



**UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA**

CAROLINE SIMÕES PEREIRA

**DINÂMICA MOLECULAR DE ENZIMAS PARA DESPOLIMERIZAÇÃO DE
CELULOSE E PROPRIEDADES DE INTERAÇÃO ENTRE XILANOS E CELULOSE**

**CAMPINAS
2016**

CAROLINE SIMÕES PEREIRA

**DINÂMICA MOLECULAR DE ENZIMAS PARA DESPOLIMERIