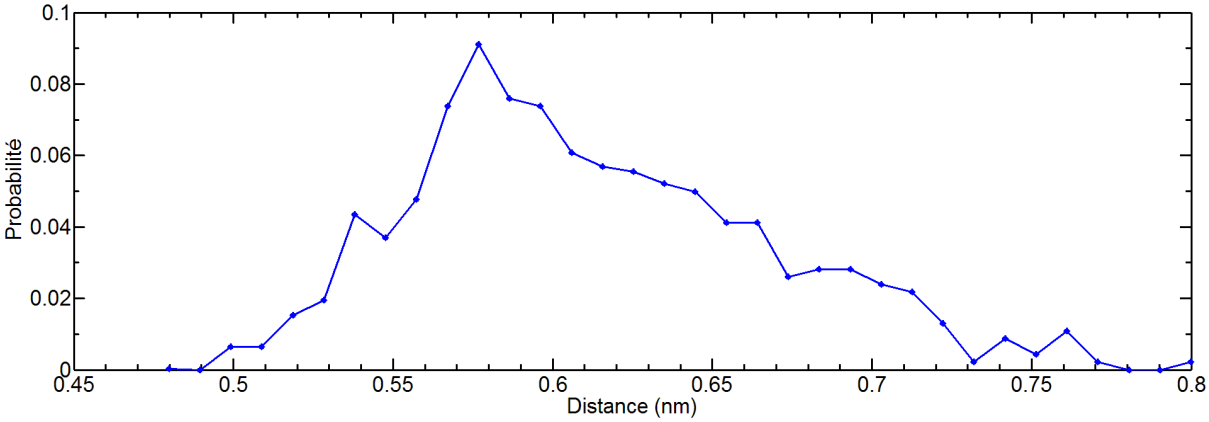

Figure II-52 Distribution de probabilité de la distance entre le grain central de la chaine de xylane et la surface de cellulose.

La configuration initiale du système moléculaire est une chaîne du xylane de DP = 5 avec une surface de 8 chaînes de DP = 8. Les paramètres des simulations sont les suivants :

Type	Ensemble	Temps total	Pas de temps	Température
Xylane-Cellulose	**NVT**	**8 ns**	**16 fs**	**300 K**

Tableau II-7 Paramètres de simulation CG MD de test du potentiel provenant du Force-Matching.

Comme on pouvait s'y attendre, la distribution issue du potentiel Lennard-Jones 12-6 est éloignée de la distribution cible (Fig. II-53).

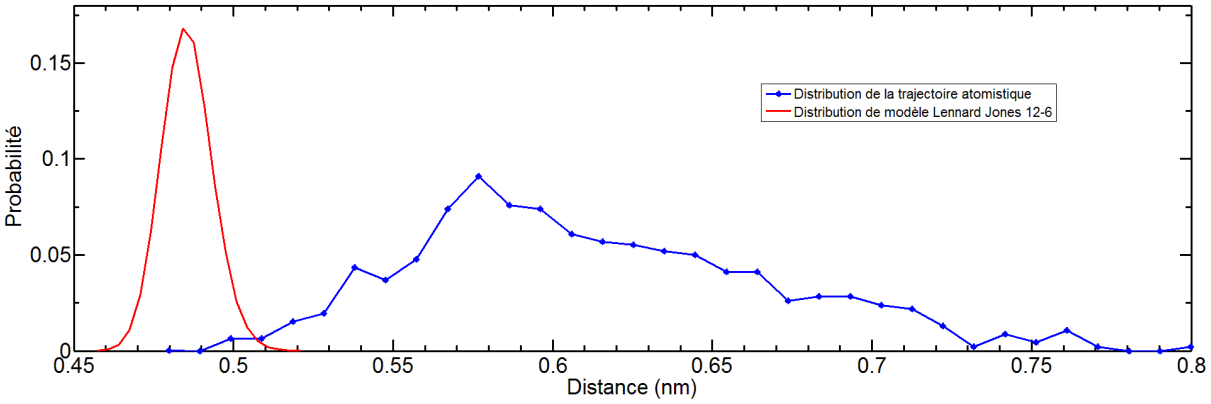

Figure II-53 Distribution de probabilité de la distance xylane-cellulose, d'une part issue de la dynamique moléculaire CG avec le potentiel de Lennard-Jones 12-6, et, d'autre part, issue de la trajectoire atomistique.

On peut expliquer cette constatation en considérant que la distribution issue du modèle, d'une part, et la distribution cible, d'autre part, ne résultent pas des mêmes interactions. En effet, alors que seules les chaines de cellulose de surface sont considérées dans les simulations CG, les monomères de xylane, et en particulier le monomère central, interagissent avec trois couches de cellulose dans les cas de la simulation atomistique. C'est pourquoi l'optimisation du potentiel est indispensable. Ainsi, l'interaction des monomères de xylane avec les monomères de cellulose des chaînes de cœur sera indirectement modélisée dans le modèle CG, sous la forme d'une interaction grain-grain effective.

II.5.3.2 Ajustement à posteriori du potentiel

L'ajustement des paramètres du potentiel xylane-cellulose est similaire à celui de du potentiel xylane-xylane. Plusieurs itérations sont réalisées pour minimiser la différence entre la distribution de la distance xylane-cellulose générée et la distribution cible. La fonction erreur (Eq. II-10) utilisée est la même que pour l'optimisation du potentiel xylane-xylane. Les étapes sont détaillées ci-dessous :

1. Le paramétrage du potentiel de type Lennard-Jones 12-6 est utilisé pour initialiser la procédure d'optimisation.
2. Le paramètre ε est fixé, σ varie autour sa valeur courante pour 20 simulations CG MD. Pour chaque simulation, la fonction erreur est calculée à partir de la différence entre la distribution générée $g(r)$ et la distribution cible $g_0(r)$. Les 20 simulations donnent un ensemble de 20 valeurs de la fonction erreur.

3. Une fonction de type polynomiale quadratique est utilisée pour fitter cet ensemble, la valeur σ correspondant au minimum est considérée comme optimale dans cette itération.
4. Le paramètre σ est fixé à la valeur optimale, le paramètre ε est varié autour de sa valeur courante pour 20 simulations CG MD. Les erreurs sont calculées entre les distributions générées et la distribution cible. La valeur optimale de ε est identifiée comme précédemment. Les deux paramètres optimaux sont par conséquent déterminés pour cette itération. L'itération suivante commence à l'étape 2.

Les paramètres des simulations CG MD sont listés ci-dessous :

Type	Ensemble	Temps total	Pas du temps	Température
Xylane-Cellulose	**NVT**	**8 ns**	**16 fs**	**300 K**

Tableau II-8 Paramètres des simulations CG MD de l'ajustement à posteriori du potentiel intermoléculaire cellulose-xylane.

La fonction erreur décroît rapidement avec le nombre d'itérations (Fig. II-54).

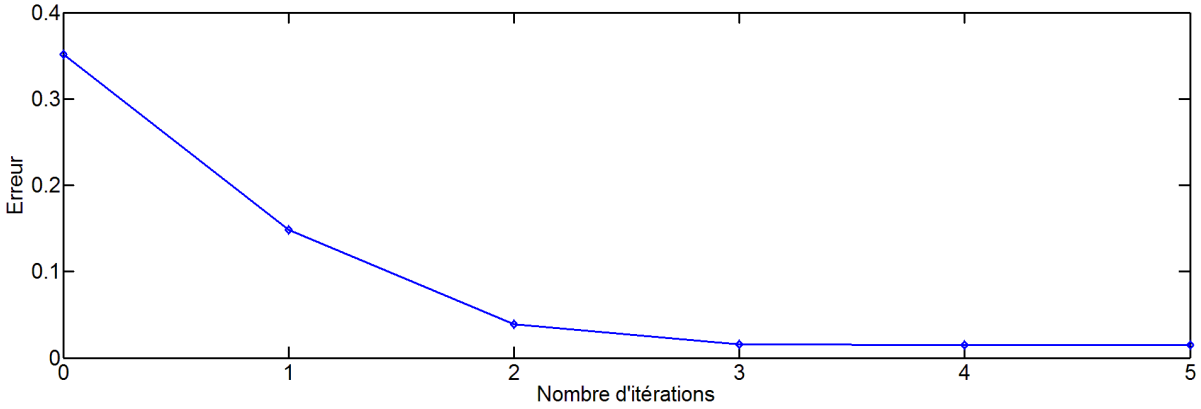

Figure II-54 Evolution de la fonction erreur entre la distribution générée et la distribution cible avec le numéro d'itération de la procédure d'optimisation.

On peut constater que 5 itérations suffisent pour atteindre une fonction erreur suffisamment faible.

Les distributions prédites à l'aide des itérations successives d'optimisation du potentiel sont de plus en plus proches de la distribution cible à mesure que les itérations se succèdent (Fig. II-55).

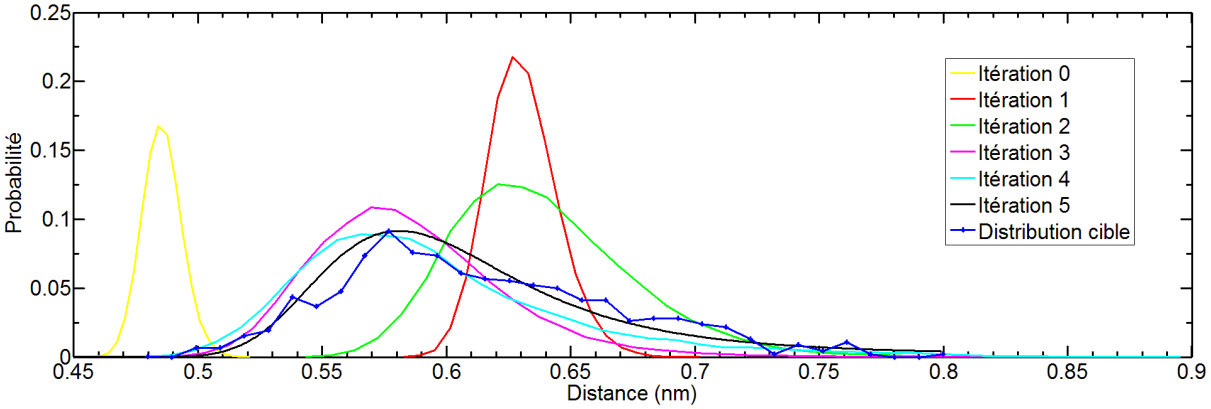

Figure II-55 Distributions de probabilité de la distance xylane-surface de cellulose prédites par simulation CG pour les cinq premières itérations d'optimisation. Comparaison avec la distribution-cible.

Les valeurs prises par les paramètres du potentiel intermoléculaire cellulose-xylane à chaque itération d'optimisation sont données ci-dessous :

	σ (*nm*)	ε (*kJ/mol*)
Itération 0	**0,572**	**32,77**
Itération 1	**0,721**	**22,87**
Itération 2	**0,677**	**3,54**
Itération 3	**0,654**	**2,22**
Itération 4	**0,633**	**2,32**
Itération 5	**0,635**	**2,13**

Tableau II-9 Paramètres du potentiel xylane-cellulose pour chaque itération d'optimisation.

On choisit finalement les paramètres issus de la cinquième itération de paramètres : σ = 0,635 *nm* et ε = 2,13 *kJ/mol*.

II.6 Conclusions

Dans ce chapitre, les modèles atomistiques du xylane et de la cellulose ont été présentés, ainsi que les matériels informatiques et la stratégie du modèle gros grain.

Les potentiels intramoléculaires CG du xylane ont été établis par l'inversion de Boltzmann à partir des trajectoires atomistiques. Les potentiels établis sont validés par la cohérence entre les distributions statistiques générées par la simulation CG et celles extraites des trajectoires atomistiques. Il est à noter que les potentiels intramoléculaires établis ne nécessitent aucune optimisation de paramètres à posteriori, cela signifie que la méthodologie développée dans la section II.3 est valable dans le cas de la modélisation intramoléculaire.

En revanche, la modélisation des potentiels intermoléculaires est différente que la modélisation intramoléculaire. La modélisation et la validation des potentiels intermoléculaires ont été réalisée via la mesure de la différence entre deux distributions statistiques : la distribution cible et la distribution générée. La distribution cible est extraite de la trajectoire atomistique. La distribution générée est extraite de trajectoire gros grain qui est générée par les potentiels CG modélisés. Dans les sections II.4 et II.5, les configurations initiales des potentiels intermoléculaires CG ont été déduites directement à partir des trajectoires atomistiques par l'inversion de Boltzmann et par le Force-Matching. Néanmoins, avec ces configurations initiales, la distribution cible ne peut être précisément reproduite par la simulation CG. Cela signifie que les configurations initiales des potentiels ne sont pas valides. Par conséquent, un procédé d'optimisation de paramètres de potentiels a été engagé afin de minimiser la différence entre la distribution cible et la distribution générée. Cette optimisation de paramètres à posteriori permet d'établir les valides paramètres de potentiels intermoléculaires. La méthodologie développée dans ces deux sections est valable dans le cas de la modélisation intermoléculaire.

Le modèle CG établi dans ce chapitre permet d'exploiter la conformation dynamique du composite cellulose-xylane à l'échelle gros grain, les différentes applications du modèle CG seront présentées dans le troisième chapitre.

III. APPLICATIONS ET RÉSULTATS

Résumé du chapitre III

Plusieurs applications du champ de force gros grain précédemment décrit sont ici exposées.

Une démarche de génération aléatoire des états initiaux est d'abord présentée. Un algorithme de génération aléatoire des chaînes de xylane est déduit du potentiel gros grain et appliqué au calcul de la longueur de persistance et les modèles gros grain de cellulose cristalline sont déduits des données expérimentales. Une stratégie de disposition des chaînes de xylane au voisinage de la cellulose cristalline est également détaillée.

Ensuite, l'adsorption d'une chaîne de xylane sur trois surfaces de cellulose cristalline est mise en œuvre. La conformation de la chaîne adsorbée est analysée. L'expérience numérique inverse, qui consiste à arracher une chaîne adsorbée à l'aide d'une pointe AFM, est également mise en œuvre.

Enfin, l'interaction d'une phase amorphe de xylane avec la surface de la cellulose cristalline est simulée dans deux situations différentes. Premièrement, une surface de cellulose cristalline 110 est utilisée, ce qui permet d'identifier un phénomène de structuration du xylane, sous la forme de couches successives, au voisinage de la cellulose. Dans un second temps, un modèle gros grain de cristal de cellulose, exposant les faces 110, 100 et 1-10, est utilisé. Une structuration du xylane contrastée en fonction des faces, est observée au voisinage de la cellulose.

III.1 Constructions des configurations initiales

III.1.1 GENERATIONS DES CHAINES DE XYLANE

III.1.1.1 Le mécanisme de génération

TIRAGE ALEATOIRE DES ENSEMBLES DES PARAMETRES

La première étape de l'application de l'approche gros grain mise au point dans cette thèse consiste à générer des états initiaux à partir desquels le logiciel MD pourra générer des trajectoires.

Dans le cas de la construction des configurations initiales des chaînes moléculaires du xylane, les longueurs des liaisons, les angles et les angles dièdres, doivent être échantillonnées à partir de la statistique de Boltzmann des potentiels CG modélisés dans le chapitre II.3.3.

Le premier pas de la génération des configurations initiales est la construction à partir des distributions de longueurs de liaisons, d'angles et d'angles dièdres, de leurs propres fonctions de répartition. Nous avons les distributions du modèle CG au niveau des liaisons, des angles et des angles dièdres ci-dessous (Fig. III-1, Fig. III-2, Fig. III-3) :

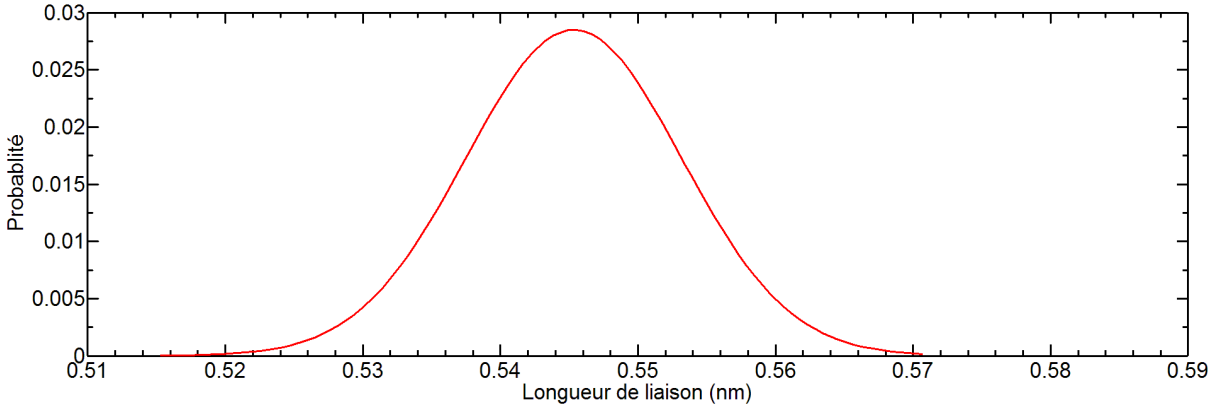

Figure III-1 Distribution de la longueur de liaison entre deux grains successifs du modèle CG de xylane.

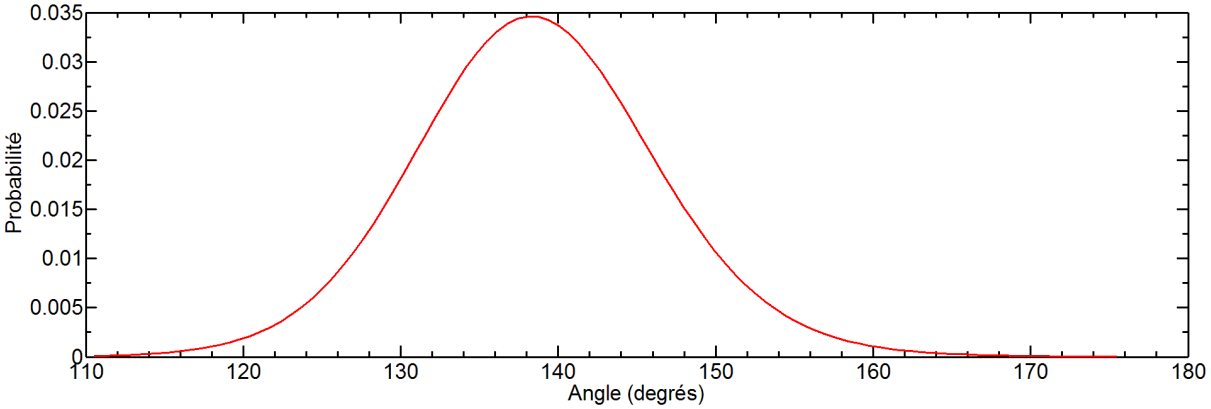

Figure III-2 Distribution de l'angle entre trois grains successifs du modèle CG de xylane.

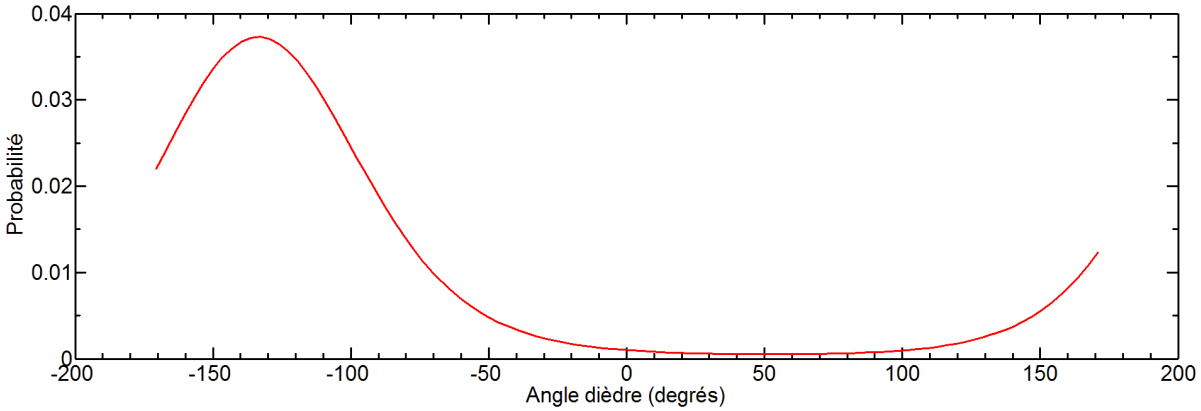

Figure III-3 Distribution de l'angle dièdre entre quatre grains successifs du modèle CG de xylane.

Nous avons respectivement pour chaque distribution, une fonction de répartition ci-dessous (Fig. III-4, Fig. III-5, Fig. III-6) :

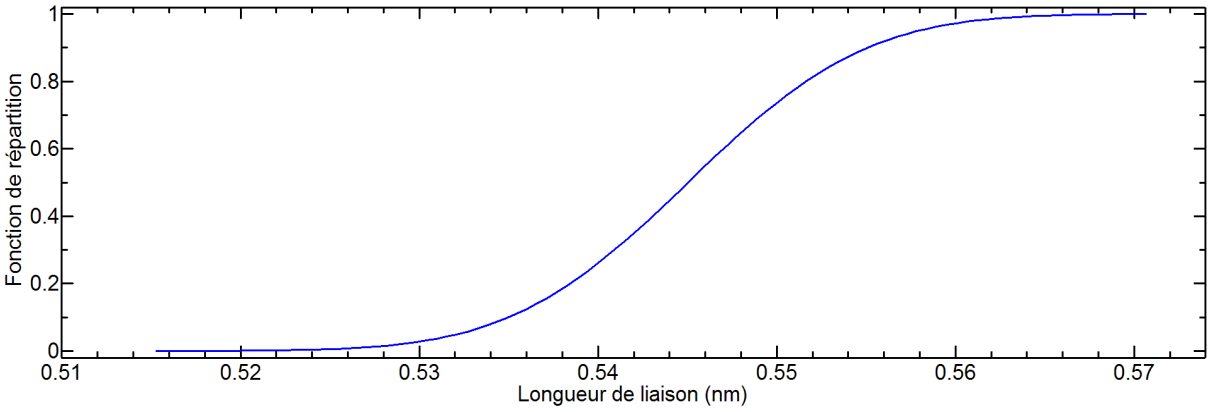

Figure III-4 Fonction de répartition de la longueur de liaison du modèle CG.

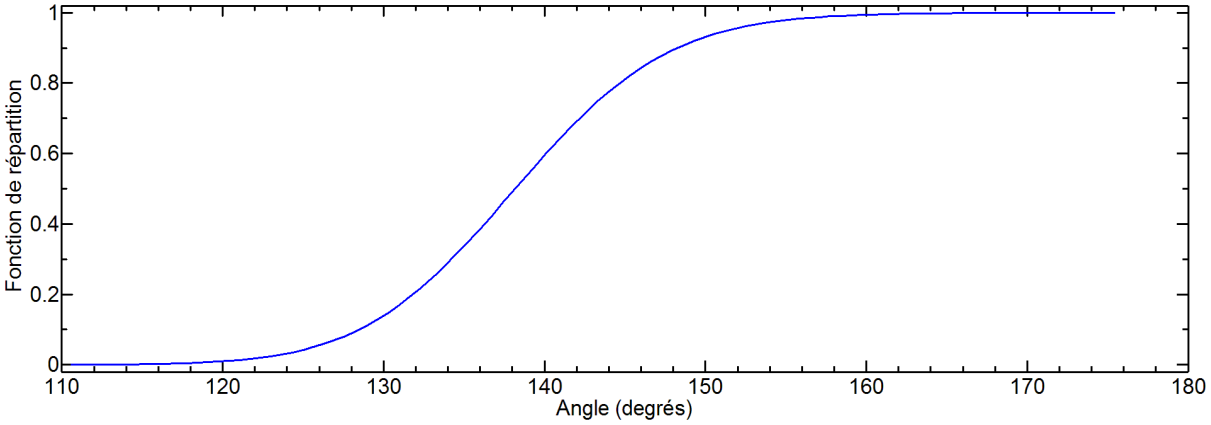

Figure III-5 Fonction de répartition de l'angle du modèle CG.

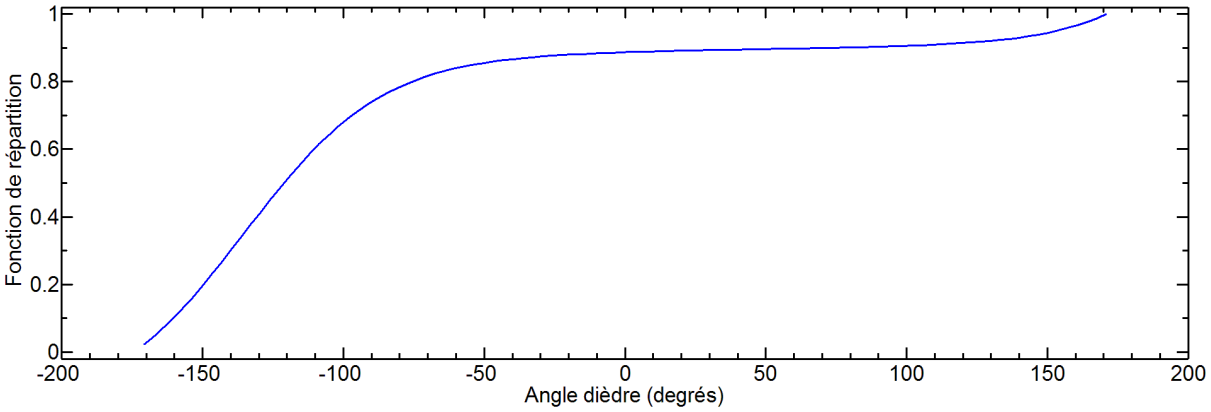

Figure III-6 Fonction de répartition de l'angle dièdre du modèle CG

Ensuite, il s'agit de sélectionner aléatoirement une suite de ces trois paramètres à partir des fonctions de répartition. On tire aléatoirement un nombre réel entre 0 et 1, correspondant à une valeur de la fonction de partition (ordonnée), et on en déduit la valeur du paramètre recherché (abscisse). Cette procédure permet d'obtenir des valeurs des paramètres conformes aux distributions du modèle CG.

CONSTRUCTION DES CHAINES DE XYLANE.

Une fois que le tirage aléatoire des paramètres de chaînes du xylane est réalisé, il s'agit ensuite de construire les chaînes de xylane, c'est-à-dire générer les coordonnées des grains. Cette construction géométrique des conformations des molécules peut être résumée par les étapes suivantes.

Etape Un : (Fig. III-7) Considérons un point initial *i* quelconque comme la position du premier grain de la chaîne de xylane. Nous copions le grain *i* et le translatons vers une direction quelconque pour construire le deuxième grain *j*. La distance de la translation est issue du tirage aléatoire de longueur de liaison. Cela forme les deux premiers grains de la chaîne du xylane.

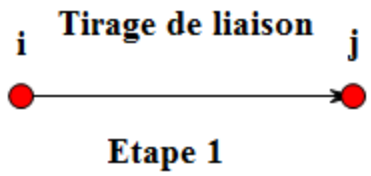

Figure III-7 La première étape de la construction géométrique d'une chaîne du xylane. « Tirage de liaison » signifie un tirage aléatoire de la valeur de liaison.

Etape Deux : (Fig. III-8) De la même façon, nous copions le grain *j* et le translatons dans la même direction *ij* avec un deuxième tirage aléatoire de la longueur de liaison, ceci nous permet d'obtenir le troisième grain temporaire *k'*.

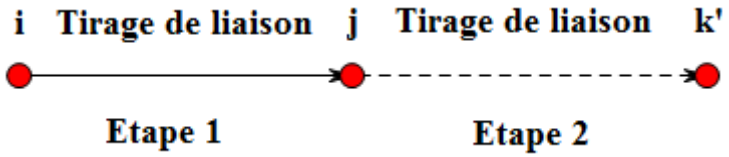

Figure III-8 La deuxième étape de la construction géométrique d'une chaîne du xylane. « Tirage de liaison » signifie un tirage aléatoire de la valeur de liaison.

Etape Trois : (Fig. III-9) La liaison *jk'* subit alors une rotation autour du vecteur perpendiculaire à la figure traversant le grain *j* dont la valeur d'angle est issue d'un tirage aléatoire de l'angle. Le grain *k'* devient donc le troisième grain définitif *k* de la chaîne de xylane.

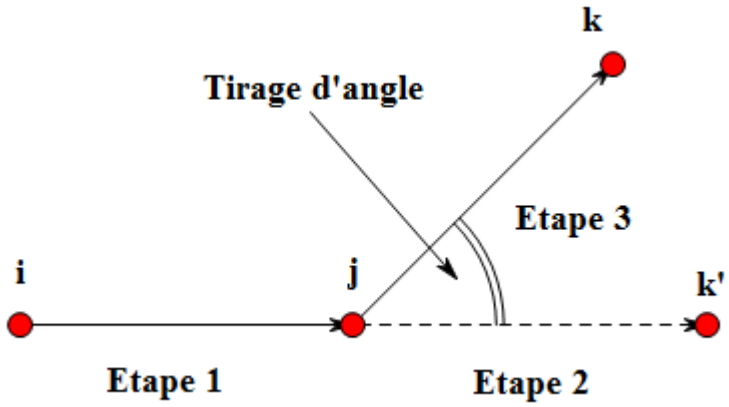

Figure III-9 La troisième étape de la construction géométrique d'une chaîne du xylane. « Tirage d'angle » signifie un tirage aléatoire de la valeur de l'angle.

Etape Quatre : (Fig. III-10) Ensuite, le grain *k* est copié et translaté dans la direction *jk*, la distance de translation est issue d'un troisième tirage aléatoire de longueur de liaison. Cela désigne le quatrième grain temporaire *l'*. A ce stade, les grains *i*, *j*, *k* et *l'* sont dans un même plan.

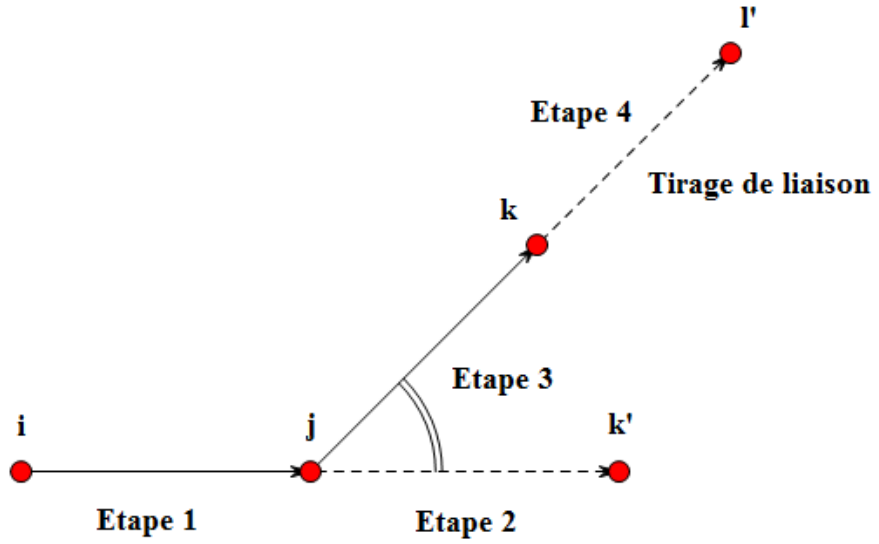

Figure III-10 La quatrième étape de la construction géométrique d'une chaîne du xylane. « Tirage de liaison » signifie un tirage aléatoire de la valeur de liaison.

Etape Cinq : (Fig. III-11) La liaison *kl'* subit une rotation autour du vecteur perpendiculaire au plan *ijk* traversant le grain *k* pour une valeur d'angle issue du deuxième tirage aléatoire d'angle. Cela forme le nouveau grain temporaire *l''*. A ce stade, les grains *i*, *j*, *k* et *l''* sont encore dans un même plan.

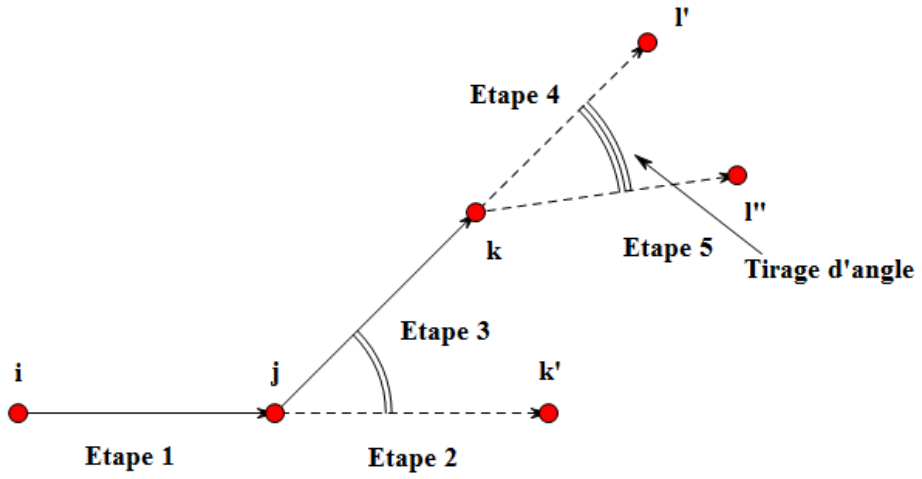

Figure III-11 La cinquième étape de la construction géométrique d'une chaîne du xylane. « Tirage d'angle » signifie un tirage aléatoire de la valeur de l'angle.

Etape Six : (Fig. III-12) La liaison *kl''* subit finalement une rotation autour du vecteur *jk* pour former l'angle dièdre dont la valeur provient du tirage aléatoire de l'angle dièdre. Pour générer les grains suivants, il suffit de répéter les mêmes procédures de la génération du quatrième grain de l'étape 4 jusqu'à 6.

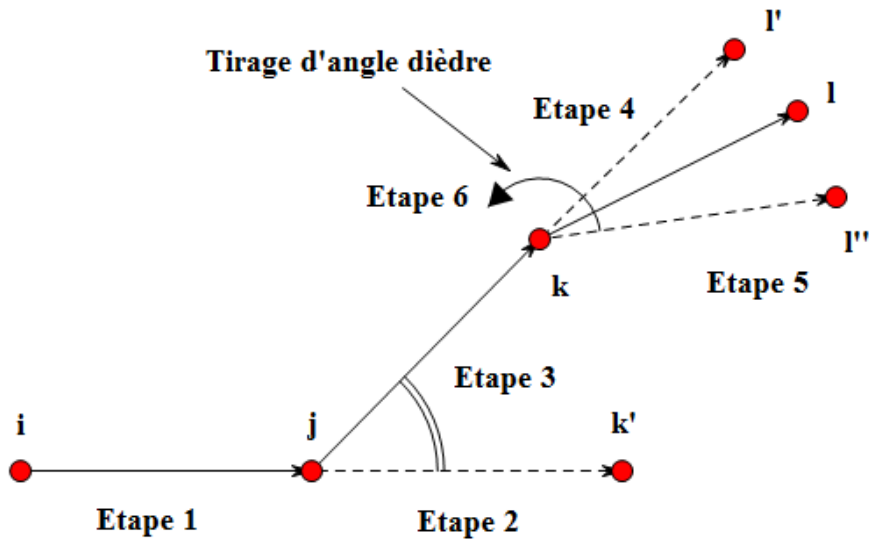

Figure III-12 La sixième étape de la construction géométrique d'une chaîne du xylane. « Tirage d'angle dièdre » signifie un tirage aléatoire de la valeur de l'angle dièdre.

Les six étapes ci-dessus nous permettent de construire des conformations initiales aléatoires des chaînes de xylane à partir des distributions de paramètres correspondants. Nous montrons dans la figure III-13 quelques exemples de chaînes de xylanes générées selon la méthode présentée.

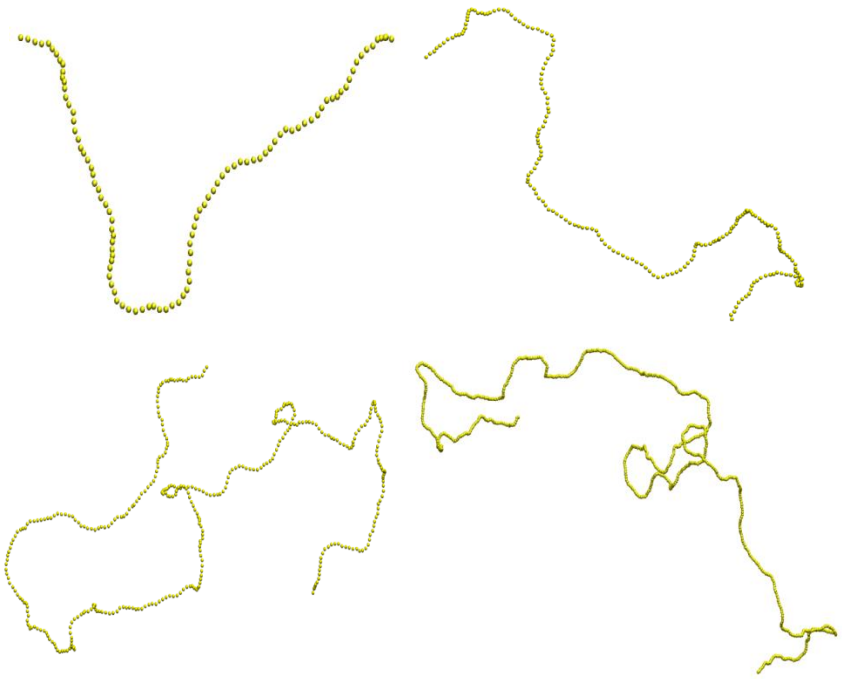

Figure III-13 Configurations initiales de chaînes de xylane. La figure en haut à gauche : une chaîne de xylane de DP = 100. La figure en haut à droite : une chaîne de xylane de DP = 200.La figure en bas à gauche : une chaîne de xylane de DP = 400.La figure en bas à droite : une chaîne de xylane de DP = 800.

III.1.1.2 Validations et Résultats

VALIDATIONS

Dans la section précédente, nous avons présenté la méthode de la génération des configurations initiales des chaînes moléculaires de xylane à l'échelle CG ainsi que quelques

chaînes, à titre d'exemple. Pour vérifier la validité de l'échantillonnage du tirage aléatoire, nous avons comparé les distributions statistiques du modèle CG du xylane avec celles des configurations initiales pour la longueur de la liaison, la valeur de l'angle et de l'angle dièdre. Dans les trois figures suivantes (Fig. III-14, Fig. III-15, Fig. III-16), les statistiques de trois termes d'interactions proviennent d'un ensemble des configurations initiales de 10000 chaînes de xylane de DP = 200.

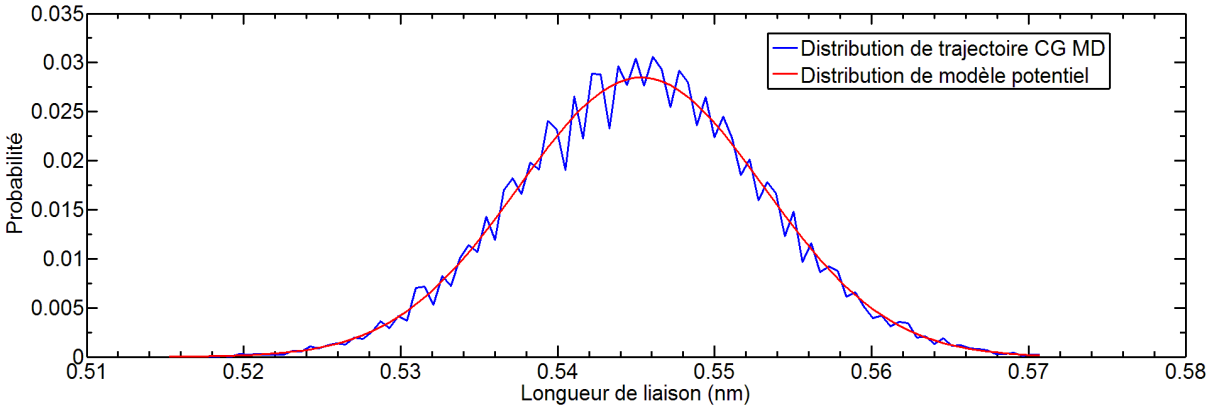

Figure III-14 Comparaison des distributions de longueur de liaison provenant du modèle CG, d'une part, et déduite de l'échantillon généré, d'autre part.

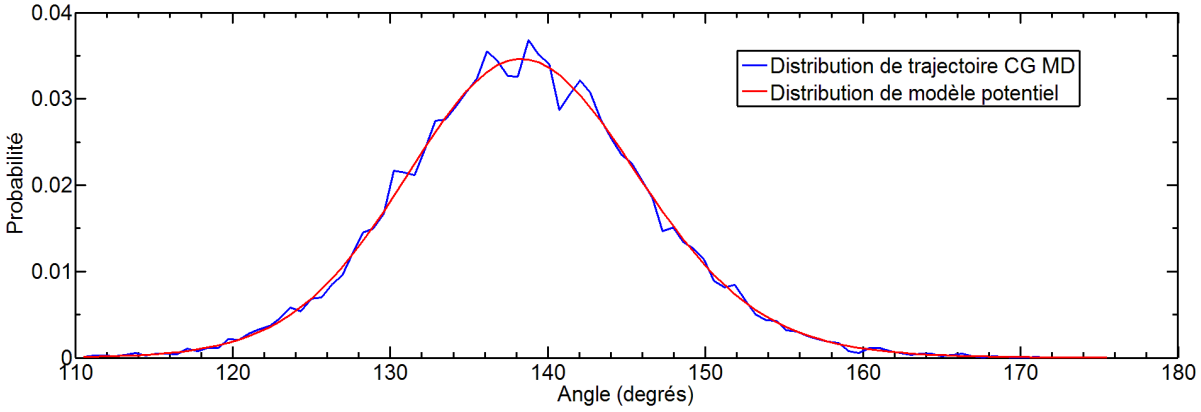

Figure III-15 Comparaison des distributions d'angle provenant du modèle CG, d'une part, et déduite de l'échantillon généré, d'autre part.

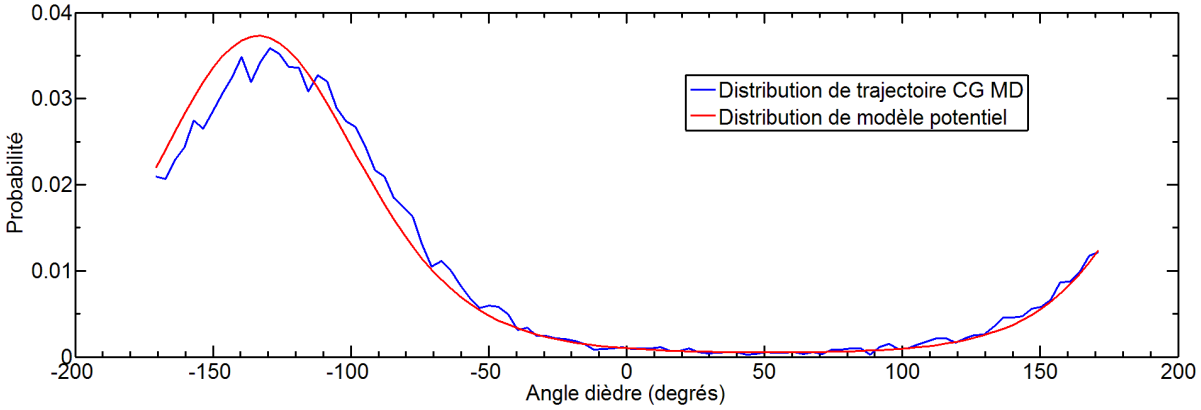

Figure III-16 Comparaison des distributions d'angle dièdre provenant du modèle CG, d'une part, et déduite de l'échantillon généré, d'autre part.

Nous pouvons constater que dans les comparaisons ci-dessus, les distributions issues de l'ensemble des configurations initiales de xylane tirées aléatoirement et celles du modèle CG

du xylane sont en accord. Cela prouve que l'ensemble des configurations initiales sont correctement construites à partir des distributions du modèle CG. La méthode de génération des configurations initiales est donc validée.

LONGUEUR DE PERSISTANCE ET RAYON DE GIRATION

La procédure de tirage aléatoire à l'échelle gros grain permet de générer rapidement de grands échantillons de chaînes de xylane isolées en investissant des temps de calcul très modestes. On peut donc valider cette procédure à partir de ses prédictions à l'échelle globale, c'est-à-dire la longueur de persistance et le rayon de giration.

La longueur de persistance, en abréviation « PL », est un paramètre qui mesure la rigidité d'une chaîne moléculaire, plus sa valeur est grande, plus la chaine est rigide. Elle est définie comme la projection du vecteur bout à bout de la chaîne moléculaire sur le vecteur de la première liaison de la même chaîne (Fig. III-17).Il est évident que le calcul de la PL donnera une valeur croissante avec le DP de la chaîne. Pour les chaînes courtes, les directions des liaisons successives sont fortement corrélées, et la valeur de PL croît avec la longueur de la chaîne test. Au-delà d'une certaine longueur, les liaisons considérées sont décorrelées de la première liaison, et augmenter la longueur de chaîne n'a plus d'influence sur la PL. On peut donc considérer la PL comme la longueur de décorrélation des liaisons. En calculant la PL pour des PL croissants, on s'attend donc à obtenir d'abord des valeurs faibles (on est bien-sûr limité par le nombre de liaisons impliquées) croissantes avec le PL, puis, pour des chaînes longues, on s'attend à observer une saturation de la valeur de PL, due à la décorrélation précédemment évoquée. C'est la valeur de PL obtenue à saturation qui sera retenue comme PL intrinsèque de la chaîne, et comme évaluation de la rigidité intrinsèque de la chaîne.

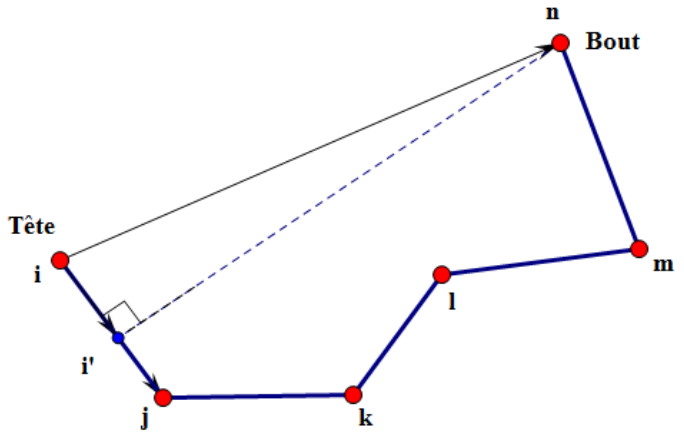

Figure III-17 Définition de la longueur de persistance. Le vecteur entre le premier et le dernier grain de la chaîne est projeté sur le vecteur liant le premier et le deuxième grain.

La méthode de calcul de la PL est différente pour les modèles atomistiques et les modèles CG. Pour tracer l'évolution de PL dans les trajectoires atomistiques, les atomes d'oxygènes de la liaison glycosidique sont utilisés à la place des grains (Fig. III-18).

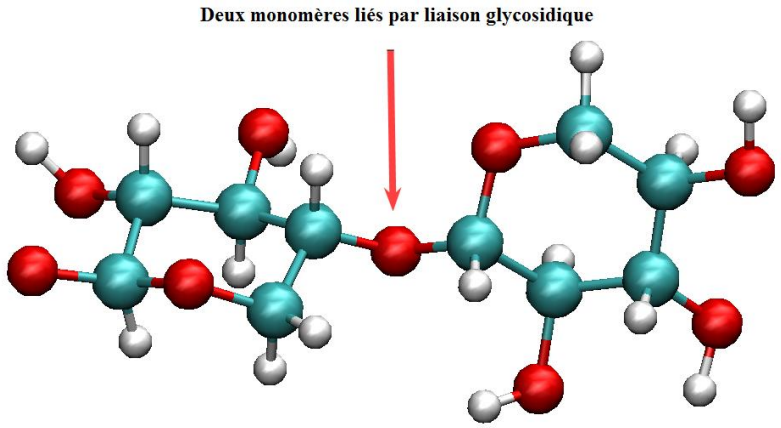

Figure III-18 Deux monomères de xylose sont connectées par l'atome oxygène de la liaison glycosidique.

Dans cette section, nous présentons la longueur de persistance issue de trois modèles :

- ✓ Trajectoires atomistiques du xylane de DP = 25.
- ✓ Version CG des trajectoires du xylane de DP = 25.
- ✓ Echantillon aléatoire de 10000 chaînes CG du xylane de DP = 200 chacune.

Nous avons extrait la LP des trajectoires atomistiques (Fig. III-19) du xylane de DP = 25.

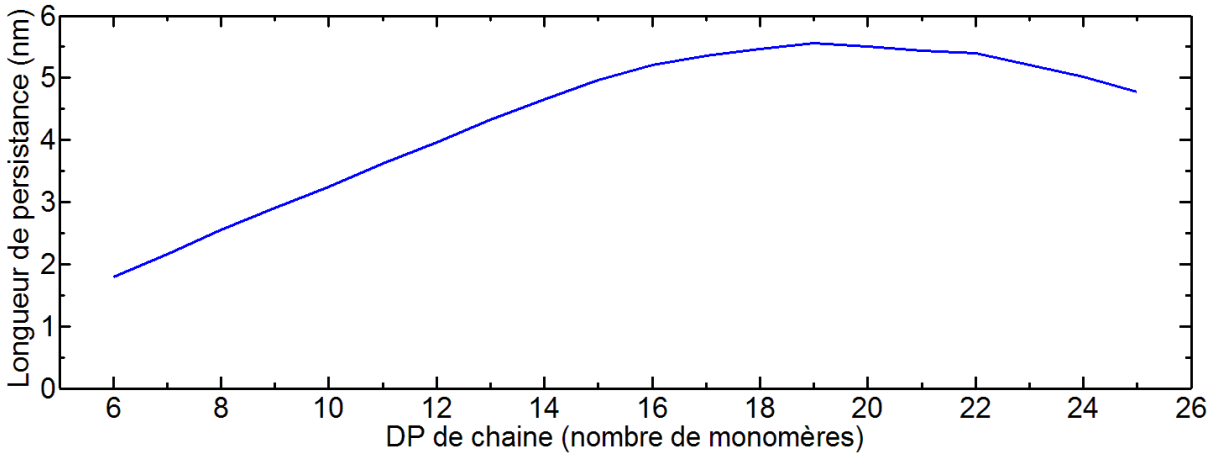

Figure III-19 Evolution de la longueur de persistance en fonction de la longueur de chaîne utilisée pour la trajectoire atomistique de la chaîne de DP = 25. PL est calculé avec les atomes d'oxygènes de la liaison glycosidique.

L'évolution de PL dans le cas DP = 25 montre bien la tendance de saturation autour de 5,5 nm, mais elle diminue rapidement à 4,75 nm après la saturation. Cette décroissance peut être facilement attribuée à un échantillon trop réduit pour les chaines longues.

A partir des configurations initiales construites selon la méthode présentée dans la section précédente, nous calculons l'évolution de PL en fonction du DP (Fig. III-20).

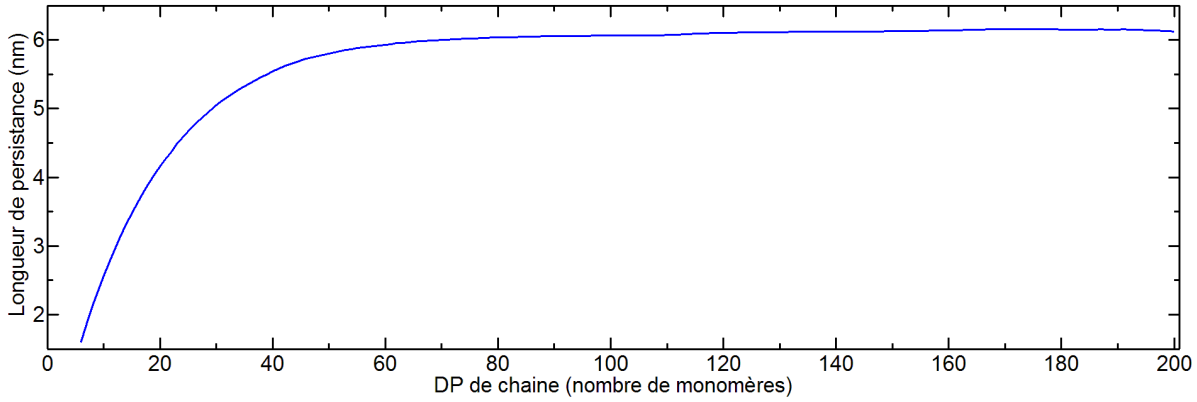

Figure III-20 Evolution de la longueur de persistance en fonction de la longueur de chaîne de xylane utilisée pour le modèle CG du xylane DP = 200.

Dans le cas des chaînes de longueur DP = 200, on constate que la saturation de l'évolution de PL commence à partir de 50 grains. La valeur de LP finale est de 6,25 nm ce qui est en accord avec les données expérimentales d'environ 6,6 ~ 9,2 nm [Dervilly-Pinel *et al.* 2001].

III.1.2 GENERATIONS DES SURFACES ET CRISTAUX DE CELLULOSE

III.1.2.1 Surface de cellulose

La surface de cellulose est construite à partir des données expérimentales [Nishiyama *et al.* 2002]. Cette surface de cellulose est construite à partir des résultats de diffractions des rayons X et de diffusion de neutrons. Les données géométriques de la cellulose sont présentées dans la section I.6.1. Ci-dessous un exemple de la surface 110 de la cellulose. Pour construire cette surface, il suffit d'abord de copier et translater un motif d'un dimère CG de cellulose selon la direction *Z* pour construire une chaîne. Ensuite, copier cette chaîne selon la direction *X* pour former une surface de taille souhaitée (Fig. III-21).

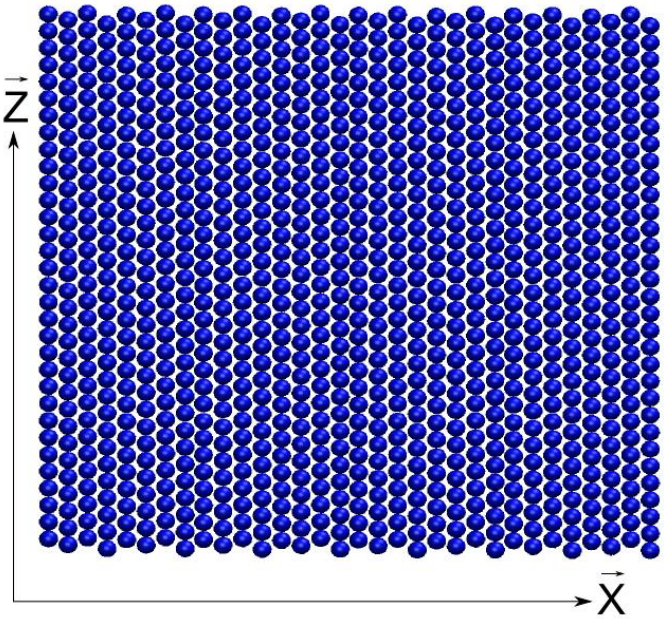

Figure III-21 Une surface 110 de cellulose cristalline *Iβ* comportant 32 chaînes de chaque chaîne de 32 grains. Z est la direction de l'axe de la cellulose, ainsi que la direction de la chaîne de la cellulose. Données géométrique provenant des expériences de diffraction des rayons X et de diffusion de neutrons [Nishiyama *et al.* 2002]. Taille de la surface : 15,92 nm selon la direction *Z* et 16,31 nm selon la direction *X*.

III.1.2.2 Cristal de cellulose

La méthode de construction des cristaux de la cellulose est similaire à celle de la surface de la cellulose. La construction du cristal de cellulose repose sur les paramètres cristallographiques de Nishiyama [Nishiyama *et la* 2002].

En l'absence de consensus au sujet de la forme du cristal de cellulose, des choix et des hypothèses doivent être faits. Les surfaces exposées sont, majoritairement les surfaces 110 et 1-10, très peu 100 [Bergenstråhle *et al.* 2008, Larsson *et al.* 2004, Larsson *et al.* 1999]. La surface 100 est limitée à 3 chaînes dans les microfibrilles composées de 20 à 30 chaînes. Pour l'instant, aucune évidence expérimentale ne peut prouver l'existence de la surface 010, elle devrait être minoritaire vis-à-vis de la surface 100 dans la morphologie externe si elle existe [Mazeau 2011]. De ce fait, les surfaces du cristal peuvent être au minimum réduites aux trois qui sont déjà mentionnées dans la littérature : 110, 100 et 1-10 (Fig. III-22, III-23).

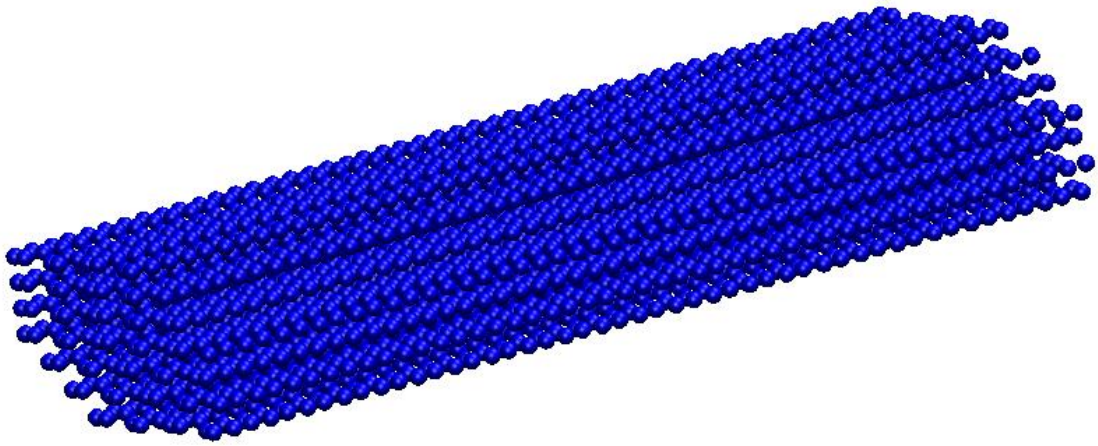

Figure III-22 Un cristal de cellulose de 65 chaînes de 48 grains chacune. Les surfaces exposées sont 110, 100 et 1-10. La longueur du cristal de la cellulose est de 24,47 nm.

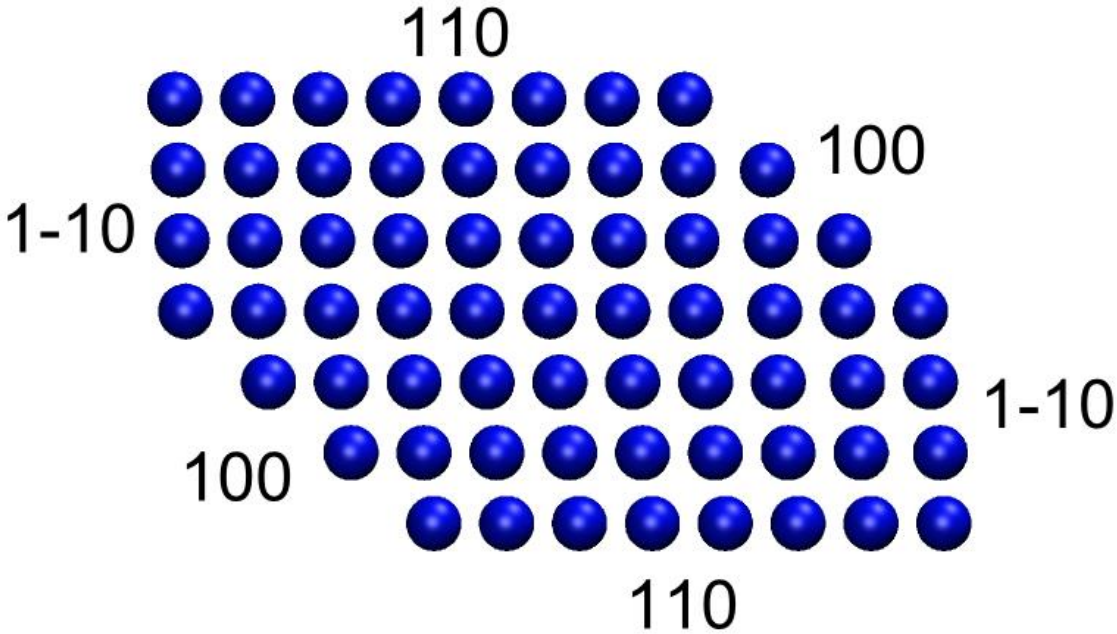

Figure III-23 Vue en coupe du cristal de cellulose. Pour chacune des surfaces exposées, le type de paramétrage pour les interactions intermoléculaires cellulose-xylane est indiqué : 110, 1-10 et 100.

III.1.3 GENERATIONS DES SYSTEMES XYLANE-CELLULOSE

Le point important pour générer un système moléculaire xylane-cellulose est la stratégie de disposition spatiale des chaînes de xylane. On s'efforcera d'éviter que les distances initiales entre grains prennent des valeurs anormalement faibles, ce qui déstabiliserait le calcul (Fig. III-24). D'autre part, il s'agira d'éviter que les chaines ne traversent les parois de la boîte de simulation, dans la mesure où le code de calcul utilisé ne gère pas cette situation à l'état initial (c'est évidemment le cas au cours du calcul).

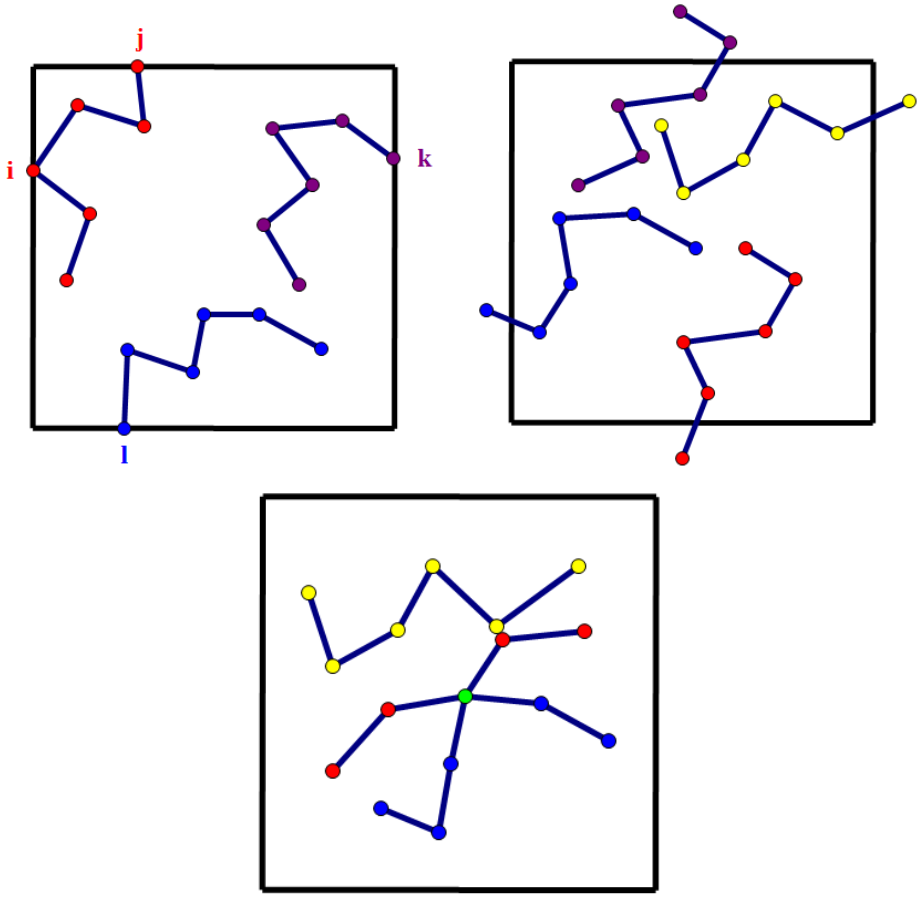

Figure III-24 Situations à proscrire lors de la construction de l'état initial. Figure en haut à gauche : les grains *i*, *j*, *k* et *l* sont très proche du bord de la boîte de simulation, *i* et *k* peuvent être superposés ou très proche en conditions aux limites périodiques, ainsi que *j* et *l*. Figure en haut à droite : des grains en dehors de la boîte de simulation. Figure en bas : des grains sont très proches, indépendamment des conditions aux limites périodiques.

III.1.3.1 Xylane et surface de cellulose

Le système moléculaire xylane-cellulose est construit pour l'étude de l'interaction entre le xylane et une surface spécifique de cellulose cristalline. Dans l'optique de la construction, on définit tout d'abord une boîte de simulation de la taille souhaitée. Les PBC sont appliquées sur les deux directions X et Z. La surface de cellulose est placée dans le plan $Y = 0$. Si le nombre de chaînes de xylane est modéré, les chaînes peuvent être aisément déployées verticalement au-dessus de la surface (Fig. III-25).

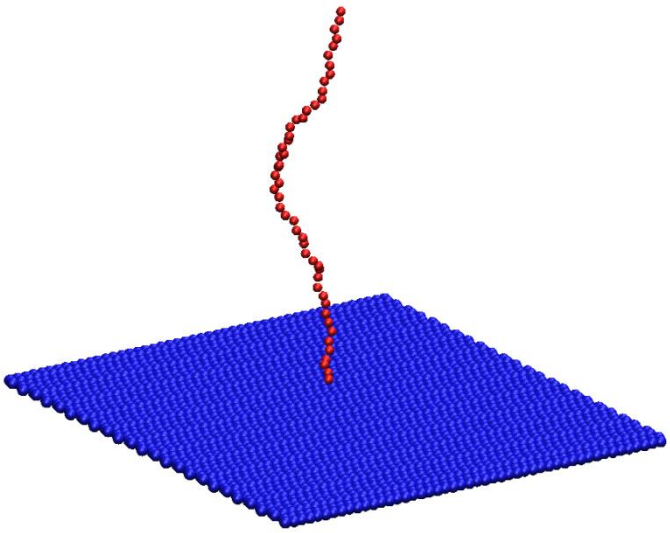

Figure III-25 Configuration initiale d'un système moléculaire constitué d'une chaîne de xylane de DP = 50 et d'une surface de cellulose de 48 chaînes de 48 grains chacune.

Lorsque de nombreuses chaînes sont impliquées, une stratégie doit être mise en œuvre afin de distribuer le mieux possible les chaînes de xylane sur la surface de la cellulose. Ci-dessous trois point spécifiques pour ce type de construction.

a. L'attribution des positions des chaînes commence par le centre géométrique de la surface et continue ver les directions *Z* et –*Z*, une fois que la première ligne est saturée, les chaînes sont déployées à partir des centres sur les lignes *X* et –*X*, et vers les deux directions *Z* et –*Z* (Fig. III-26).

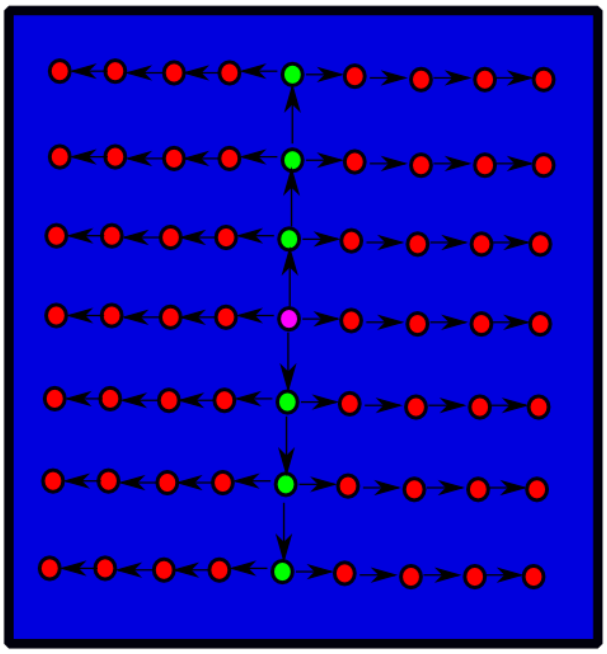

Figure III-26 Ordre de déploiement des chaînes de xylane sur la surface de la cellulose. Le point rose est la position de la première chaîne de xylane déployée, les flèches indiquent les directions de déploiement pour le reste de chaînes, les points verts et rouges sont les positions initiales pour le reste des chaînes à générer.

b. Les chaînes déployées sont ensuite tournées jusqu'à ce que leurs vecteurs bout-à-bout soient perpendiculaires à la surface de la cellulose (Fig. III-27). Cette

opération permet d'organiser systématiquement les orientations globales des chaînes de xylane.

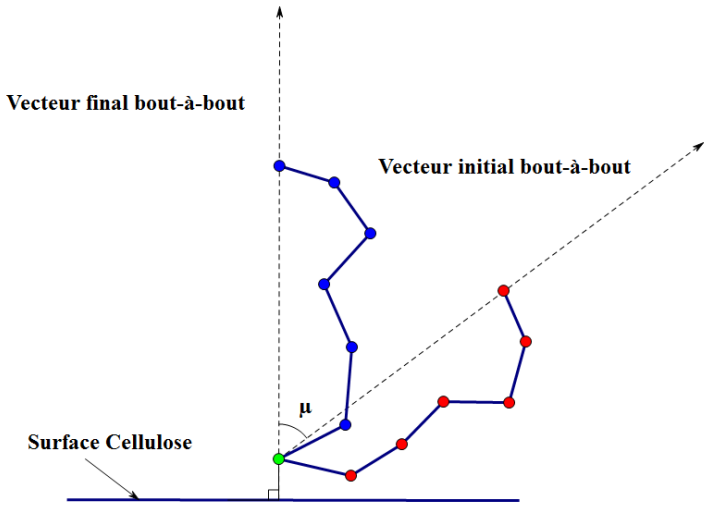

Figure III-27 Rotation d'une chaîne de xylane de façon à ce que le vecteur bout-à-bout de la chaîne soit perpendiculaire à la surface de cellulose.

c. Pour éviter la superposition des grains de xylane avec les grains de cellulose, les chaînes de xylane sont placées à une distance minimale au-dessus de la surface de la cellulose. La superposition entres les grains du xylane est évitée par deux étapes : la première étape consiste à mesurer la distance minimale entre la chaîne de xylane à placer et les chaînes déjà placées. La position de la chaîne à placer sera acceptée si cette distance minimale est supérieure au seuil qui a été fixé à 0,6 nm. Sinon, la chaîne à placer sera translatée légèrement et répétitivement de 0,05 nm dans la direction *Y* perpendiculaire à la surface de la cellulose (Fig. III-28). Cette translation répétitive s'arrête jusqu'à la distance minimale soit supérieure ou égale au seuil.

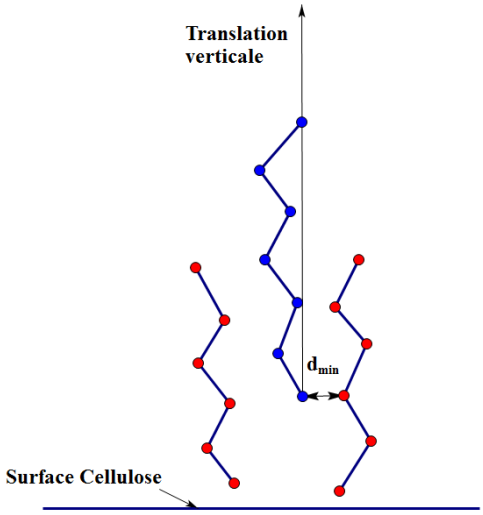

Figure III-28 Translation verticale de la chaîne de xylane à placer, la chaîne bleue, dans la direction *Y* pour éviter la superposition avec les grains de xylane déjà placés.

En respectant les trois points décrits, nous pouvons générer des configurations initiales correctes. La figure III-29 donne un exemple d'une configuration initiale générée à partir de la

stratégie décrite. On peut noter que cette stratégie n'est applicable que dans la mesure où les chaines de xylane sont suffisamment rigides, comme l'ont montré les évaluations de leur longueur de persistance.

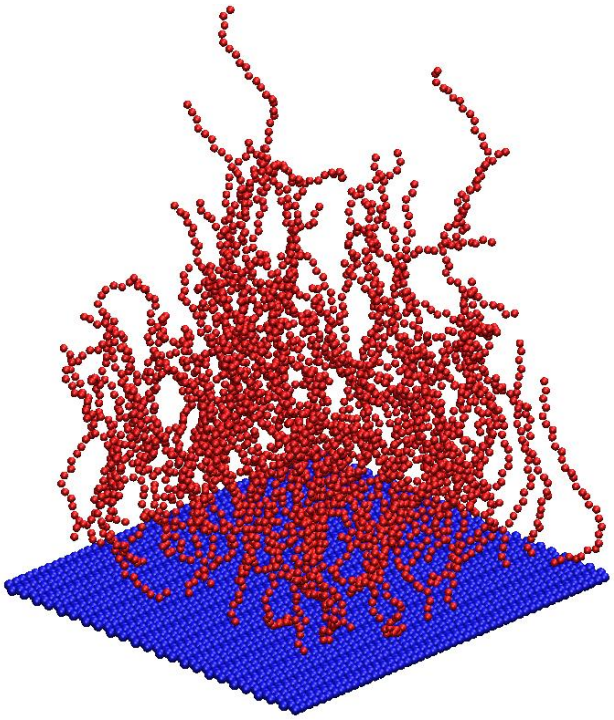

Figure III-29 Configuration initiale d'un système moléculaire de 120 chaînes de xylane de DP = 30 et une surface de cellulose de 48 chaînes de 48 grains chacune.

La stratégie choisie pour cette génération de configuration initiale implique des difficultés. La procédure de translation selon Y visant à éviter les superpositions xylane-xylane implique qu'un certain nombre de chaines de xylane soient assez éloignées de la surface, ce qui implique un risque que celles-ci n'interagissent pas suffisamment avec le reste du système et s'éloignent de la surface. Cet effet peut être renforcé par le fait que le potentiel de Morse présente une extinction trop rapide à longue distance. Pour remédier au problème, une paroi répulsive est introduite dans la boîte de simulation dans la première phase de la simulation, de façon à forcer les chaînes de xylane à rester proches du composite xylane-cellulose (Fig. III-30). La force de la paroi s'applique sur les grains dépassant au-dessus de la paroi. Cette force est décrite par une force harmonique donnée comme :

$$\vec{F} = k \cdot \left(\overrightarrow{r_{paroi}} - \overrightarrow{r_{grains}}\right) \qquad \text{Équation III-1}$$

Où k = 20 $pN*nm^{-1}$ est l'intensité de force, $\overrightarrow{r_{paroi}}$ est la position de la paroi, $\overrightarrow{r_{grains}}$ les positions des grains traversant au-dessus de la paroi.

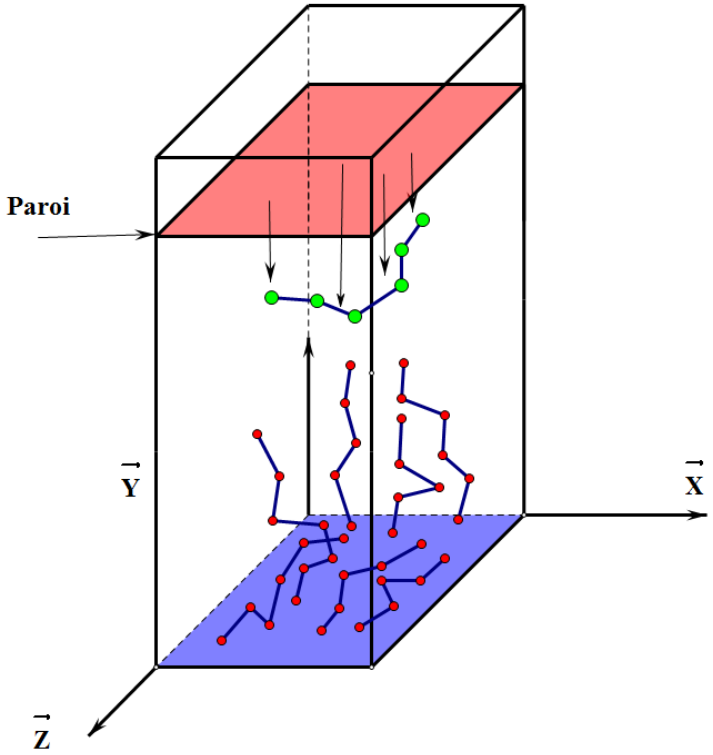

Figure III-30 Confinement des chaines de xylane au début de la simulation. La surface bleue est la surface de cellulose, la surface rouge est la paroi répulsive.

III.1.3.2 Xylane et monocristal de cellulose

La génération des configurations initiales du composite xylane-cristal de cellulose est une variante de la génération des configurations initiales du composite xylane-surface de cellulose car il s'agit de distribuer les chaînes du xylanes sur les différentes surfaces de la cellulose. Le cas le plus simple consiste à disposer aléatoirement une chaîne du xylane autour du cristal de cellulose (Fig. III-31).

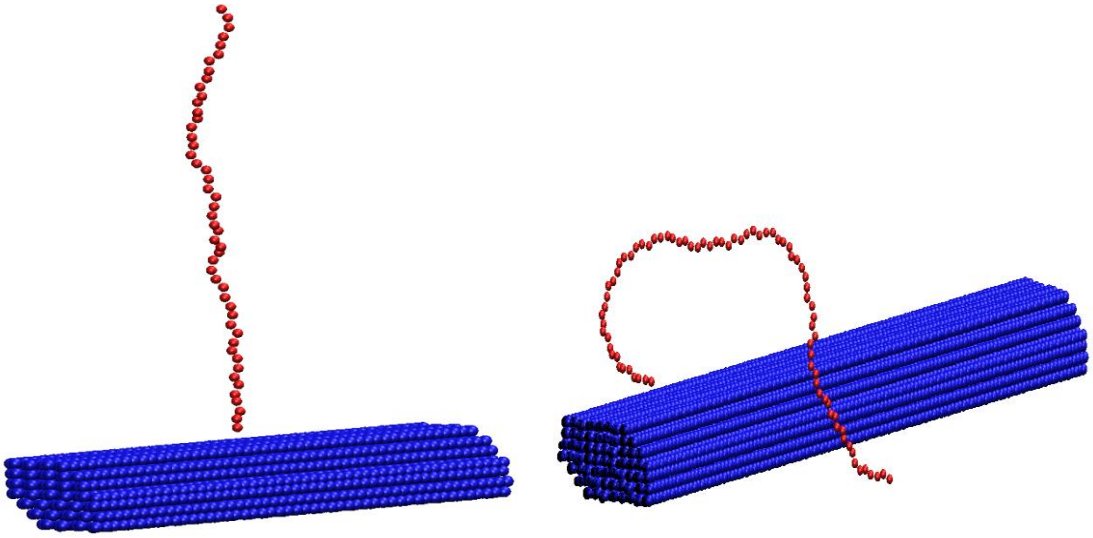

Figure III-31 Configuration initiale d'un système moléculaire constitué d'une chaîne de xylane de DP = 50 et d'un cristal de la cellulose de 65 chaînes de 48 grains chacune. A gauche : la chaîne de xylane est perpendiculaire à la surface 110 du cristal. A droite : chaîne de xylane dont l'orientation est aléatoire.

Il est bien-sûr plus compliqué de déposer de nombreuses chaines de xylane autour du cristal de cellulose.

Premièrement, il est important de noter que la boîte de simulation pour le système xylane-cristal de cellulose est différente que celle du système xylane-surface de cellulose. Les PBC sont appliquées sur la direction *Z* qui est l'axe du cristal, la périodicité étant la longueur du cristal. Les mêmes conditions sont appliquées dans les directions *X* et *Y* mais la taille de la boîte dans ces directions a été choisie très grande. En pratique il n'existe donc pas de conditions aux limites périodiques selon *X* et *Y*, et tout se passe comme si la boîte était infinie dans ces directions.

Le cristal est orienté de telle façon que la surface 1-10 soit dans le plan *YZ* et la surface 110 soit dans le plan *XZ* (Fig. III-32).

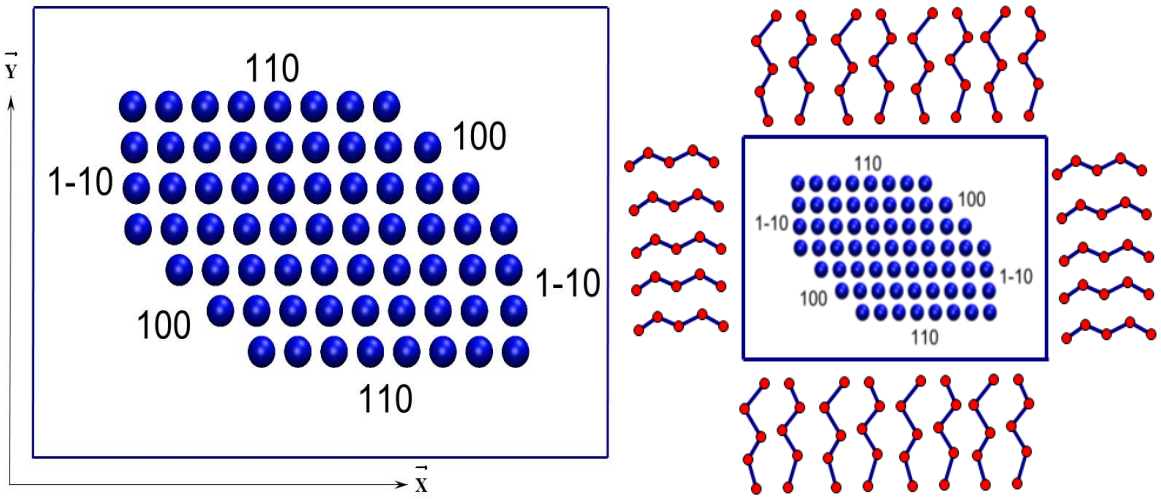

Figure III-32 Disposition initiale du système. A gauche : orientation du cristal par rapport aux axes. A droite : disposition initiale des chaînes de xylane.

Ensuite, le déploiement des chaînes de xylane suit une procédure similaire à celle qui a été utilisée pour la construction du système xylane-surface de cellulose. Les chaînes sont tournées jusqu'à ce que leurs vecteurs bout-à-bout soient perpendiculaires à l'axe du cristal *Z*. Puis, elles sont disposées selon les directions *X* et *Y* (Fig. III-33).

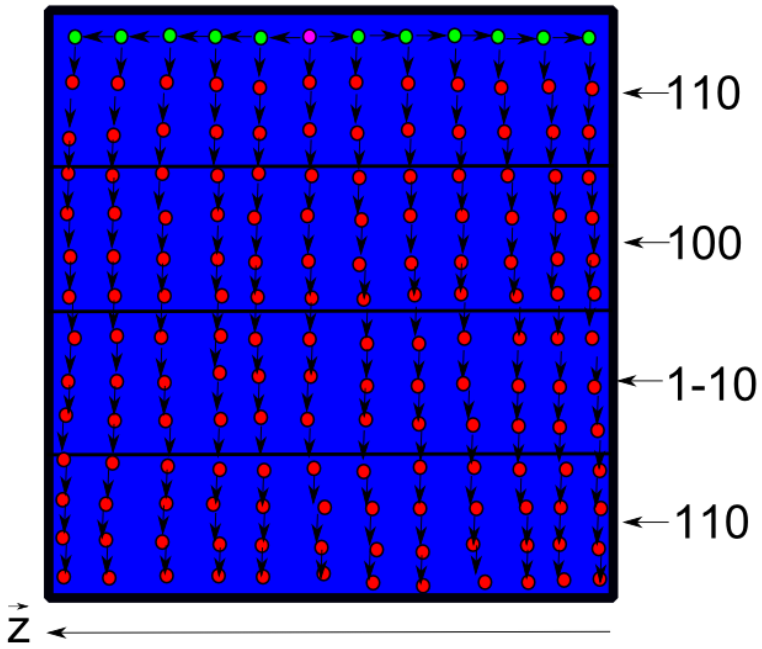

Figure III-33 Ordre de déploiement des chaînes de xylane sur les surfaces 110, 1-10 et 100 du cristal de cellulose. Le point rose est la position de la première chaîne du xylane déployée, les flèches indiquent les directions de déploiement pour le reste de chaînes, les points verts et rouges sont les positions initiales pour le reste des chaînes à générer.

Pour remédier au problème de la superposition des grains, les chaînes superposées sont translatées par pas successifs de façon à les éloigner du cristal et jusqu'à ce que la distance minimale entre les chaînes translatées et les chaines déjà en place soit supérieure à 0,5 nm. Nous exposons dans (Fig. III-34) un état initial d'un monocristal entouré par xylane.

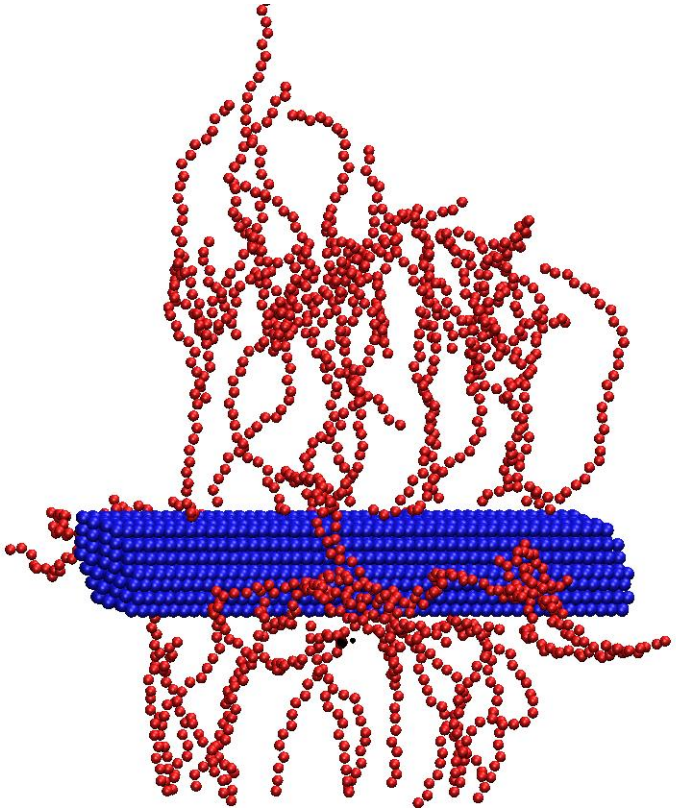

Figure III-34 Configuration initiale d'un système moléculaire de 50 chaînes de xylane de DP = 30 et un cristal de cellulose de 65 chaînes de 48 grains chacune.

Cette procédure présente le même inconvénient que dans le cas des systèmes xylane-surface de cellulose : il existe un risque que certaines chaines n'interagissent pas suffisamment avec le composite cellulose-xylane et s'éloignent de la zone d'intérêt. Des parois répulsives sont alors introduites dans la première phase des simulations, de façon à confiner les chaines de xylane au voisinage du cristal de cellulose (Fig. III-35).

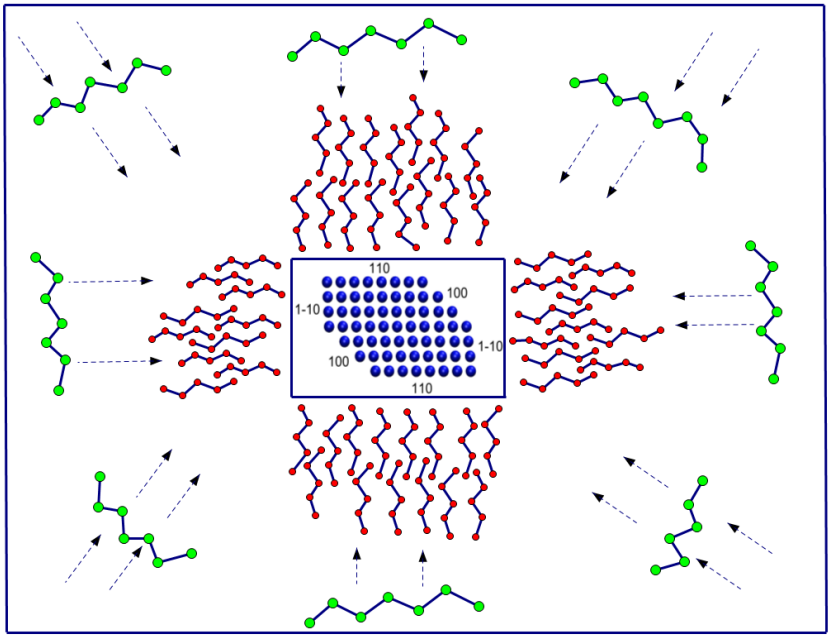

Figure III-35 Les parois répulsives (cadre bleu entourant le système) confinent les molécules éloignées du cristal (indiquées en vert).

Dans la section III.1, la génération des états initiaux des chaines de xylane, des cristaux de cellulose et du composite cellulose-xylane ont été présentés. Les états initiaux des chaines du xylane ont été validés par la comparaison des distributions statistiques avec les modèles mathématiques établis dans le chapitre précédent. Les cristaux de cellulose ont été construits à partir des données expérimentales en considérant trois types de surface dominants : 110, 1-10 et 100.

Le but de cette section est pour la préparation des applications au sujet de la dynamique du composite cellulose-xylane dans les sections suivantes.

III.2 Interaction d'une chaine de xylane avec une surface de cellulose

III.2.1 ADSORPTION D'UNE CHAINE DE XYLANE SUR UNE SURFACE DE CELLULOSE

L'interaction entre une chaîne d'hémicellulose et une surface de cellulose a été déjà décrite à l'échelle atomistique [Karim and Charlier 2012, Zhang *et al.* 2011, Hanus and Mazeau 2006]. Cependant, les longueurs de chaine de xylane dans ces modèles atomistiques sont limitées par le grand nombre de degrés de liberté du système moléculaire. Le DP maximal des chaînes d'hémicellulose est de 15 [Zhang *et al.* 2011] et plus précisément 5 pour le xylane [Karim and Charlier 2012]. Afin de surmonter les limites imposées par le poids numérique des modèles atomistiques, nous proposons dans cette section un modèle CG composé d'une chaîne de xylane de DP = 36 et une surface de cellulose de 40 chaînes de DP = 40 dans l'ensemble *NVT*.

III.2.1.1 Détails de la simulation

Dans le modèle proposé, la surface de la cellulose est placée en $Y = 0$, la chaîne de xylane est générée aléatoirement et équilibré pendant quelques nanosecondes, elle est ensuite placée perpendiculairement à la surface (Fig. III-36).

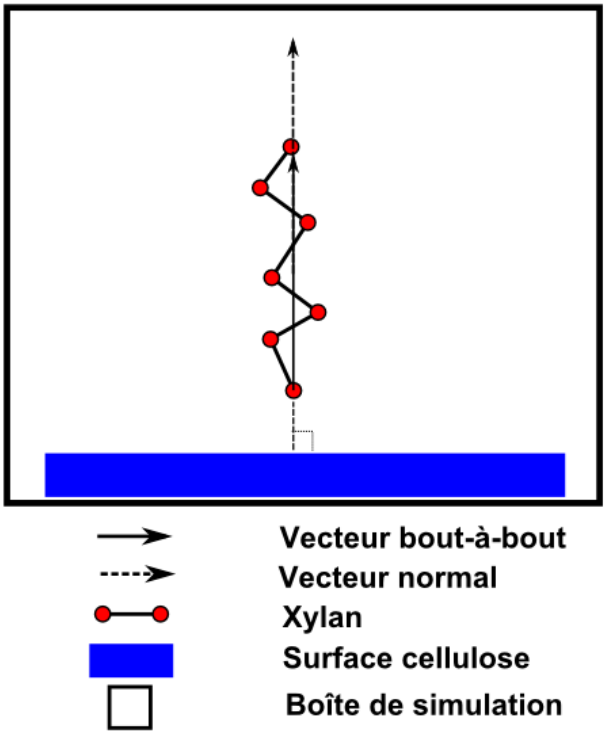

Figure III-36 Etat initial de la chaîne de xylane au voisinage de la surface de cellulose. Le vecteur bout-à-bout de la chaine de xylane est perpendiculaire à la surface de cellulose.

Ci-dessous les paramètres de la simulation.

Type	Température	Temps total	Pas de temps	Coefficient de Mass Q	Temps de simulation
Xylane/Cellulose surface	**300K**	**50 ns**	**25 fs**	**1,0**	**~2662 s**

Tableau III-1 Paramètres des simulations d'adsorption d'une chaîne de xylane sur une surface de cellulose.

La simulation est répétée 6 fois avec différentes configurations initiales pour chaque type de surface de cellulose avec les mêmes paramètres. Au total, 18 simulations ont été réalisées. L'intensité du potentiel xylane-cellulose pour chaque type de surface est ajustée selon les ratios de potentiel issus de calculs atomistiques et fournis par Karim Mazeau. Les coefficients pour les surfaces 110, 100 et 1-10 sont respectivement 1,0, 1,303 et 1,221.

III.2.1.2 Résultats

CONFORMATION DE LA CHAINE

Dans les trajectoires obtenues pour les trois types de surface de cellulose, nous constatons que la chaîne de xylane, placée perpendiculairement au-dessus de la surface, est adsorbée rapidement sur la surface de cellulose et maintenue adsorbée jusqu'à la fin de la simulation. La chaîne peut être rectiligne, entièrement alignée sur le réseau cristallin de la cellulose, ou présenter des coudes (Fig. III-37).

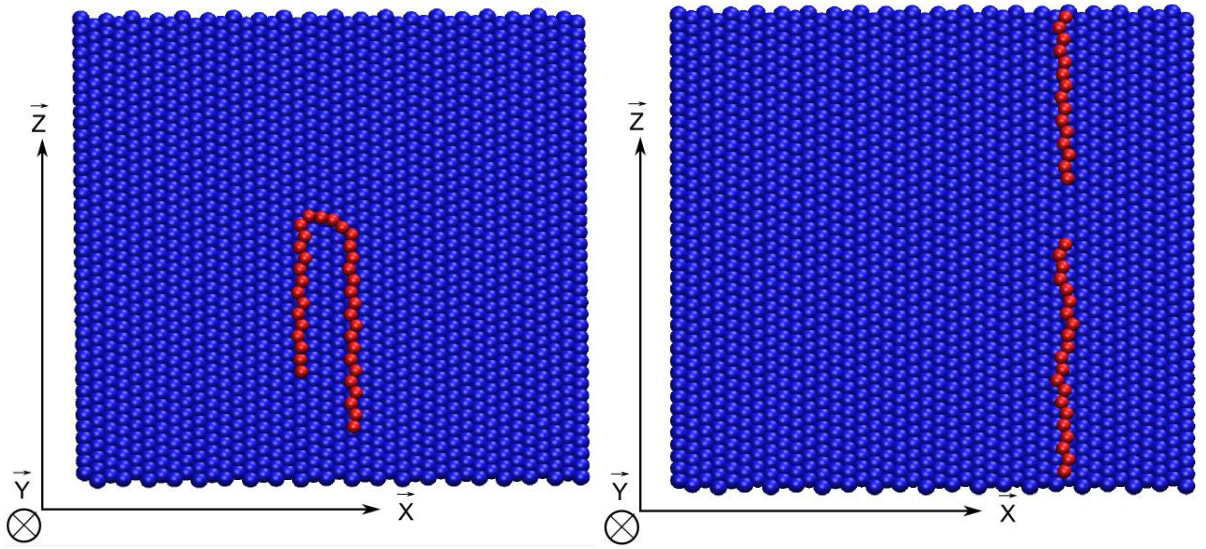

Figure III-37 Etat final de l'adsorption d'une chaîne de xylane de DP = 36 sur la surface (110) de la cellulose de 40 chaînes de 40 monomères. Les sphères rouges représentent les grains du xylane, les sphères bleues représentent les grains de la cellulose. A gauche : la chaîne du xylane forme des coudes. A droite : la chaîne de xylane s'aligne selon l'axe des chaines de cellulose.

ENERGIES

L'évolution de l'énergie d'interaction au cours de la simulation est similaire pour toutes les trajectoires générées. Le cas d'une des trajectoires est présenté en figure III-38.

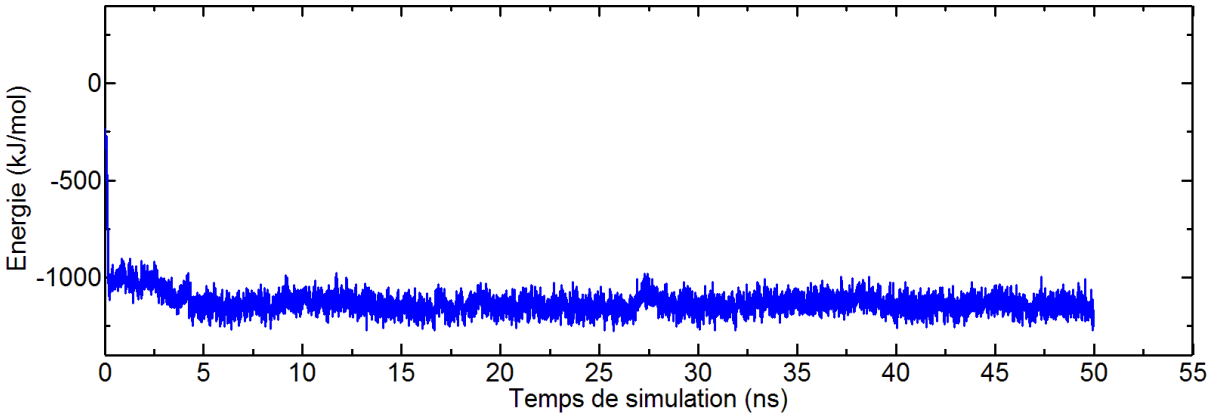

Figure III-38 Evolution temporelle de l'énergie potentielle.

L'énergie cinétique est toujours conservée pendant toute la durée de la simulation, l'ensemble statistique choisi étant l'ensemble *NVT*. L'énergie d'interaction xylane-cellulose décroit rapidement pendant l'adsorption du xylane, jusqu'à atteindre rapidement à une valeur proche de -1157 *kJ/mol*. Toutes les simulations réalisées prédisent le même type d'évolution de l'énergie, ainsi que des valeurs d'énergie d'interaction à l'équilibre très proches, et ce, pour chaque type de surface de cellulose. Les énergies d'adsorption obtenues pour les six simulations et pour les trois surfaces de cellulose sont rassemblées dans le tableau ci-dessous :

Energie Absorption (kJ/monomère*mol)	S1	S2	S3	S4	S5	S6	Energie Moyen
Surface 110	**23.6491**	**24.8173**	**24.7346**	**25.7744**	**24.6643**	**25.5159**	**24.8593**
Surface 100	**35.2903**	**36.3951**	**35.2287**	**35.6136**	**32.8841**	**34.1443**	**34.9260**
Surface 1-10	**32.5841**	**33.7529**	**32.2468**	**35.0863**	**30.1163**	**31.1678**	**32.4924**

Tableau III-2Energie d'absorption pour les trois surfaces (110), (100) et (1-10) de cellulose.

Les énergies d'adsorption obtenues pour les simulations de la surface 110 sont proches de celles obtenues par dynamique moléculaire atomistique [Mazeau and Charlier 2012] c'est-à-dire une énergie comprise entre 25,104 et 33,472 *kJ/monomère*mol*. Cet intervalle est également confirmé par AFM [Morris *et al.* 2004].

CONFORMATIONS DU XYLANE SUR LA SURFACE : ANALYSE QUANTITATIVE

La distance bout-à-bout

Un objectif essentiel des simulations présentées ici est de prédire la conformation du xylane sur la surface de cellulose. Cette conformation sera ici caractérisée quantitativement par deux paramètres : la distance bout-à-bout de la chaîne du xylane et les projections du vecteur bout-à-bout sur l'axe de la cellulose (Fig.III-39). Les valeurs obtenues permettront une comparaison avec les calculs atomistiques.

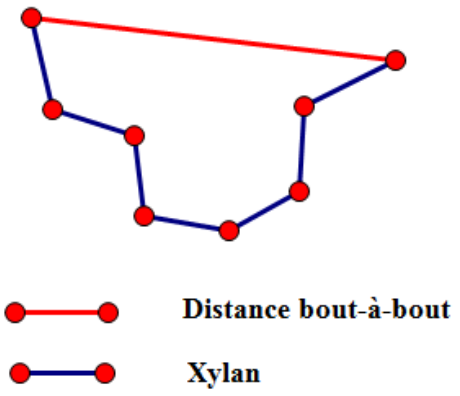

Figure III-39 La distance bout-à-bout de la chaîne de xylane.

La distance bout-à-bout de la chaîne de xylane permet d'évaluer son état d'extension lorsqu'elle est adsorbée sur la surface de cellulose. Seront présentées ci-après l'évolution temporelle de cette valeur au cours des six simulations (Fig. III-40), ainsi que la distribution de probabilité, et ce, pour chacune des surfaces de cellulose.

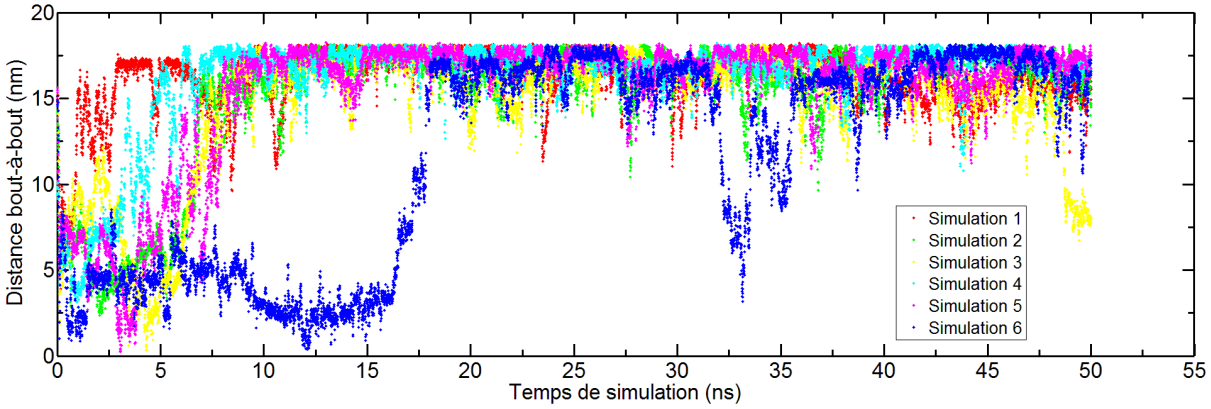

Figure III-40 Evolution temporelle de la distance bout-à-bout de la chaîne de xylane sur la surface 110 de cellulose.

Dans le cas de la surface 110, la distance bout-à-bout varie fortement durant la phase d'adsorption du xylane, au début de chaque simulation. Puis, après un régime transitoire d'une durée comprise entre 5 ns et 10 ns, la chaîne de xylane est majoritairement étendue lorsqu'elle est adsorbée sur la surface de cellulose.

La distribution de probabilité confirme cette observation, puisque les valeurs de la distance bout-à-bout sont majoritairement supérieures à 16 nm, la chaîne est donc presque systématiquement complètement étendue (Fig. III-41).

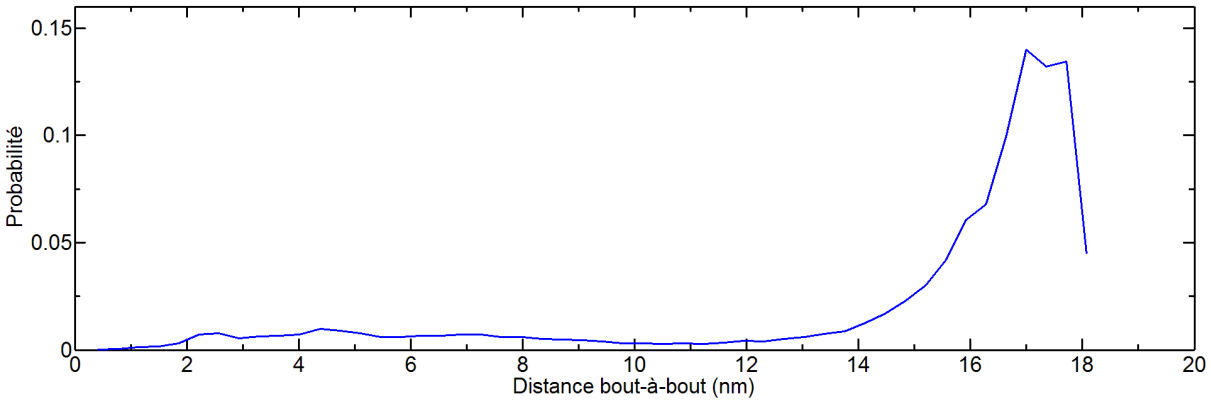

Figure III-41 Distribution de probabilité de la distance bout-à-bout de la chaîne de xylane sur la surface 110 de cellulose.

Cependant, cette extension n'est pas observée pour tous les types de surface. Pour les simulations réalisées avec la surface 100, la chaîne n'est pas systématiquement étendue (Fig. III-42), puisque ce n'est le cas que pour deux simulations sur les six qui ont été réalisées.

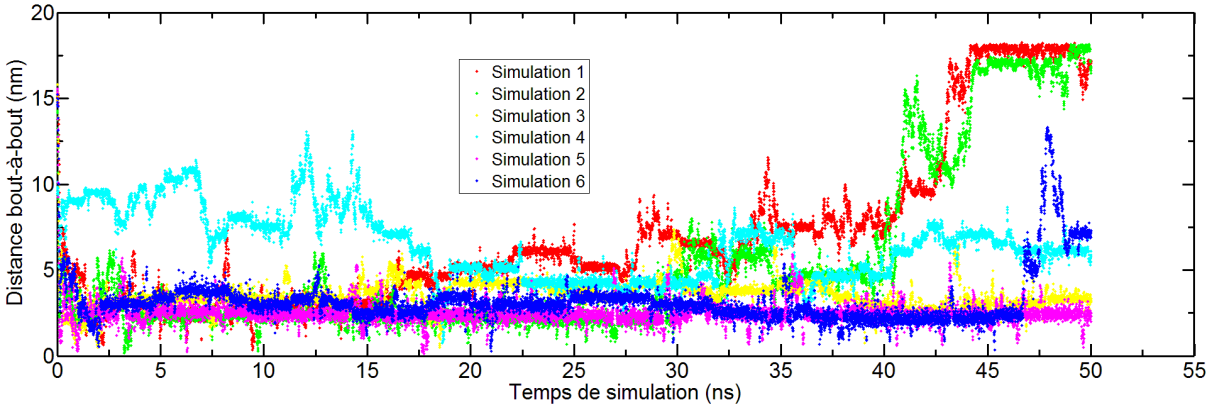

Figure III-42 Evolution temporelle de la longueur bout-à-bout de la chaîne de xylane sur la surface 100 de la cellulose.

Cette constatation est confirmée par la distribution de probabilité (Fig. III-43). La distance bout-à-bout prédite par les simulations est majoritairement basse et la conformation étendue est peu probable.

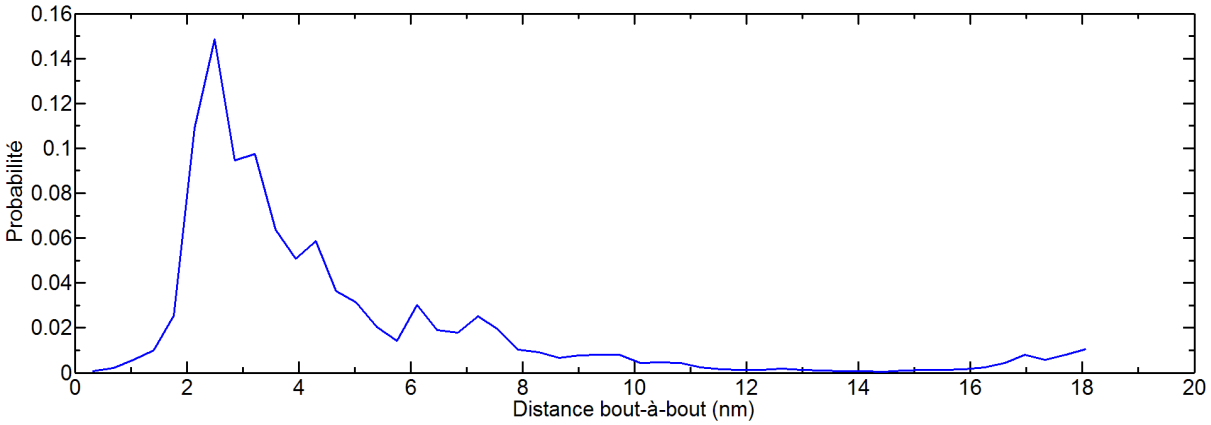

Figure III-43 Distribution de la distance bout-à-bout de la chaîne de xylane sur la surface 100 de la cellulose.

Dans le cas de la surface 1-10, la distance bout-à-bout atteint sa valeur maximale pour quatre simulations sur six, mais après un régime transitoire bien plus long que dans le cas de la surface 110 (Fig. III-44). Pour les autres simulations, la distance bout-à-bout conserve une valeur basse.

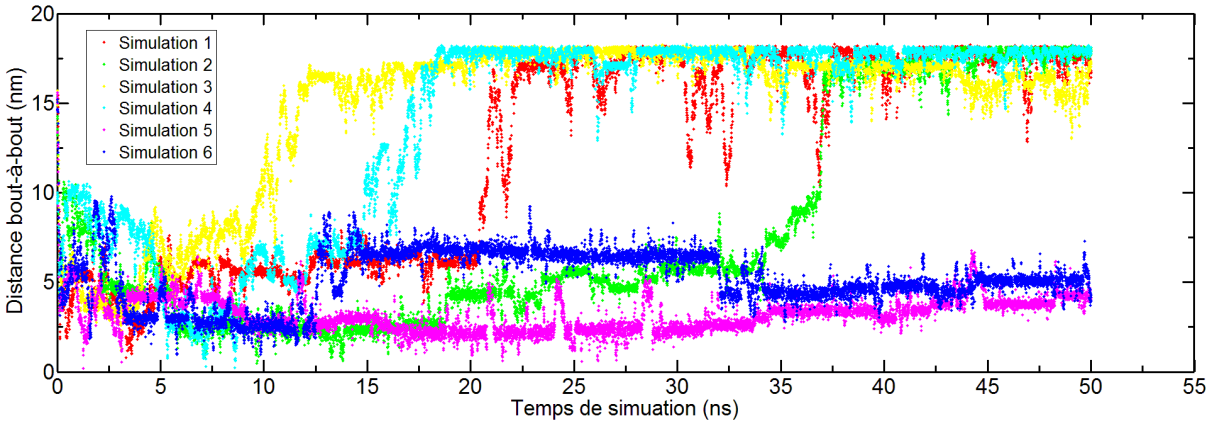

Figure III-44 Evolution temporelle de la distance bout-à-bout de la chaîne de xylane pour la surface 1-10 de cellulose.

En cohérence avec ces constatations, la distribution de probabilité de la distance bout-à-bout est bimodale, la conformation du xylane restant tout de même majoritairement coudée (Fig. III-45).

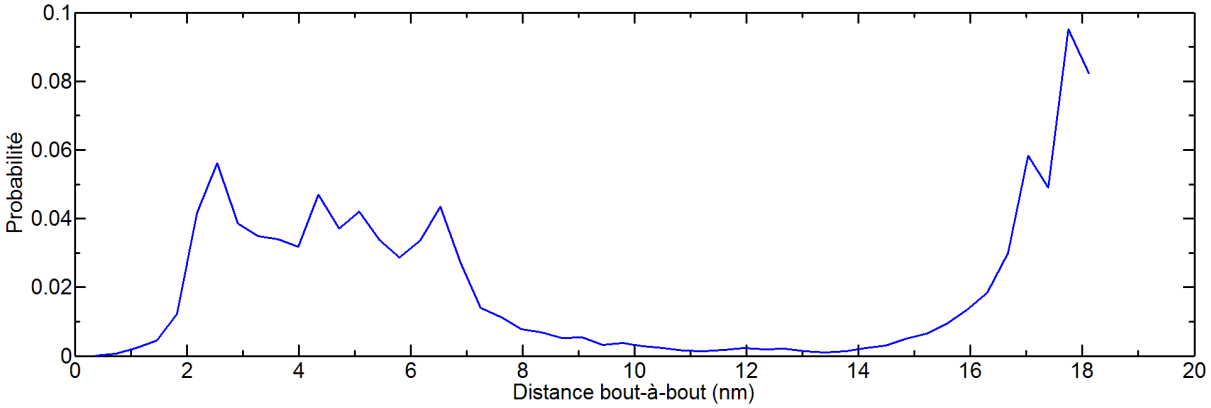

Figure III-45 Distribution de la distance bout-à-bout de la chaîne de xylane sur la surface 1-10 de la cellulose.

Au vu des cinétiques de la distance bout-à-bout, la surreprésentation des conformations non étendue est majoritairement liée au fait que les chaines mettent plus de temps à atteindre une conformation étendue dans le cas des surfaces 100 et 1-10. Un état d'équilibre stable n'est peut-être pas atteint à l'issue des simulations numériques présentées ici. On peut illustrer cette hypothèse à travers un exemple d'évolution temporelle de l'énergie (Fig. III-46).

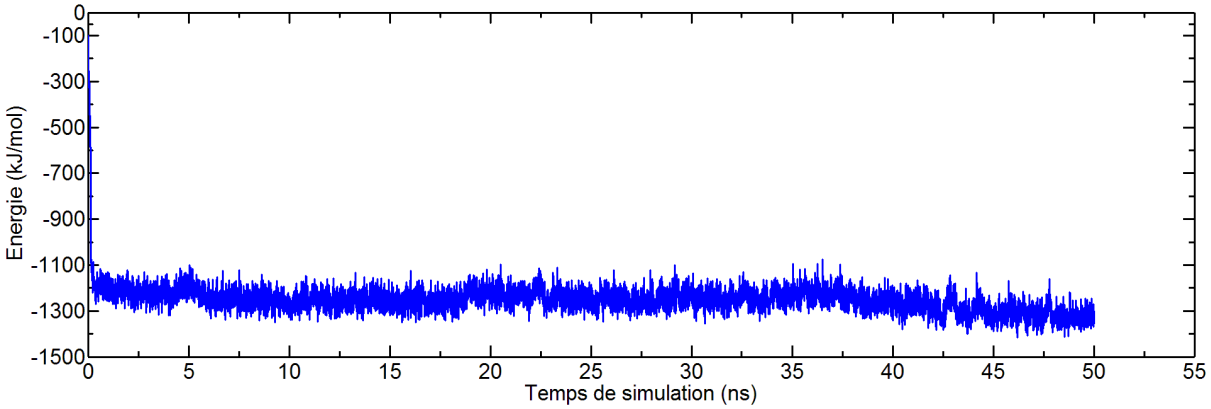

Figure III-46 Evolution d'énergie totale de la simulation n°=2 dans le cas de la surface 1-10.

Si le xylane s'adsorbe rapidement sur la cellulose, la conformation de la chaine confinée sur la surface semble obéir à une cinétique plus lente, dans la mesure où l'énergie continue d'évoluer au-delà de 5 ns. La mise en œuvre de nouvelles simulations, sur des durées plus longues, sera nécessaire pour éclaircir ce point.

Les projections du vecteur bout-à-bout

La projection du vecteur bout-à-bout sur la direction *Z* permet d'évaluer la conformation de la chaîne du xylane par rapport à la surface (Fig. III-47). On rappelle que la direction *Y* est perpendiculaire à la surface, *Z* est la direction de l'axe des chaînes de cellulose.

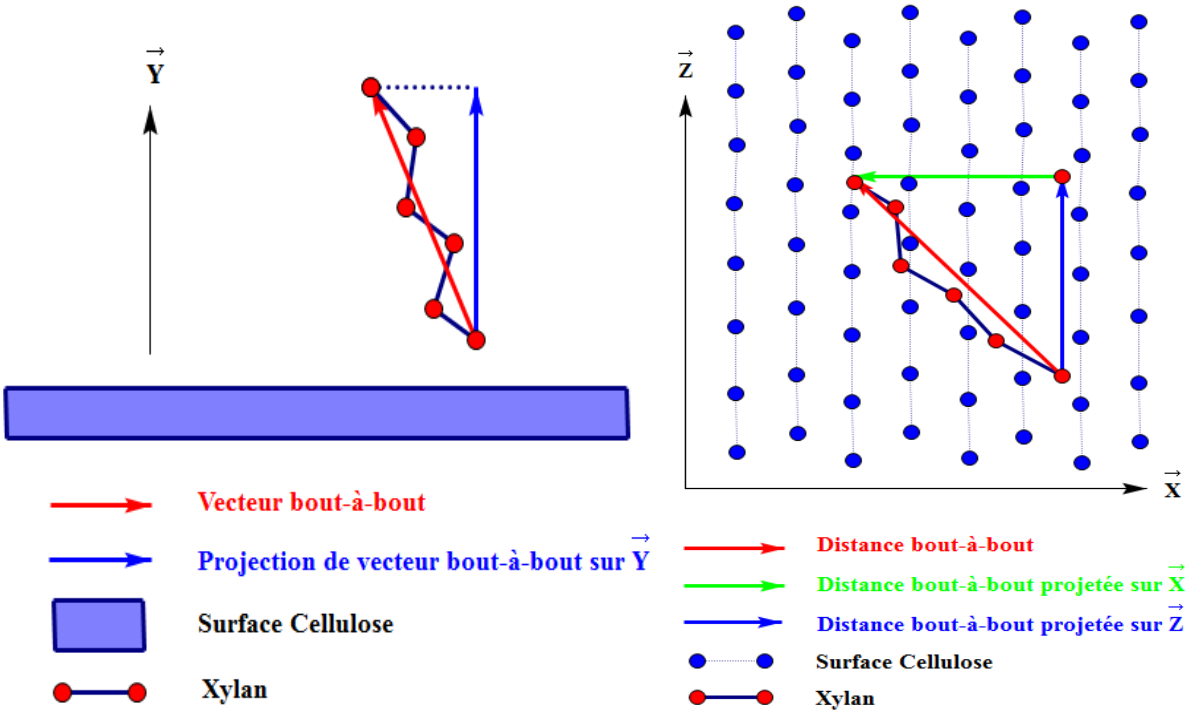

Figure III-47 Projection du vecteur bout-à-bout. A gauche : la projection sur *Y* du vecteur bout-à-bout de la chaîne de xylane. A droite : les projections sur X et Z du vecteur bout-a-bout de la chaîne de xylane.

La projection du vecteur bout-à-bout sur la direction *Z* permet de décrire l'orientation de la chaîne de xylane par rapport à l'axe de la cellulose (Fig. III-48, Fig. III-49).

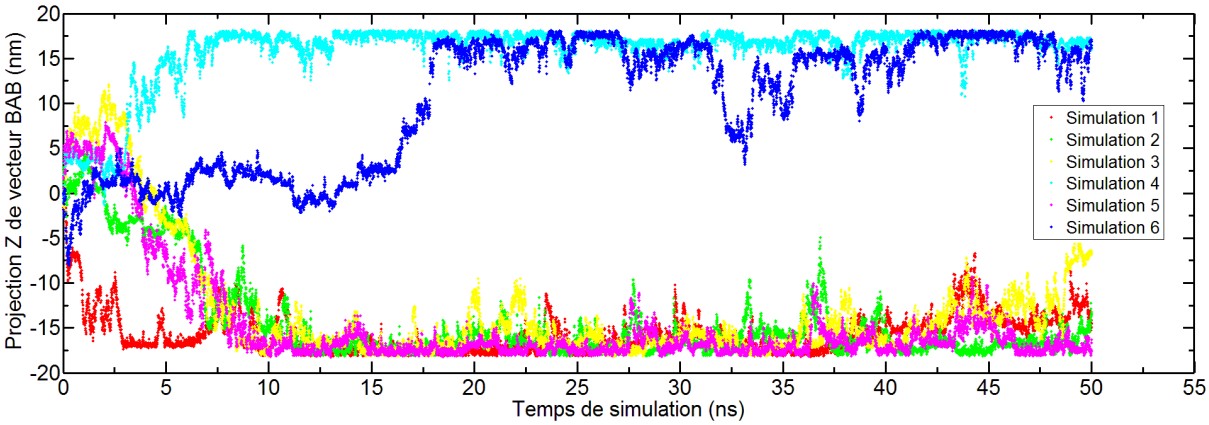

Figure III-48 Evolution temporelle de la projection du vecteur bout-à-bout de la chaîne de xylane projetée sur l'axe *Z* de la surface 110.

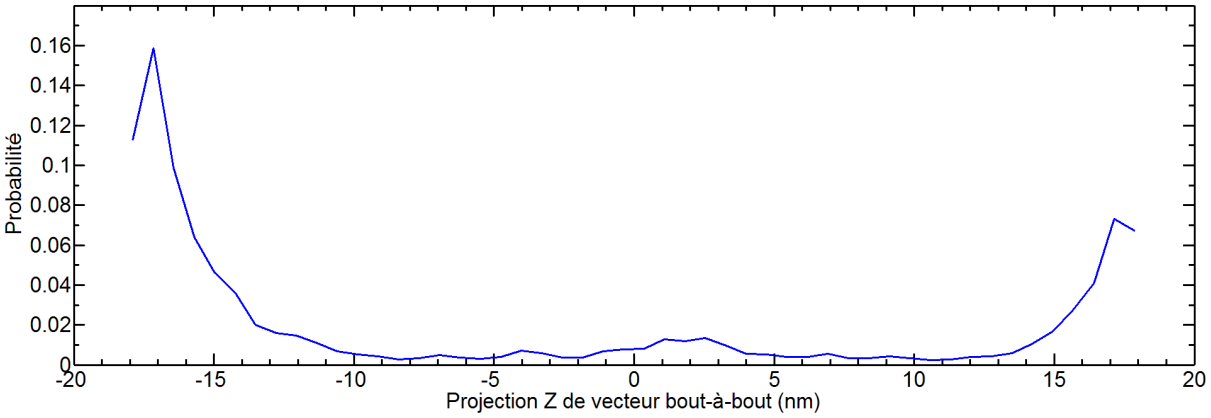

Figure III-49 Distribution de la projection du vecteur bout-à-bout de la chaîne de xylane sur la direction *Z* pour la surface 110.

On constate que, pour la surface 110, la chaîne de xylane est très majoritairement alignée avec les chaînes de cellulose, puisque les valeurs absolues des projections sur *Z* sont très proches de la distance bout-à-bout de la chaîne dans l'état étendu. Ces résultats sont en accord avec ceux obtenus par dynamique moléculaire atomistique [Mazeau and Charlier 2012, Hanus and Mazeau, 2006].

Les valeurs comprises entre les deux extrêmes signifient que le xylane est incliné selon l'axe de cellulose ou qu'il forme des coudes. La formation de coudes n'était pas observée dans les modèles atomistiques car le DP du xylane dans ces modèles est trop faible. Ces valeurs intermédiaires sont en partie dues, dans le modèle atomistique, à des mouvements de retournement global de la chaîne à 180°. De tels comportements n'ont pas été observés dans les trajectoires CG. De toute évidence, le DP plus grand des chaines de xylane implique que ce type de mouvement est beaucoup plus coûteux énergétiquement.

Le comportement de la chaîne de xylane sur la surface 100 est différent. Comme l'indique la projection du vecteur bout-à-bout du xylane sur l'axe *Z* (Fig. III-50, Fig. III-51), la chaîne du xylane n'est pas alignée selon l'axe de la cellulose.

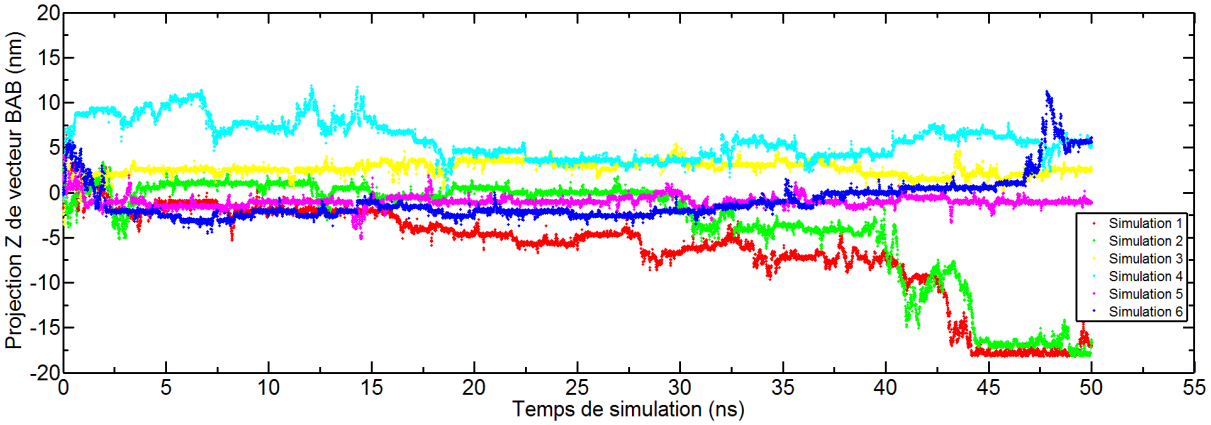

Figure III-50 Evolution temporelle de la projection du vecteur bout-à-bout de la chaîne de xylane sur l'axe *Z* pour la surface 100.

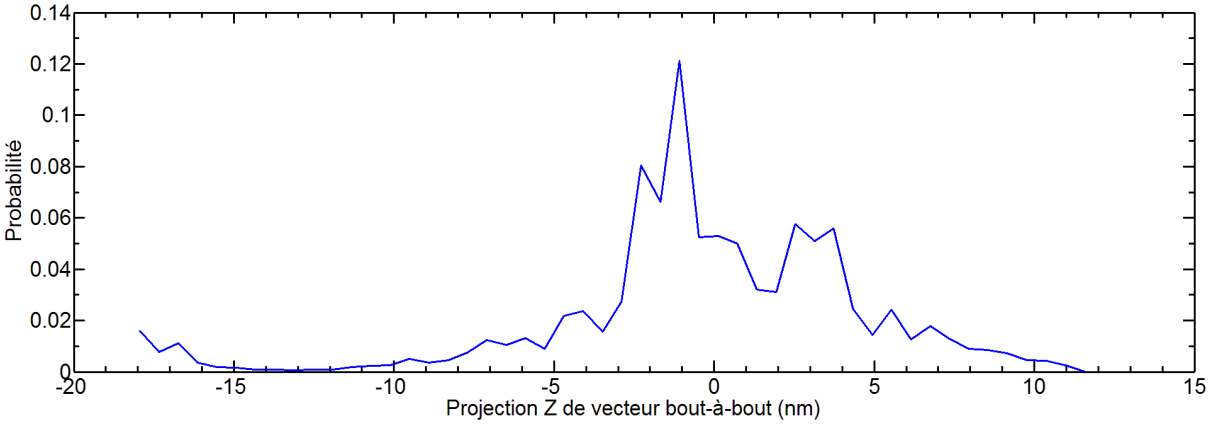

Figure III-51 Distribution de la projection du vecteur bout-à-bout de la chaîne de xylane sur la direction *Z* pour la surface 100.

Au vu des trajectoires, le xylane forme davantage de coudes sauf à la fin des simulations 1 et 2 pour lesquelles l'alignement du xylane sur l'axe de la cellulose a lieu à partir de 43 ns. Ces constatations sont cohérentes avec ce qui a été constaté au vu de la distance bout-à-bout.

Des phénomènes similaires apparaissent dans les simulations d'adsorption du xylane sur la surface 1-10 (Fig. III-52, III-53). Dans les trajectoires des simulations 5 et 6, le xylane forme des coudes après adsorption sur la surface. Dans les autres simulations, le xylane s'aligne avec l'axe de la cellulose, mais après un intervalle de temps plus important que dans le cas de la surface 110. On constate également que l'évolution de la projection du vecteur bout-à-bout sur *Z* est cohérente avec l'évolution de la distance bout-à-bout. La projection sur *Z* et la distance bout-à-bout de la chaîne atteignent leurs valeurs maximales simultanément, comme on s'y attend.

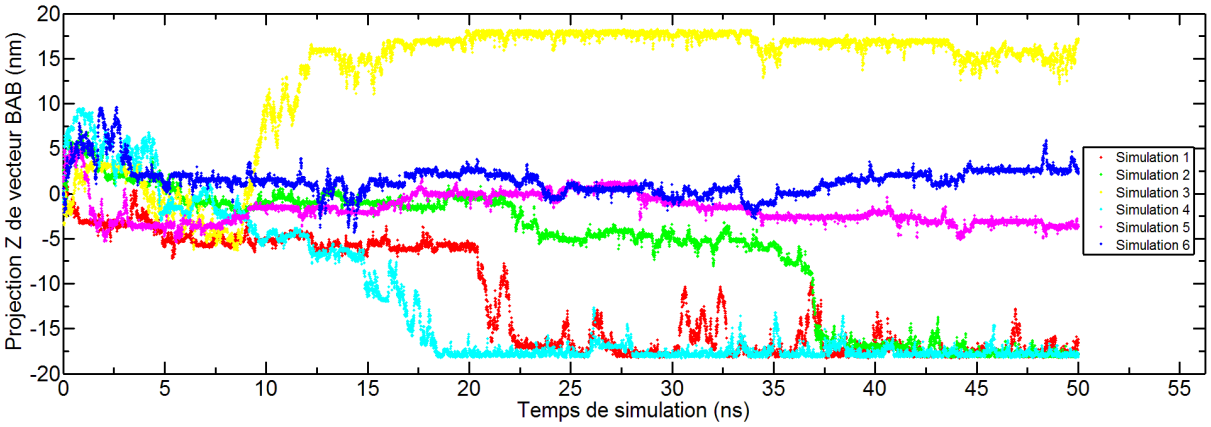

Figure III-52 Evolution temporelle de la projection du vecteur bout-à-bout de la chaîne de xylane sur l'axe *Z* de la surface 1-10.

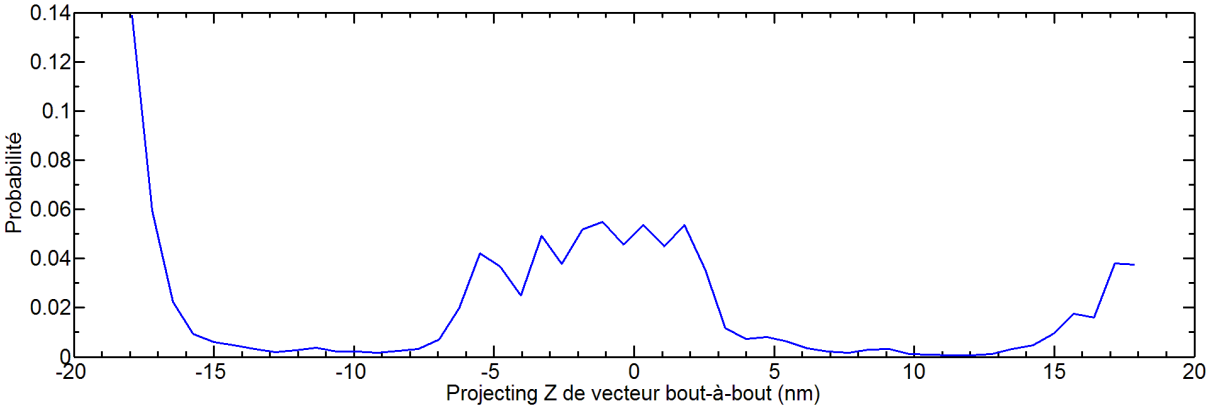

Figure III-53 Distribution de la projection du vecteur bout-à-bout de la chaîne de xylane sur l'axe *Z* pour la surface 1-10.

Au vu de ces résultats, on peut conclure que la tendance des chaînes de xylane à s'aligner avec la cellulose varie avec la surface de cellulose considérée. Si le xylane s'aligne très facilement avec la cellulose pour la surface 110, cette tendance est nettement moindre pour la surface 1-10, et la surface 100. Si le xylane est susceptible de cristalliser partiellement sur la cellulose, ce phénomène semble donc dépendre fortement de la surface considérée. Au vu des trajectoires obtenues, cette dépendance semble davantage d'origine cinétique que proprement énergétique, d'autant plus que la surface induisant le plus un alignement de la chaîne de xylane est en même temps la surface présentant l'interaction la moins intense avec le xylane, parmi les trois surfaces étudiées.

III.2.2 ARRACHEMENT D'UNE CHAINE DE XYLANE D'UNE SURFACE DE CELLULOSE

Les pointes de microscope à force atomique (AFM) permettent de réaliser des expériences de détachement de polymères initialement adsorbés sur des surfaces [Ritort 2006] afin de tester certaines propriétés physiques des polymères [Kierfeld 2006]. Ce type d'expérience a fait l'objet de simulations numériques, notamment par Monte-Carlo [Morris *et al.* 2004, Paturej *et al.* 2012]. Une application du champ de force CG mis au point dans le cadre de cette thèse à ce type de situation est ici présentée.

III.2.2.1 Détails de la simulation

Afin d'explorer le comportement d'une chaîne de xylane pendant son détachement d'une surface de cellulose, un pseudo-dispositif AFM est mis en place dans le logiciel de simulation. Dans les expériences numériques présentées dans cette section, le micro-levier d'AFM est lié au grain d'un bout de la chaîne de xylane, l'interaction entre eux est du type harmonique.

$$\vec{F} = k \cdot (\vec{r}_{micro-levier} - \vec{r}_{grain})$$ Équation III-2

Les paramètres utilisés sont issus de la littérature [Bergenstråhle *et al.* 2009, Morris *et al.* 2004, Su and Purohit 2009], la raideur de l'interaction xylane-micro-levier est fixée à 830 $pN*nm^{-1}$, la distance d'équilibre est fixée à 0, la vitesse de déplacement du micro-levier d'AFM est de 0,6 *m/s* , ce qui est inférieur aux valeurs utilisés dans les modèles atomistiques [Bergenstråhle *et al.* 2009]. Bien entendu, cette valeur est très supérieure à celle mise en œuvre aux cours des expériences réelles, qui sont de l'ordre du *µm/s* [Franca *et al.* 2010, Lee *et al.* 2007, Kierfeld 2006]. Pendant les simulations d'arrachement, la surface de la cellulose est toujours figée. La chaîne de xylane est tirée perpendiculairement à la surface, suivant la direction *Y* (Fig. III-54, Fig. III-55).

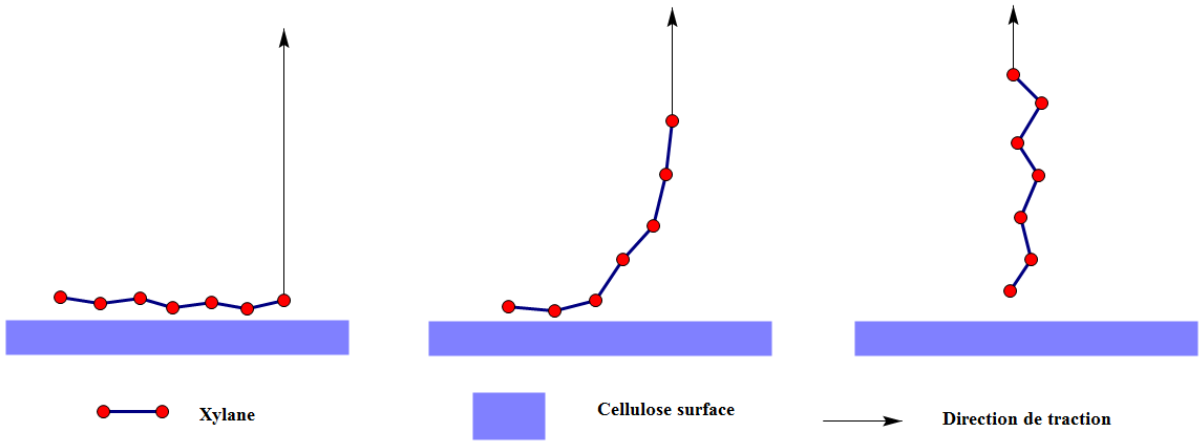

Figure III-54 Schéma de principe de l'arrachement d'une chaîne de xylane par le micro-levier d'AFM. La flèche indique la direction de la force de traction exercée.

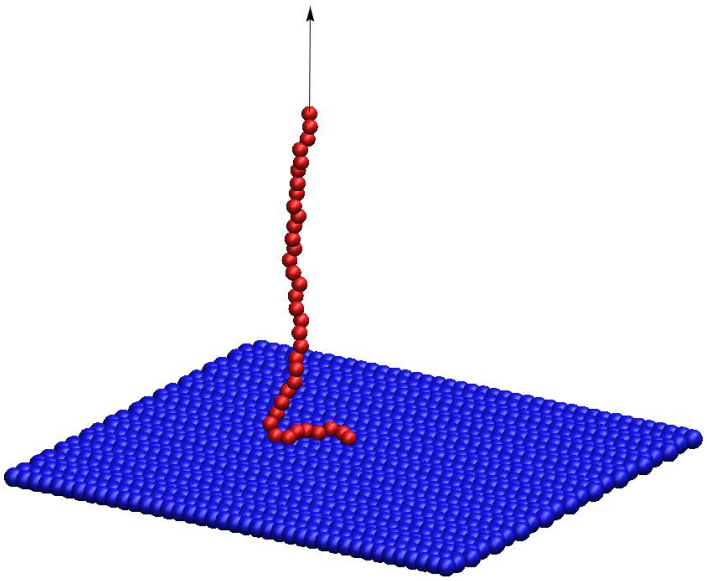

Figure III-55 Aperçu de la simulation de l'arrachement d'une chaîne de xylane perpendiculairement à une surface (110) de cellulose.

Cinq types de configurations initiales ont été utilisés, toutes issues des simulations CG de l'adsorption du xylane sur la surface :

1. La chaîne de xylane est étendue et forme un angle de 0º avec l'axe des chaînes de cellulose.

2. La chaîne de xylane est étendue et forme un angle de 45º avec l'axe des chaines de cellulose.
3. La chaîne de xylane est étendue et forme un angle de 90º avec l'axe des chaines de cellulose.
4. La chaîne de xylane forme des coudes sur la surface de la cellulose et présente plusieurs parties alignées sur les chaînes de cellulose.
5. La chaîne de xylane est adsorbée sur la surface de la cellulose de manière désordonnée.

Un aperçu de ces 5 cas est présenté en figure III-56.

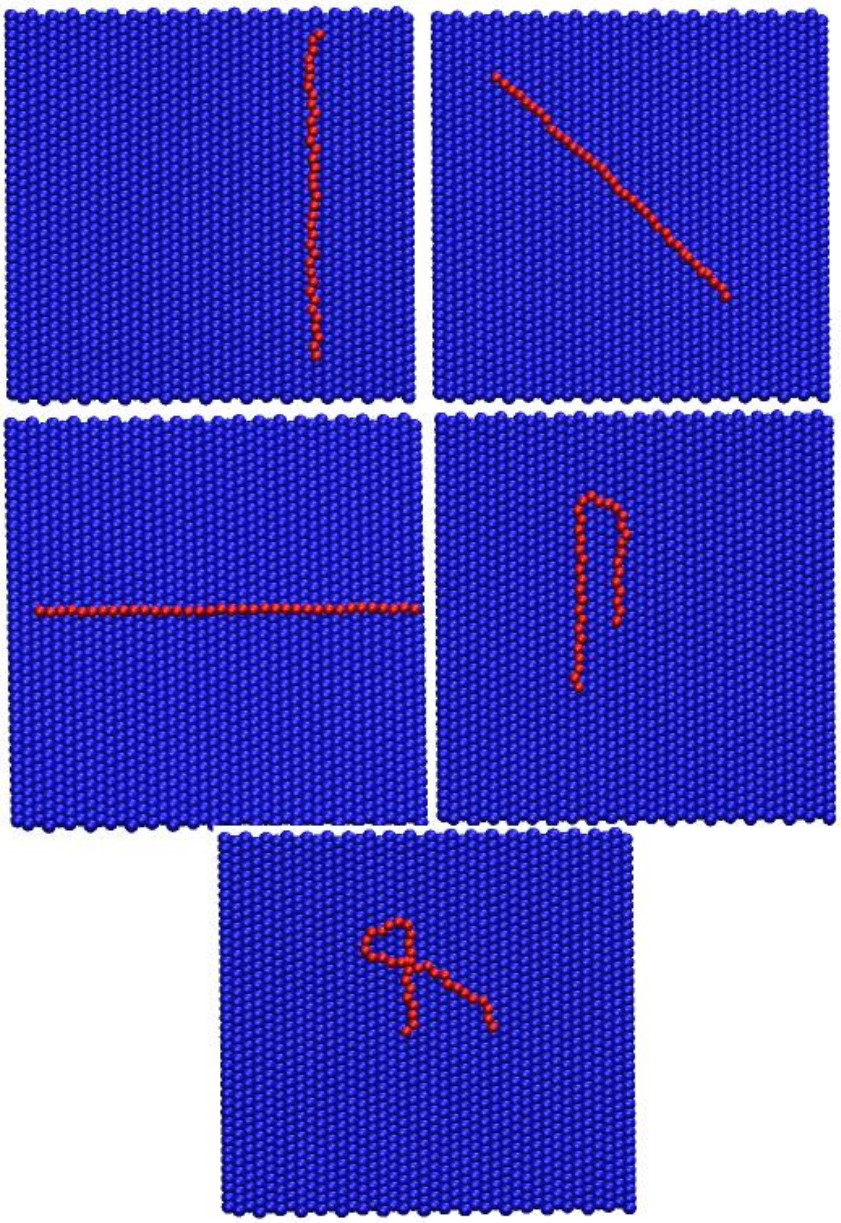

Figure III-56 Aperçu des configurations initiales d'une chaîne de xylane de DP = 36 sur une surface de 40 chaînes de cellulose de DP = 40. En haut à gauche : configuration 0º. En haut à droite : configuration 45º. Au milieu à gauche : configuration 90º. Au milieu à droite : configuration coudée. En bas : configuration aléatoire.

La simulation d'arrachement de la chaîne de xylane est réalisée avec ces 5 configurations initiales pour chaque type de surface sous le même paramétrage. Les paramètres de simulation sont détaillés ci-dessous :

Type	Température	Temps total	Pas de temps	Coefficient de Mass Q	Temps de simulation
Détachement Xylane/Cellulose surface	**300K**	**37,5 ns**	**25 fs**	**1,0**	**~2661 s**

Tableau III-3 Paramètres des simulations d'arrachement.

III.2.2.2 Résultats

L'énergie totale a été calculée au cours de chaque simulation (Fig. III-57, III-58, III-59). L'énergie totale augmente au cours du temps, du fait du travail de la force exercée par la pointe AFM sur le dernier grain de la chaîne de xylane. Cette énergie compense l'énergie d'interaction du xylane avec la surface de cellulose jusqu'à ce que la chaine soit totalement arrachée, pour une distance pointe AFM-surface imposée d'environ 18 nm. Au moment où le xylane est entièrement décollé de la surface, l'énergie du système est constante.

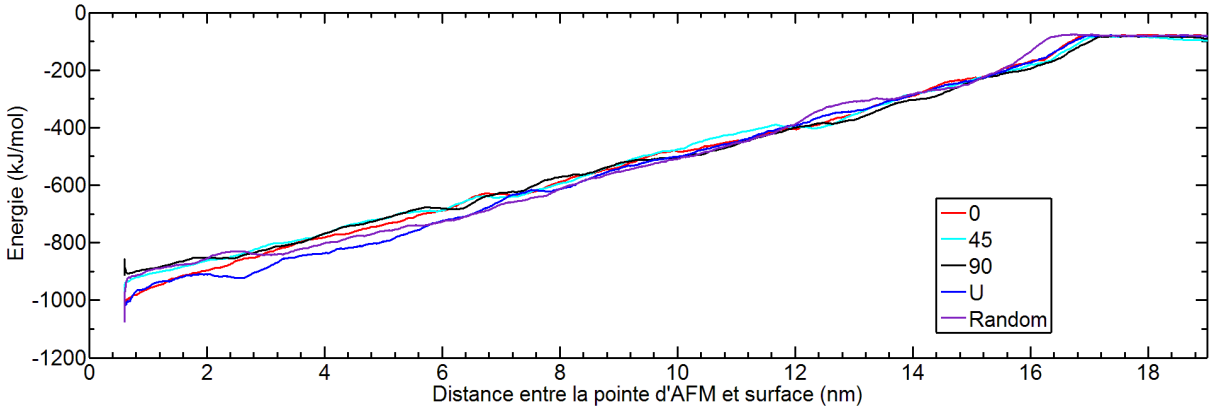

Figure III-57 Evolution temporelle de l'énergie totale du système lors de l'arrachement d'une chaine de xylane pour la surface 110.

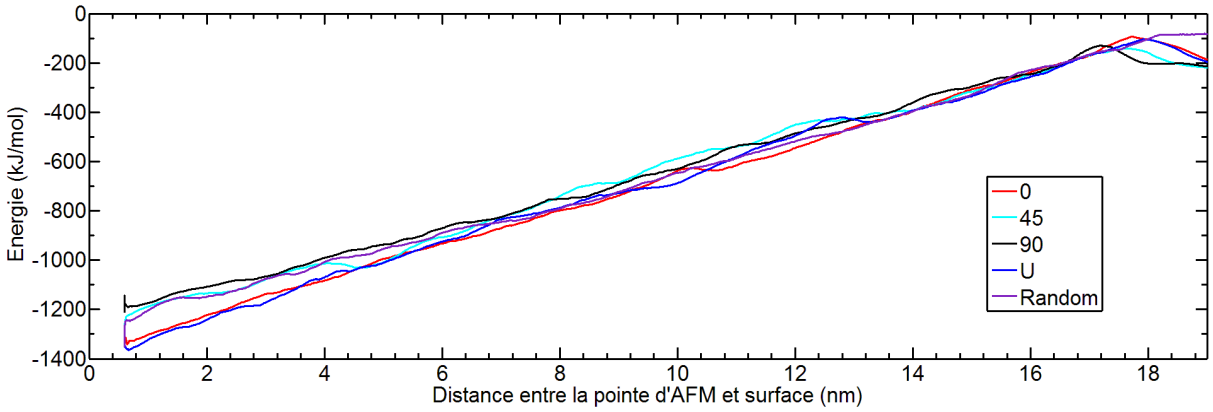

Figure III-58 Evolution temporelle de l'énergie totale du système lors de l'arrachement d'une chaine de xylane pour la surface 100.

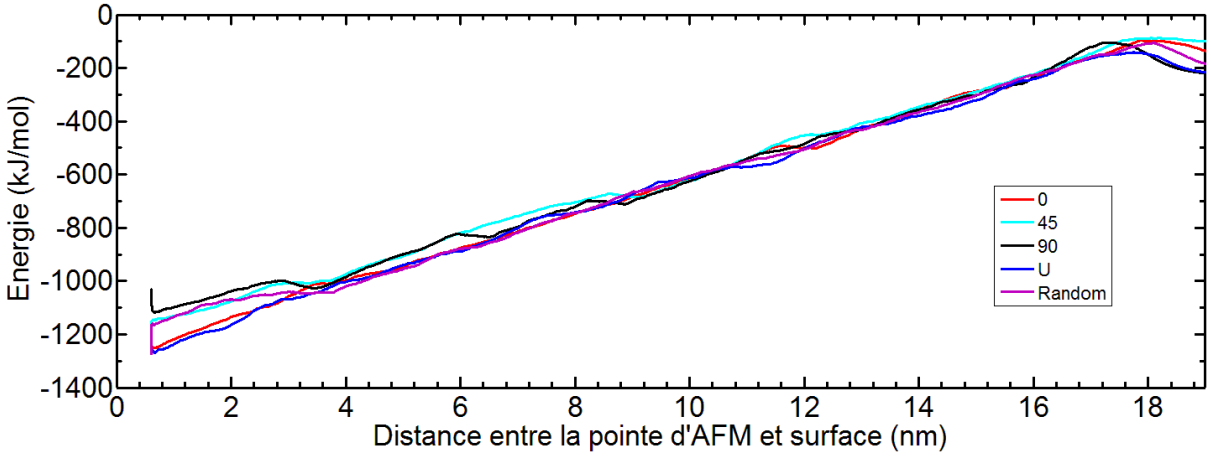

Figure III-59 Evolution temporelle de l'énergie totale du système lors de l'arrachement d'une chaine de xylane pour la surface 1-10.

On constate qu'il n'existe pas de différence significative entre les 5 configurations initiales pour un même type de surface. Cependant, il est intéressant de constater dans tous les cas, l'énergie de départ des configurations initiales parallèles aux chaines de cellulose et coudées est toujours plus basse que celle des autres cas. Cette expérience numérique, en quelque sorte inverse de l'expérience d'adsorption, donne accès à une évaluation de l'énergie d'interaction effective entre le xylane et la surface de cellulose considérée, à travers l'énergie nécessaire pour arracher la chaine de xylane. Les résultats sont rassemblés dans le tableau ci-dessous :

Energie Détachement (kJ/monomère*mol)	0	45	90	Aléatoire	Coudée	Energie Moyen
Surface 110	**29,7882**	**26 ,8950**	**26,2666**	**28,0388**	**30,3770**	**28,2731**
Surface 100	**37,8450**	**34,1188**	**33,4148**	**36,0912**	**37,7962**	**35,8532**
Surface 1-10	**34,9354**	**33,1035**	**31,1694**	**33,3317**	**34,7592**	**33,4598**

Tableau III-4 Energies de détachement pour les trois types de surface.

Ces énergies sont cohérentes avec celles qui ont été obtenues à partir des expériences numériques d'adsorption. En particulier, le cas de la surface 110 est en accord avec les résultats du modèle atomistique [Mazeau and Charlier 2012] qui indique que cette énergie est entre 25,104 et 33,472 *kJ/monomère*mol*.

Dans cette section III.2, deux applications du modèle CG au sujet de l'interaction locale cellulose-xylane ont été présentées : l'adsorption et la séparation du composite cellulose-xylane. Les résultats, au sujet d'énergie et de la conformation moléculaire, issus de ces deux applications sont en cohérence avec ceux issus des modèles atomistiques et des expériences.

III.3 Multicouche de xylane sur la cellulose

Les résultats des simulations de l'interaction d'une chaîne de xylane avec chacune des surfaces de cellulose considérées ont montré que l'interaction avec la cellulose cristalline a un impact majeur sur la conformation des chaînes de xylane. On peut donc supposer qu'une phase dense amorphe de xylane présentera, au voisinage d'une surface de cellulose [Fleer *et al.* 1993, Wu 1982], une structuration dépendant de la surface. Afin d'éclaircir ce point, nous proposons dans cette section deux modèles de composites constitués d'une phase amorphe de xylane et de cellulose cristalline. Nous considèrerons une surface de cellulose et un monocristal.

III.3.1 PHASE AMORPHE DE XYLANE SUR UNE SURFACE DE CELLULOSE

III.3.1.1 Détails de la simulation

Nous proposons dans cette section d'étudier la structuration d'une phase amorphe de xylane au voisinage d'une surface de cellulose par dynamique moléculaire gros grain.. Le système moléculaire comporte une surface (110) de cellulose de 24 chaînes de DP = 24 et 300 chaînes de xylane de DP = 30, la taille de la surface est d'environ 14 nm * 14 nm et l'épaisseur finale de la phase de xylane est d'environ 6 nm. Le nombre total de grains de xylane participant à l'interaction xylane-cellulose est le 9576.

Une difficulté dans la réalisation d'une telle simulation réside dans la mise en place des chaînes de xylane. La procédure d'initialisation du système décrite dans la III.1.3 implique que le nombre de chaînes de xylane pouvant être placées au voisinage de la surface de cellulose est limité. Afin de remédier à ce problème, l'état initial est construit en suivant les étapes suivantes :

1 Une surface de 24 chaînes de cellulose de 24 grains chacune est générée et placée dans le plan $Y = 0$. La surface choisie est du type 110.
2 Ensuite, on génère 50 chaînes de xylane de DP = 30 et on les place verticalement au voisinage de la surface de cellulose. La simulation est lancée pour ce système jusqu'à ce que toutes les chaînes de xylane soient adsorbées sur la surface. Durant cette étape, une paroi répulsive (Voir III.1.3) est introduite, de façon à confiner les chaînes de xylane.
3 Nous recommençons l'étape 2 jusqu'à ce que 9000 grains de xylane soient au voisinage de la surface de cellulose.

La simulation est exécutée à partir de cette configuration initiale, moyennant des conditions aux limites périodiques dans les directions *X* et *Z* et une boîte de dimension très supérieure à la longueur caractéristique du système dans la direction *Y*. Les paramètres de la simulation sont présentés dans le tableau ci-dessous. Le système a fait l'objet d'une relaxation de 1 ns sous 300 K avant le lancement de la simulation proprement dite.

Type	Température	Temps total	Pas de temps	Coefficient de Mass Q	Temps de simulation
Multi-couches du Xylane/Cellulose Surface	**300K**	**12,5 ns**	**25 fs**	**20,0**	**7 jours**

Tableau III-5 Paramètres de la simulation multicouche de xylane sur une surface de cellulose.

III.3.1.2 Résultats

Au vu de l'aperçu du système, le xylane forme bien une phase amorphe dense au voisinage de la surface de cellulose (Fig. III-60).

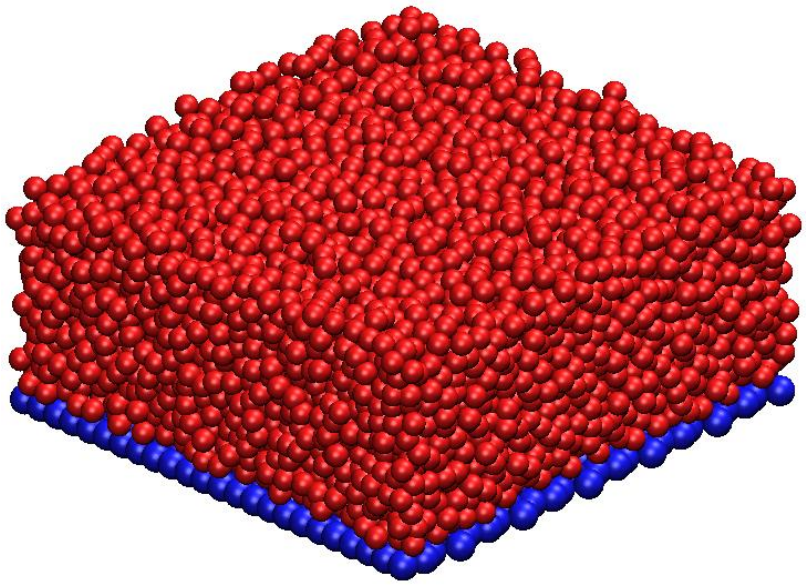

Figure III-60 Aperçu du système surface du type 110 de cellulose cristalline-xylane multicouche. Les grains rouges correspondent à la phase amorphe du xylane, les grains bleus représentent la surface de cellulose.

Afin d'analyser la structure du xylane au voisinage de la surface, nous calculons la densité de probabilité en fonction de la coordonnée *Y* (Fig. III-61).

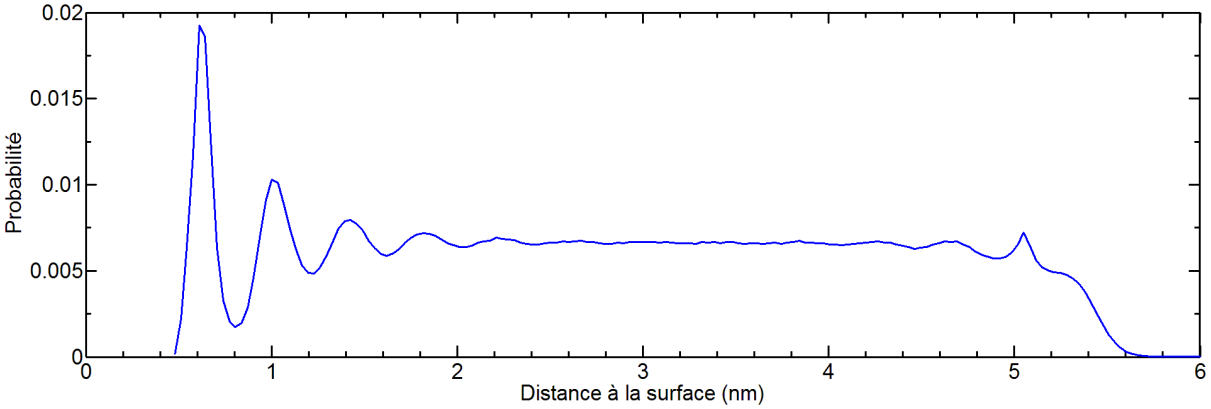

Figure III-61 Distribution de probabilité la distance grains de xylane–surface de cellulose.

On observe très nettement une structuration locale de la phase amorphe au voisinage de la surface de cellulose, dans la mesure où des pics très nets apparaissent. Cette structuration est atténuée avec la distance à la surface et disparaît au-delà de 2,5 nm, la densité de probabilité fonction de Y étant uniforme au-delà. On peut identifier respectivement 5 couches distantes les unes des autres de 0,39 nm, 0,42 nm, 0,39 nm et 0,39 nm. Le dernier pic, à 5 nm, est un artefact.

III.3.2 PHASE AMORPHE DE XYLANE SUR UN MONOCRISTAL DE CELLULOSE

III.3.2.1 Détails de la simulation

Nous proposons dans cette section un modèle de composite constitué d'une phase amorphe de xylane couvrant un monocristal de cellulose. Dans ce modèle, le nombre de chaînes de xylane est de 200, chaque chaîne comportant 30 grains. Le cristal de cellulose, expose trois types de surface : 110, 100 et 1-10 (types de paramétrages de l'interaction cellulose-xylane), et comporte 65 chaînes de 48 grains chacune, dont 26 chaînes superficielles du cristal sont actives pour l'interaction avec le xylane. Le nombre total de grains, en-dehors des grains de cœur du cristal de cellulose, et de 7248. Des conditions aux limites périodiques sont imposées dans la direction Z et la boîte de simulation est suffisamment grande dans les deux autres directions pour pouvoir être considérée comme infinie.

Comme dans le cas d'un composite xylane-surface de cellulose, il est difficile de construire directement la phase amorphe de xylane en une seule fois. La construction de l'état initial a donc consisté en 3 étapes successives, chacune permettant de déposer 2000 grains de xylane sur le composite

1 On a généré d'abord un monocristal de 65 chaînes de 48 grains chacune. Ensuite, 2000 grains de xylane ont été placés autour du cristal suivant la procédure décrite dans la section III.1.3.

2 Les vecteurs bout-à-bout des chaînes de xylane sont ajustés perpendiculairement à l'axe du cristal et à la surface correspondante pour optimiser l'organisation du xylane dans l'espace. Ensuite, on lance une simulation dans l'ensemble *NVT* d'une durée de 0,835 ns pour que les chaînes du xylanes soient adsorbées sur le cristal. Des parois répulsives, présentées dans la section III.1.3, permettent de confiner le xylane au voisinage du cristal.

3 Puis, on a généré 2000 grains de xylane autour du composite déjà formé, la distance initiale entre les nouvelles chaînes de xylane et le cristal étant choisie de façon à éviter la superposition des nouvelles chaînes avec celles qui ont été précédents déposées sur le cristal. Une étape de relaxation similaire à la précédente a permis d'adsorber la deuxième couche de xylane sur la première.

4 Une dernière étape du même type a permis d'obtenir le système final, constitué de 6000 grains de xylane et de 1248 grains de cellulose.

Une dernière étape de relaxation d'une durée de 1,67 ns permet d'équilibrer le système. Les parois répulsives sont, durant cette étape, progressivement éloignées du composite. La simulation *NVT* est ensuite lancée. Le paramètre est listé au ci-dessous.

Type	Température	Temps total	Pas de temps	Coefficient de Mass Q	Temps de simulation
Multicouhes Xylane/Cellulose Cristal	**300K**	**8,37 ns**	**16,7 fs**	**20,0**	**2,2 jours**

Tableau III-6 Paramètres de la simulation pour le système phase amorphe de xylane-cristal de cellulose.

III.3.2.2 Résultats

Comme on peut le constater au vu de l'aperçu du système (Fig. III-62), toutes les chaînes de xylane sont impliquées dans le composite et couvrent toutes les faces exposées du cristal.

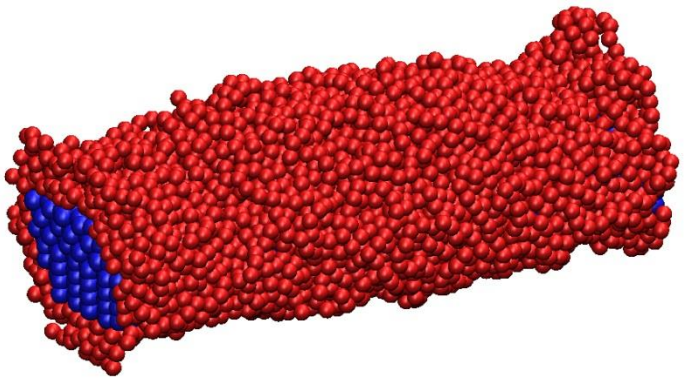

Figure III-62 Aperçu du système phase amorphe de xylane-monocristal de cellulose. Le nombre de grains du xylane est 6000, le nombre de grains du cristal est 1248. Les grains rouges correspondent à la phase amorphe du xylane, les grains bleus représentent la surface de cellulose.

Afin d'effectuer une analyse structurale quantitative, il est nécessaire d'être capable d'attribuer chaque grain de xylane à une des trois faces exposées. Pour ce faire, nous

proposons une stratégie pour l'attribution de grains en fonction de type de surface. Comme illustré en figure III-63, pour la classification, on ne considère que le plan *XY*, la direction de l'axe du cristal *Z* peut être ignorée. Pour distinguer les grains en fonction de leur appartenance à telle ou telle surface, on considère qu'un grain appartient à une surface donnée s'il se trouve dans le prisme de section triangulaire dont l'axe est parallèle à celui du cristal et dont deux faces contiennent chacune une des deux intersections de la face considérée du cristal avec ses voisines. Deux plans doivent être crées reliant respectivement l'axe du cristal, positionnés au barycentre du cristal aux deux côtés de la surface de cristal considérée. Les grains se trouvant entre ces deux plans séparateurs appartiennent à la surface considérée. Cette méthode de classification est appliquée pour toutes les surfaces de cellulose. Comme chaque type de surface est représenté deux fois, les grains appartenant à la surface de même type peuvent être unifiés dans l'analyse statistique. On définit par ailleurs, pour chacune des surfaces, la distance entre un grain de xylane qui lui est attribué et la surface à l'aide de la normale à la surface.

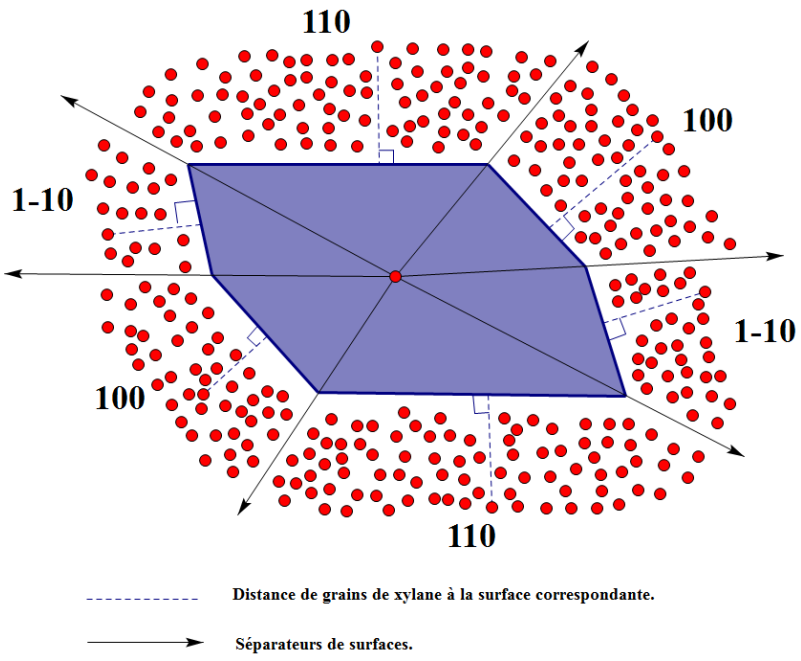

Figure III-63 Classification des grains selon différents types de surfaces du cristal (paramétrage de l'interaction cellulose-xylane). Le cristal est représenté en bleu, le point rouge au centre du cristal est son barycentre, les flèches indiquent les plans séparant les domaines d'espace attribués aux surfaces. La distance entre les grains du xylane et la surface à laquelle ils appartiennent est indiquée en pointillés.

Par cette méthode de classification, la distribution de probabilité en fonction de la distance grain-surface pour les trois types de surface a été extraite de la trajectoire CG (Fig. III-64).

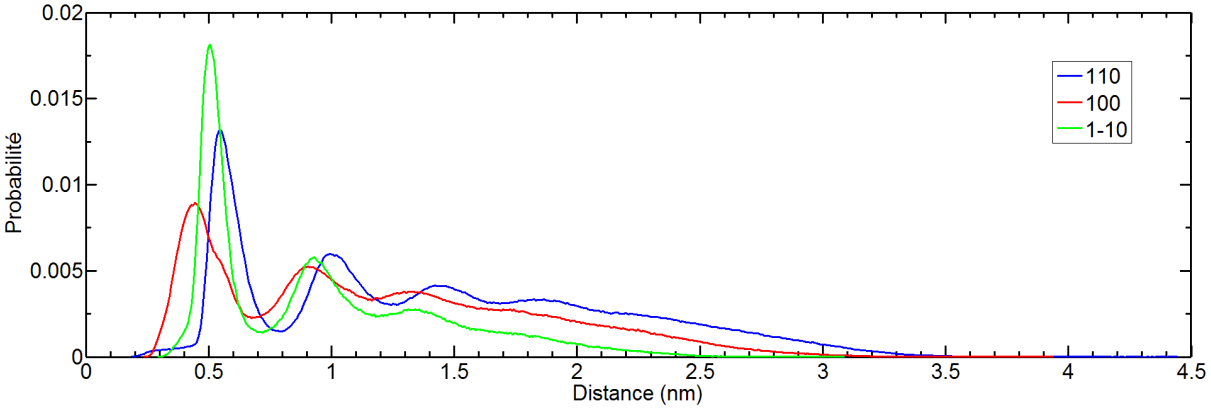

Figure III-64 Distribution de probabilité de la distance grains de xylane-surface pour chaque type de surface.

On peut constater un effet de structuration au voisinage des surfaces similaire à celui observé pour le modèle précédent : des couches successives se forment sur le cristal, couches matérialisées dans la distribution sous forme de pics. De même, cette structuration s'atténue avec la distance pour disparaître aux environs de 2 nm. Il reste cependant difficile de comparer la distance d'atténuation avec celle observée dans le cas d'une surface de cellulose infinie, dans la mesure où l'épaisseur de la couche de xylane est plus faible. Pour la même raison, seules trois couches peuvent être clairement identifiées pour les trois surfaces. On peut constater que l'épaisseur des couches varie avec l'énergie d'interaction, et non pas avec la tendance du xylane à s'aligner avec la cellulose. En effet, plus l'énergie d'interaction xylane-surface est élevée, plus l'épaisseur de couche est faible. On peut donc s'attendre à ce que la densité locale de xylane soit croissante avec l'énergie d'interaction.

A partir de la classification des grains de xylane en fonction des types de surface de cellulose, on définit la densité surfacique de population pour chaque type de surface comme le nombre total de grains de xylane attribués au type de surface considéré normalisée par l'aire de surface correspondante. Afin de comparer ces trois densités, celles-ci sont normalisées par la valeur obtenue pour la surface 110 (Fig. III-65).

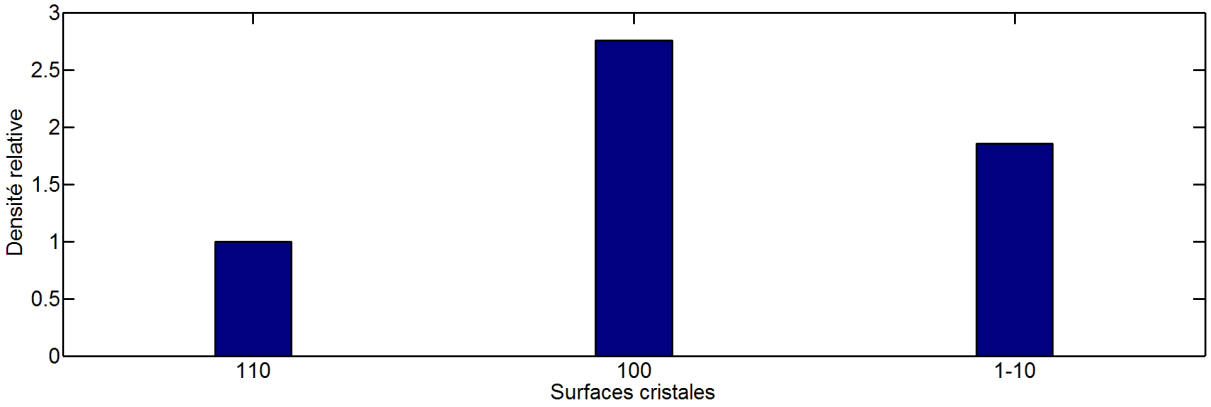

Figure III-65 Densité surfacique relative de population pour les trois types de surface du cristal de cellulose.

Les densités relatives pour les trois types de surface 110, 100, 1-10 sont respectivement 1,0, 2,75 et 1,85. On peut donc en conclure que, comme on s'y attendait, la densité locale est déterminée par l'intensité de l'interaction xylane-surface de cellulose, et non pas par l'alignement du xylane sur la cellulose, constaté dans le cas d'une chaîne de xylane isolée.

Ces constatations corroborent l'hypothèse d'une nature essentiellement cinétique de cet effet d'alignement.

Cette première simulation de la dynamique d'un composite monocristal-xylane n'est qu'une première étape, d'autres investigations seront nécessaires pour comprendre la formation de ce type de système.

Dans cette section III.3, deux applications du modèle CG de la dynamique de composite constitué d'une phase amorphe de xylane et de cellulose cristalline ont été présentées. Analysant les trajectoires dynamiques, la structuration locale de la phase amorphe du xylane au voisinage de la surface solide de cellulose a été dévoilée.

III.4 Conclusions

Plusieurs applications sont présentées dans ce chapitre aux deux sujets :

- ✓ Génération d'états initiaux du xylane et de la cellulose.
- ✓ Interaction de la cellulose et du xylane.

Les résultats produits sont également validés par les modèles atomistiques et les expériences. Cependant, deux défauts au sujet du modèle CG doivent être mentionnés par rapport aux modèles proposés [Terashima *et al.* 2009, Salmén 2004] au niveau de la paroi végétale :

- ➢ Les polymères considérées sont non-substituées.
- ➢ Les groupes fonctionnels ne sont pas compris dans le modèle.

Ces deux défauts signifient que l'introduction de divers types de monomère des molécules végétales est indispensable pour la construction plus complète d'un modèle à l'échelle CG de la paroi plus réaliste. L'absence de telle diversité peut influencer légèrement les résultats obtenus qui seraient plus précis, cependant, qui ne seraient pas fondamentalement basculés.

Conclusions et Perspectives

Conclusions

Une approche de dynamique moléculaire gros grain, à l'échelle du monomère, a été mise au point et appliquée à des modèles de composite cellulose cristalline-xylane.

Dans le modèle gros grain construit dans le cadre de cette thèse, un monomère est assimilé à un grain et les potentiels d'interaction sont paramétrés à l'aide de trajectoires de dynamique moléculaire atomistique. Les potentiels d'interaction intramoléculaires sont déduits d'une inversion de Boltzmann et les potentiels d'interaction intermoléculaires sont issus de processus itératifs d'optimisation calés sur des distributions de probabilité issues des trajectoires atomistiques.

Un code de calcul, programmé en FORTRAN et parallélisé en mémoire partagé par OpenMP a été développé, ainsi que plusieurs outils de construction des états initiaux à l'échelle gros grain.

L'interaction entre le xylane et la surface de la cellulose cristalline a fait l'objet de simulations numériques dans quatre contextes différents.

D'une part, une chaine de xylane de degré de polymérisation réaliste a été adsorbée sur trois types de surface de cellulose cristalline : 110, 100 et1-10. L'énergie d'interaction effective monomère de xylane/surface, la cinétique d'adsorption, ainsi que la probabilité d'alignement du xylane parallèlement aux chaines de cellulose sont contrastées en fonction du type de surface considéré.

D'autre part, la simulation de l'arrachement, par une pointe d'AFM, de la même chaine de xylane, initialement adsorbée sur chacun des trois types de surface de cellulose, donne des résultats en cohérence avec les simulations d'adsorption.

A plus grande échelle, la simulation de l'interaction d'une phase amorphe de xylane avec une surface 110 de cellulose cristalline a permis d'observer un effet de structuration locale du xylane au voisinage de la surface. Le xylane y est structuré sous forme de couches successives, ce qui suggère une cristallisation partielle du xylane amorphe par l'interaction avec la cellulose cristalline.

Enfin, une expérience numérique de recouvrement, par du xylane amorphe, d'un monocristal de cellulose exposant les surfaces 110, 100 et 1-10, a permis de prédire l'apparition de cet effet structurant local au sein du xylane, l'épaisseur des couches ainsi formées dépendant du type de surface, en cohérence avec l'énergie d'interaction xylane/surface déduite des simulations impliquant des chaînes de xylane isolées.

Perspectives

Modèle gros grain

Le modèle gros grain développé dans le cadre de cette thèse comporte une restriction forte qui limite ses applications potentielles. En effet, qu'il s'agisse de surfaces ou d'un cristal entier, la cellulose y est systématiquement figée à l'échelle du grain. Cette approximation a été validée au vu de la trajectoire atomistique générée par Karim Mazeau. Si cette limite n'implique pas de difficulté spécifique pour modéliser l'interaction de la cellulose avec d'autres polymères pariétaux à l'état stationnaire, elle représente un handicap pour les applications relevant de la mécanique, et impliquant potentiellement des déformations de la cellulose. Une solution simple pourrait être de coupler le champ de force développé lors de cette thèse avec le modèle de Srinivas *et al*. [Srinivas *et al.* 2011], qui s'appuie sur un coarse graining identique à celui mis en œuvre dans cette thèse.

Le modèle CG a été établi à partir de modèles atomistiques homopolymère non-substitués, il serait plus intéressant d'imposer de manière statistique les types variables de grain qui représentent des monomères substitués, cela permet de construire un modèle plus réaliste de parois végétales. Le modèle amélioré permet de produire des résultats plus précis, cependant, il serait peu probable que cela bascule fondamentalement les résultats obtenus.

Les calculs de dynamique moléculaire atomistique ont montré que la présence de groupes latéraux sur une chaîne de xylane isolée ne modifie pas fondamentalement l'interaction de cette chaine avec une surface de cellulose (en l'occurrence 110) [Mazeau and Charlier, 2012]. La situation est peut-être différente lorsque le xylane se présente sous la forme d'une phase dense amorphe. L'influence de ces groupes latéraux peut s'avérer déterminante dans la phase amorphe de xylane elle-même, en particulier lorsqu'une déformation globale du système est imposée. De manière générale, la prise en compte d'architectures moléculaires plus réalistes pour le xylane est essentielle, et constitue une perspective prioritaire de cette thèse.

Un des constituants de la paroi cellulaire du bois est absent du modèle développé pendant cette thèse : la lignine. Du fait des difficultés à étudier expérimentalement la structure 3D de cette molécule [Boerjan *et al.* 2003, Kilpelaeinen *et al.* 1994], l'utilisation de la dynamique moléculaire est ici particulièrement pertinente. La construction de potentiels d'interaction impliquant les blocs élémentaires de la lignine et leur intégration dans le champ de force CG développé pendant cette thèse pourrait s'appuyer sur les trajectoires produites via les modèles atomistiques [Besombes and Mazeau 2003, Besombes and Mazeau 2005[a], Besombes and Mazeau 2005[b]].

Le deuxième constituant absent de la paroi cellulaire du bois est l'eau. L'importation de molécules d'eau aurait renforcé fortement la qualité du modèle CG. Il serait possible d'introduire la molécule d'eau de diverse manière :

- Introduction de l'eau à l'échelle CG en remplaçant H_2O par un grain spécifique.
- Introduction de l'eau à l'échelle atomistique, par conséquent, le modèle présenté dans la thèse serait de l'échelle mélange : gros grain et atomistique.
- Omission de la présence de l'eau dans le modèle CG mais en tenir en compte dans les simulations atomistiques.

Avec les améliorations présentées dans cette section, il sera possible de construire un modèle CG complet de la réelle paroi végétale.

Applications

Plusieurs chaines de xylanes de DP variés sont présentées dans les applications du troisième chapitre. Il serait intéressant d'évoluer l'influence de la variation du DP de chaines de xylane sur les résultats obtenus. Cela permet de vérifier les effets de structuration locale sur la cellulose cristalline.

Le modèle développé ici peut, en l'état, faire l'objet d'applications supplémentaires. Il serait possible de former un modèle de composite plus réaliste, en intercalant une couche de xylane entre deux surfaces de cellulose, voire entre deux cristaux de cellulose. Pour aller plus loin, un tel composite peut subir une déformation imposée, ce qui permettrait d'étudier les effets des structures supramoléculaires sur le comportement mécanique du matériau.

Outils de calcul

Le code de calcul développé dans le cadre de cette thèse a été parallélisé par OpenMP un paradigme de parallélisation en mémoire partagée par directives compilateur. La limite essentielle de l'OpenMP est qu'il ne permet pas d'exécuter un calcul sur plusieurs serveurs différents interconnectés en réseau. Afin d'augmenter la taille des systèmes pouvant être simulés, il est nécessaire de calculer sur cluster. Pour ce faire, le code doit être parallélisé par échange de messages, en s'appuyant par exemple sur des routines MPI. Une telle refonte du code de calcul représente une perspective à moyen terme du présent travail.

RÉFÉRENCES BIBLIOGRAPHIQUES

Abrams, C. F., & Kremer, K. (2001). The effect of bond length on the structure of dense bead–spring polymer melts. *The Journal of Chemical Physics*, 115, 2776.

Adler, D. C., & Buehler, M. J. (2013). Mesoscale mechanics of wood cell walls under axial strain. *Soft Matter*.

Åkerholm, M., & Salmén, L. (2001). Interactions between wood polymers studied by dynamic FT-IR spectroscopy. *Polymer*, 42(3), 963-969.

Alder, B. J., & Wainwright, T. (1957). Phase transition for a hard sphere system. *The Journal of Chemical Physics*, 27(5), 1208-1209.

Alder, B. J., & Wainwright, T. E. (1959). Studies in molecular dynamics. I. General method. *The Journal of Chemical Physics*, 31, 459.

Allen, M. P., & Tildesley, D. J. (Eds.). (1989). Computer simulation of liquids. *Oxford university press*.

Allinger, N. L., Yuh, Y. H., & Lii, J. H. (1989). Molecular mechanics. The MM3 force field for hydrocarbons. 1. *Journal of the American Chemical Society*,111(23), 8551-8566.

Allinger, N. L., Rahman, M., & Lii, J. H. (1990). A molecular mechanics force field (MM3) for alcohols and ethers. *Journal of the American Chemical Society*,112(23), 8293-8307.

Almond, A., & Sheehan, J. K. (2003). Predicting the molecular shape of polysaccharides from dynamic interactions with water. *Glycobiology*, 13(4), 255-264.

Andersen, H. C. (1980). Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics*, 72(4), 2384.

Basdevant, N., Borgis, D., & Ha-Duong, T. (2007). A coarse-grained protein-protein potential derived from an all-atom force field. *The Journal of Physical Chemistry B*, 111(31), 9390-9399.

Bellesia, G., Chundawat, S. P., Langan, P., Redondo, A., Dale, B. E., & Gnanakaran, S. (2012). Coarse-Grained Model for the Interconversion between Native and Liquid Ammonia-Treated Crystalline Cellulose. *The Journal of Physical Chemistry B*, 116(28), 8031-8037.

Berendsen, H. J., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. R. H. J., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81, 3684.

Bergenstråhle, M., Berglund, L. A., & Mazeau, K. (2007). Thermal response in crystalline Iβ cellulose: a molecular dynamics study. *The Journal of Physical Chemistry B*, 111(30), 9138-9145.

Bergenstråhle, M., Wohlert, J., Larsson, P. T., Mazeau, K., & Berglund, L. A. (2008). Dynamics of cellulose-water interfaces: NMR spin-lattice relaxation times calculated from atomistic computer simulations. *The Journal of Physical Chemistry B*, 112(9), 2590-2595.

Bergenstråhle, M., Thormann, E., Nordgren, N., & Berglund, L. A. (2009). Force pulling of single cellulose chains at the crystalline cellulose−liquid interface: A molecular dynamics study. *Langmuir*, 25(8), 4635-4642.

Besombes, S., Robert, D., Utille, J. P., Taravel, F. R., & Mazeau, K. (2003). Molecular modeling of syringyl and p-hydroxyphenyl β-O-4 dimers. Comparative study of the computed and experimental conformational properties of lignin β-O-4 model compounds. *Journal of agricultural and food chemistry*, 51(1), 34-42.

Besombes, S., & Mazeau, K. (2004). Molecular dynamics simulations of a guaiacyl β-O-4 lignin model compound: Examination of intramolecular hydrogen bonding and conformational flexibility. *Biopolymers*, 73(3), 301-315.

Besombes, S., & Mazeau, K. (2005)[a]. The cellulose/lignin assembly assessed by molecular modeling. Part 1: adsorption of a *threo* guaiacyl β-O-4 dimer onto a I*β* cellulose whisker. *Plant Physiology and Biochemistry*, 43(3), 299-308.

Besombes, S., & Mazeau, K. (2005)[b]. The cellulose/lignin assembly assessed by molecular modeling. Part 2: seeking for evidence of organization of lignin molecules at the interface with cellulose. *Plant Physiology and Biochemistry*, 43(3), 277-286.

Bird, R. B., Armstrong, R. C., & Hassager, O. (1987). Dynamics of polymeric fluids. *Fluid Dynamics, John Wiley*. New York.

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin biosynthesis. *Annual review of plant biology*, 54(1), 519-546.

Boutherin, B., Mazeau, K., & Tvaroska, I. (1997). Conformational statistics of pectin substances in solution by a Metropolis Monte Carlo study. *Carbohydrate polymers*, 32(3), 255-266.

Bu, L., Beckham, G. T., Crowley, M. F., Chang, C. H., Matthews, J. F., Bomble, Y. J., ... & Nimlos, M. R. (2009). The energy landscape for the interaction of the family 1 carbohydrate-binding module and the cellulose surface is altered by hydrolyzed glycosidic bonds. *The Journal of Physical Chemistry B*, 113(31), 10994-11002.

Buchete, N. V., Straub, J. E., & Thirumalai, D. (2004). Continuous anisotropic representation of coarse-grained potentials for proteins by spherical harmonics synthesis. *Journal of Molecular Graphics and Modelling*, 22(5), 441-450.

Carbone, P., Varzaneh, H. A. K., Chen, X., & Müller-Plathe, F. (2008). Transferability of coarse-grained force fields: The polymer case. *The Journal of chemical physics*, 128, 064904.

Charlier, L., & Mazeau, K. (2012). Molecular modeling of the structural and dynamical properties of secondary plant cell walls: Influence of lignin chemistry. *The Journal of Physical Chemistry B*, 116(14), 4163-4174.

Core, H. A., Coté, W. A., & Day, A. C. (1979). Wood structure and identification. *Syracuse University Press*.

Curcó, D., Nussinov, R., & Alemán, C. (2007). Coarse-graining the Self-assembly of β-helical Protein Building Blocks. *The Journal of Physical Chemistry B*, 111(50), 14006-14011.

Dervilly-Pinel, G., Thibault, J. F., & Saulnier, L. (2001). Experimental evidence for a semi-flexible conformation for arabinoxylans. *Carbohydrate Research*, 330(3), 365-372.

Eichhorn, S. J., Young, R. J., & Davies, G. R. (2005). Modeling crystal and molecular deformation in regenerated cellulose fibers. *Biomacromolecules*, 6(1), 507-513.

Elazzouzi-Hafraoui, S., Nishiyama, Y., Putaux, J. L., Heux, L., Dubreuil, F., & Rochas, C. (2007). The shape and size distribution of crystalline nanoparticles prepared by acid hydrolysis of native cellulose. *Biomacromolecules*, 9(1), 57-65.

Engelsen, S. B., Cros, S., Mackie, W., & Pérez, S. (1996). A molecular builder for carbohydrates: application to polysaccharides and complex carbohydrates. *Biopolymers*, 39(3), 417-433.

Ercolessi, F., & Adams, J. B. (1994). Interatomic potentials from first-principles calculations: the force-matching method. *EPL (Europhysics Letters)*, 26(8), 583.

Evans, D. J., & Holian, B. L. (1985). The nose-hoover thermostat. *The Journal of chemical physics*, 83, 4069.

Fernández, A., & Colubri, A. (2002). Pathway heterogeneity in protein folding. *Proteins: Structure, Function, and Bioinformatics*, 48(2), 293-310.

Fleer, G. J., Stuart, C., MA, S., & JMHM, C. T., Vincent, B., (1993). *Polymers at interfaces*.

Flory, P. J. Statistical mechanics of chain molecules. 1969. *New York: Interscience*.

Franca, E. F., Amarante, A. M., & Leite, F. L. (2010). Introduction to Atomic Force Microscopy Simulation.

French, A. D., & Dowd, M. K. (1993). Exploration of disaccharide conformations by molecular mechanics. *Journal of Molecular Structure: THEOCHEM*, 286, 183-201.

Gindl, W., Gupta, H. S., Schöberl, T., Lichtenegger, H. C., & Fratzl, P. (2004). Mechanical properties of spruce wood cell walls by nanoindentation. *Applied Physics A*, 79(8), 2069-2073.

Gonnet, P. (2012). Pairwise Verlet lists: Combining cell lists and Verlet lists to improve memory locality and parallelism. *Journal of Computational Chemistry*, 33(1), 76-81.

Grest, G. S., Dünweg, B., & Kremer, K. (1989). Vectorized link cell Fortran code for molecular dynamics simulations for a large number of particles. *Computer Physics Communications*, 55(3), 269-285.

Haile, J. M. (1992). Molecular dynamics simulation: elementary methods. *John Wiley & Sons, Inc.*.

Hanley, S. J., Giasson, J., Revol, J. F., & Gray, D. G. (1992). Atomic force microscopy of cellulose microfibrils: comparison with transmission electron microscopy. *Polymer*, 33(21), 4639-4642.

Hanus, J., & Mazeau, K. (2006). The xyloglucan–cellulose assembly at the atomic scale. *Biopolymers*, 82(1), 59-73.

Heinz, T. N., & Hünenberger, P. H. (2004). A fast pairlist-construction algorithm for molecular simulations under periodic boundary conditions. *Journal of computational chemistry*, 25(12), 1474-1486.

Helbert, W., Nishiyama, Y., Okano, T., & Sugiyama, J. (1998)[a]. Molecular Imaging of Halocynthia papillosa Cellulose. *Journal of structural biology*, 124(1), 42-50.

Helbert, W., Sugiyama, J., Kimura, S., & Itoh, T. (1998)[b]. High-resolution electron microscopy on ultrathin sections of cellulose microfibrils generated by glomerulocytes inPolyzoa vesiculiphora. *Protoplasma*, 203(1-2), 84-90.

Henao, L., & Mazeau, K. (2008). The molecular basis of the adsorption of bacterial exopolysaccharides on montmorillonite mineral surface. *Molecular Simulation*, 34(10-15), 1185-1195.

Henriksson, Å., & Gatenholm, P. (2001). Controlled assembly of glucuronoxylans onto cellulose fibres. *Holzforschung*, 55(5), 494-502.

Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3), 1695-1697.

Houtman, C. J., & Atalla, R. H. (1995). Cellulose-Lignin Interactions (A Computational Study). *Plant physiology*, 107(3), 977-984.

Izvekov, S., & Voth, G. A. (2005). A multiscale coarse-graining method for biomolecular systems. *The Journal of Physical Chemistry B*, 109(7), 2469-2473.

Kabel, M. A., van den Borne, H., Vincken, J. P., Voragen, A. G., & Schols, H. A. (2007). Structural differences of xylans affect their interaction with cellulose. *Carbohydrate Polymers*, 69(1), 94-105.

Keller. R. (1991). Structure Composition Formation du Bois.

Kierfeld, J. (2006). Force-induced desorption and unzipping of semiflexible polymers. *Physical review letters*, 97(5), 058302.

Kilpelaeinen, I., Sipilae, J., Brunow, G., Lundquist, K., & Ede, R. M. (1994). Application of two-dimensional NMR spectroscopy to wood lignin structure determination and identification of some minor structural units of hard-and softwood lignins. *Journal of agricultural and food chemistry*, 42(12), 2790-2794.

Knotts IV, T. A., Rathore, N., Schwartz, D. C., & de Pablo, J. J. (2007). A coarse grain model for DNA. *The Journal of chemical physics*, 126, 084901.

Köhnke, T., Lund, K., Brelid, H., & Westman, G. (2010). Kraft pulp hornification: a closer look at the preventive effect gained by glucuronoxylan adsorption. *Carbohydrate Polymers*, 81(2), 226-233.

Köhnke, T., Östlund, Å., & Brelid, H. (2011). Adsorption of arabinoxylan on cellulosic surfaces: influence of degree of substitution and substitution pattern on adsorption characteristics. *Biomacromolecules*, 12(7), 2633-2641.

Kremer, K. (2003). Computer simulations for macromolecular science. *Macromolecular Chemistry and Physics*, 204(2), 257-264.

Larsson, P. T., Hult, E. L., Wickholm, K., Pettersson, E., & Iversen, T. (1999). CP/MAS 13C-NMR spectroscopy applied to structure and interaction studies on cellulose I. *Solid State Nuclear Magnetic Resonance*, 15(1), 31-40.

Larsson, A. C., Ivanov, A. V., Antzutkin, O. N., Gerasimenko, A. V., & Forsling, W. (2004). Complexation of lead (II) with O, O'-dialkyldithiophosphate ligands: 31P and 13C CP/MAS NMR and single-crystal X-ray diffraction studies. *Inorganica chimica acta*, 357(9), 2510-2518.

Lapierre, C. (2010), Les lignines, des polymères uniques au monde, *Séance Académie d'Agriculture du 17 février 2010*.

Lee, C. K., Wang, Y. M., Huang, L. S., & Lin, S. (2007). Atomic force microscopy: determination of unbinding force, off rate and energy barrier for protein–ligand interaction. *Micron*, 38(5), 446-461.

Lemak, A. S., & Balabaev, N. K. (1994). On the Berendsen thermostat. *Molecular Simulation*, 13(3), 177-187.

Levy, S., York, W. S., Stuike-Prill, R., Meyer, B., & Staehelin, L. A. (1991). Simulations of the static and dynamic molecular conformations of xyloglucan. The role of the fucosylated sidechain in surface-specific sidechain folding. *The Plant Journal*, 1(2), 195-215.

Li, X., Ma, X., Huang, L., & Liang, H. (2005). Developing coarse-grained force fields for cis-poly (1, 4-butadiene) from the atomistic simulation. *Polymer*, 46(17), 6507-6512.

Linder, Å., Bergman, R., Bodin, A., & Gatenholm, P. (2003). Mechanism of assembly of xylan onto cellulose surfaces. *Langmuir*, 19(12), 5072-5077.

López, C. A., Rzepiela, A. J., de Vries, A. H., Dijkhuizen, L., Hünenberger, P. H., & Marrink, S. J. (2009). Martini coarse-grained force field: extension to carbohydrates. *Journal of Chemical Theory and Computation*, 5(12), 3195-3210.

Manyuhina, O. V., Fasolino, A., & Katsnelson, M. I. (2007). Slow dynamics in a model of the cellulose network. *Polymer*, 48(16), 4849-4854.

Marrink, S. J., de Vries, A. H., & Mark, A. E. (2004). Coarse grained model for semiquantitative lipid simulations. *The Journal of Physical Chemistry B*, 108(2), 750-760.

Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., & de Vries, A. H. (2007). The MARTINI force field: coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry B*, 111(27), 7812-7824.

Martyna, G. J., Klein, M. L., & Tuckerman, M. (1992). Nosé–Hoover chains: the canonical ensemble via continuous dynamics. The Journal of chemical physics,97, 2635.

Matthews, J. F., Skopec, C. E., Mason, P. E., Zuccato, P., Torget, R. W., Sugiyama, J., ... & Brady, J. W. (2006). Computer simulation studies of microcrystalline cellulose Iβ. *Carbohydrate research*, 341(1), 138-152.

Mattson, W., & Rice, B. M. (1999). Near-neighbor calculations using a modified cell-linked list method. *Computer Physics Communications*, 119(2), 135-148.

Maupetit, J., Tuffery, P., & Derreumaux, P. (2007). A coarse grained protein force field for folding and structure prediction. Proteins: Structure, Function, and Bioinformatics, 69(2), 394-408.

Mazeau, K., & Heux, L. (2003). Molecular dynamics simulations of bulk native crystalline and amorphous structures of cellulose. *The Journal of Physical Chemistry B*, 107(10), 2394-2403.

Mazeau, K., Moine, C., Krausz, P., & Gloaguen, V. (2005). Conformational analysis of xylan chains. *Carbohydrate research*, 340(18), 2752-2760.

Mazeau, K. (2005). Structural micro-heterogeneities of crystalline Iβ-cellulose. *Cellulose*, 12(4), 339-349.

Mazeau, K., & Charlier, L. (2012). The molecular basis of the adsorption of xylans on cellulose surface. *Cellulose*, 19(2), 337-349.

Metropolis, N., & Ulam, S. (1949). The monte carlo method. *Journal of the American statistical association*, 44(247), 335-341.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21, 1087-1092.

Miyoshi, K., Uezu, K., Sakurai, K., & Shinkai, S. (2006). Inter-chain and arrayed hydrogen bonds in β-1, 3-d-xylan triple helix predicted by quantum mechanics calculation. *Carbohydrate polymers*, 66(3), 352-356.

Monticelli, L., Kandasamy, S. K., Periole, X., Larson, R. G., Tieleman, D. P., & Marrink, S. J. (2008). The MARTINI coarse-grained force field: extension to proteins. *Journal of Chemical Theory and Computation*, 4(5), 819-834.

Mora, F., Ruel, K., Comtat, J., & Joseleau, J. P. (1986). Aspect of native and redeposited xylans at the surface of cellulose microfibrils. *Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood*, 40(2), 85-91.

Morris, S., Hanna, S., & Miles, M. J. (2004). The self-assembly of plant cell wall components by single-molecule force spectroscopy and Monte Carlo modelling. *Nanotechnology*, 15(9), 1296.

Müller-Plathe, F. (2002). Coarse Graining in Polymer Simulation: From the Atomistic to the Mesoscopic Scale and Back. *ChemPhysChem*, 3(9), 754-769.

Noid, W. G., Chu, J. W., Ayton, G. S., Krishna, V., Izvekov, S., Voth, G. A., ... & Andersen, H. C. (2008). The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. *The Journal of chemical physics*, 128, 244114.

Nosé, S. (1984)[a]. A molecular dynamics method for simulations in the canonical ensemble. *Molecular physics*, 52(2), 255-268.

Nosé, S. (1984)[b]. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81, 511-519.

Nishiyama, Y., Langan, P., & Chanzy, H. (2002). Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction. *Journal of the American Chemical Society*, 124(31), 9074-9082.

Nguyen, H. D., Marchut, A. J., & Hall, C. K. (2004)[a]. Solvent effects on the conformational transition of a model polyalanine peptide. *Protein science*, 13(11), 2909-2924.

Nguyen, H. D., & Hall, C. K. (2004)[b]. Phase diagrams describing fibrillization by polyalanine peptides. *Biophysical journal*, 87(6), 4122-4134.

Nguyen, H. D., & Hall, C. K. (2004)[c]. Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46), 16180-16185.

Olsson, A. M., Bjurhager, I., Gerber, L., Sundberg, B., & Salmén, L. (2011). Ultra-structural organisation of cell wall polymers in normal and tension wood of aspen revealed by polarisation FTIR microspectroscopy. *Planta*, 233(6), 1277-1286.

Petridis, L., Pingali, S. V., Urban, V., Heller, W. T., O'Neill, H. M., Foston, M., ... & Smith, J. C. (2011). Self-similar multiscale structure of lignin revealed by neutron scattering and molecular dynamics simulation. *Physical Review E*, 83(6), 061911.

Petridis, L., & Smith, J. C. (2009). A molecular mechanics force field for lignin. *Journal of computational chemistry*, 30(3), 457-467.

Pérez, S., & Delage, M. M. (1991). A database of three-dimensional structures of monosaccharides from molecular-mechanics calculations. *Carbohydrate research*, 212, 253-259.

Padding, J. T., & Briels, W. J. (2002). Time and length scales of polymer melts studied by coarse-grained molecular dynamics simulations. *The Journal of chemical physics*, 117, 925.

Peter, C., & Kremer, K. (2009). Multiscale simulation of soft matter systems from the atomistic to the coarse-grained level and back. *Soft Matter*, 5(22), 4357-4366.

Paturej, J., Milchev, A., Rostiashvili, V. G., & Vilgis, T. A. (2012). Polymer Detachment Kinetics from Adsorbing Surface: Theory, Simulation and Similarity to Infiltration into Porous Medium. *Macromolecules*, 45(10), 4371-4380.

Queyroy, S., Neyertz, S., Brown, D., & Müller-Plathe, F. (2004). Preparing relaxed systems of amorphous polymers by multiscale simulation: application to cellulose. *Macromolecules*, 37(19), 7338-7350.

Rappé, A. K., Casewit, C. J., Colwell, K. S., Goddard Iii, W. A., & Skiff, W. M. (1992). UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society*, 114(25), 10024-10035.

Rappé, A. K., & Goddard III, W. A. (1991). Charge equilibration for molecular dynamics simulations. The Journal of Physical Chemistry, 95(8), 3358-3363.

Revol, J. F. (1982). On the cross-sectional shape of cellulose crystallites in Valonia ventricosa. *Carbohydrate Polymers*, 2(2), 123-134.

Reith, D., Pütz, M., & Müller-Plathe, F. (2003). Deriving effective mesoscale potentials from atomistic simulations. *Journal of computational chemistry*, 24(13), 1624-1636.

Reith, D., Meyer, H., & Müller-Plathe, F. (2001). Mapping atomistic to coarse-grained polymer models using automatic simplex optimization to fit structural properties. *Macromolecules*, 34(7), 2335-2345.

Roberts, J. C., & Awad El-Karim, S. (1983). The behaviour of surface adsorbed xylans during the beating of a bleached kraft pine pulp. *Cellulose chemistry and technology*, 17(4), 379-386.

Ritort, F. (2006). Single-molecule experiments in biological physics: methods and applications. *Journal of Physics: Condensed Matter*, 18(32), R531.

Sadus, R. J. (2002). Molecular simulation of fluids: theory, algorithms and object-orientation. Access Online via Elsevier.

Salmén, L. (2004). Micromechanical understanding of the cell-wall structure. *Comptes Rendus Biologies*, 327(9), 873-880.

Salmén, L., & Olsson, A. M. (1998). Interaction between hemicelluloses, lignin and cellulose: structure-property relationships. *Journal of pulp and paper science*, 24(3), 99-103.

Sassi, J. F., & Chanzy, H. (1995). Ultrastructural aspects of the acetylation of cellulose. *Cellulose*, 2(2), 111-127.

Sassi, J. F., Tekely, P., & Chanzy, H. (2000). Relative susceptibility of the Iα and Iβ phases of cellulose towards acetylation. *Cellulose*, 7(2), 119-132.

Satge, C. (2002), Etude de nouvelles stratégies de valorisation de mono et polysaccharides, 207 pp, *Thèse de doctorat, Université de Limoges*, Limoges.

Schönberg, C., Oksanen, T., Suurnäkki, A., Kettunen, H., & Buchert, J. (2001). The importance of xylan for the strength properties of spruce kraft pulp fibres. *Holzforschung*, 55(6), 639-644.

Schulz, R., Lindner, B., Petridis, L., & Smith, J. C. (2009). Scaling of multimillion-atom biological molecular dynamics simulation on a petascale supercomputer. *Journal of Chemical Theory and Computation*, 5(10), 2798-2808.

Shelley, J. C., Shelley, M. Y., Reeder, R. C., Bandyopadhyay, S., & Klein, M. L. (2001). A coarse grain model for phospholipid simulations. *The Journal of Physical Chemistry B*, 105(19), 4464-4470.

Shen, T., & Gnanakaran, S. (2009). The stability of cellulose: a statistical perspective from a coarse-grained model of hydrogen-bond networks. *Biophysical journal*, 96(8), 3032-3040.

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., ... & Youngs, H. (2004). Toward a systems approach to understanding plant cell walls. *Science*, 306(5705), 2206-2211.

Srinivas, G., Cheng, X., & Smith, J. C. (2011). A solvent-free coarse grain model for crystalline and amorphous cellulose fibrils. *Journal of Chemical Theory and Computation*, 7(8), 2539-2548.

Stevanic, J. S., & Salmen, L. (2009). Orientation of the wood polymers in the cell wall of spruce wood fibres. *Holzforschung*, 63(5), 497-503.

Su, T., & Purohit, P. K. (2009). Mechanics of forced unfolding of proteins. *Acta Biomaterialia*, 5(6), 1855-1863.

Sugiyama, J., Harada, H., Fujiyoshi, Y., & Uyeda, N. (1985). Lattice images from ultrathin sections of cellulose microfibrils in the cell wall of Valonia macrophysa Kütz. *Planta*, 166(2), 161-168.

Sugiyama, J., Vuong, R., & Chanzy, H. (1991). Electron diffraction study on the two crystalline phases occurring in native cellulose from an algal cell wall. *Macromolecules*, 24(14), 4168-4175.

Tanaka, F., & Iwata, T. (2006). Estimation of the elastic modulus of cellulose crystal by molecular mechanics simulation. *Cellulose*, 13(5), 509-517.

Terashima, N., Fukushima, K., He, L. F., & Takabe, K. (1993). Comprehensive model of the lignified plant cell wall. *Forage Cell Wall Structure and Digestibility*, (foragecellwalls), 247-270.

Terashima, N., Kitano, K., Kojima, M., Yoshida, M., Yamamoto, H., & Westermark, U. (2009). Nanostructural assembly of cellulose, hemicellulose, and lignin in the middle layer of secondary wall of ginkgo tracheid. *Journal of wood science*, 55(6), 409-416.

Terashima, N., Atalla, R. H., Ralph, S. A., Landucci, L. L., Lapierre, C., Monties, B., & Thiverval-Grignon, F. (1996). New preparations of lignin polymer models under conditions that approximate cell wall lignification. *Holzforschung*,50(1), 9-14.

Terashima, N., Atalla, R. H., Ralph, S. A., Landucci, L. L., Lapierre, C., & Monties, B. (1995). New preparations of lignin polymer models under conditions that approximate cell wall lignification. I. Synthesis of novel lignin polymer models and their structural characterization by 13 C NMR. *Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood*, 49(6), 521-527.

Terashima, N., Atalla, R. H., Ralph, S. A., Landucci, L. L., Lapierre, C., & Monties, B. (1996). New preparations of lignin polymer models under conditions that approximate cell wall lignification. II. Structural characterization of the models by thioacidolysis. *Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood*, 50(1), 9-14.

Tozzini, V. (2005). Coarse-grained models for proteins. *Current opinion in structural biology*, 15(2), 144-150.

Tokoh, C., Takabe, K., Sugiyama, J., & Fujita, M. (2002). CP/MAS 13C NMR and electron diffraction study of bacterial cellulose structure affected by cell wall polysaccharides. *Cellulose*, 9(3-4), 351-360.

Trouy-Triboulot, M.C., Triboulot, P. (2012). Matériau bois.

Tschöp, W., Kremer, K., Batoulis, J., Bürger, T., & Hahn, O. (1998). Simulation of polymer melts. I. Coarse-graining procedure for polycarbonates. *Acta Polymerica*, 49(23), 61-74.

Umemura, M., & Yuguchi, Y. (2005). Conformational folding of xyloglucan side chains in aqueous solution from molecular dynamics simulation. *Carbohydrate research*, 340(16), 2520-2532.

Van Daele, Y., Revol, J. F., Gaill, F., & Goffinet, G. (1992). Characterization and supramolecular architecture of the cellulose-protein fibrils in the tunic of the sea peach (*Halocynthia papillosa*, Ascidiacea, Urochordata). *Biology of the Cell*, 76, 87-96.

Viëtor, R. J., Mazeau, K., Lakin, M., & Pérez, S. (2000). A priori crystal structure prediction of native celluloses. *Biopolymers*, 54(5), 342-354.

Verlet, L. (1967). Computer" experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical review*,159(1), 98-103.

Wang, Y., & Voth, G. A. (2010). Molecular dynamics simulations of polyglutamine aggregation using solvent-free multiscale coarse-grained models.*The Journal of Physical Chemistry B*, 114(26), 8735-8743.

Winter, H., Barakat, A., Cathala, B., & Saake, B. (2005). Preparation of arabinoxylan and its sorption on bacterial cellulose during cultivation. *In Macromolecular symposia* (Vol. 232, No. 1, pp. 74-84). WILEYVCH Verlag.

Wohlert, J., & Berglund, L. A. (2011). A coarse-grained model for molecular dynamics simulations of native cellulose. *Journal of Chemical Theory and Computation*, 7(3), 753-760.

Wu, S. (1982). Polymer interface and adhesion. CRC Press.

Wu, S., Zhan, H. Y., Wang, H. M., & Ju, Y. (2012). Secondary Structure Analysis of Native Cellulose by Molecular Dynamics Simulations with Coarse Grained Model. *Chinese Journal of Chemical Physics*, 25(2), 191.

Yao, Z., Wang, J. S., Liu, G. R., & Cheng, M. (2004). Improved neighbor list algorithm in molecular simulations using cell decomposition and data sorting method. *Computer physics communications*, 161(1), 27-35.

Yui, T., & Hayashi, S. (2007). Molecular Dynamics Simulations of Solvated Crystal Models of Cellulose Iα and IIII. *Biomacromolecules*, 8(3), 817-824.

Zhang, Q., Brumer, H., Ågren, H., & Tu, Y. (2011). The adsorption of xyloglucan on cellulose: effects of explicit water and side chain variation. *Carbohydrate research*, 346(16), 2595-2602.

Zhong, L., Matthews, J. F., Hansen, P. I., Crowley, M. F., Cleary, J. M., Walker, R. C., ... & Brady, J. W. (2009). Computational simulations of the *Trichoderma reesei* cellobiohydrolase I acting on microcrystalline cellulose Iβ: the enzyme–substrate complex. *Carbohydrate research*, 344(15), 1984-1992.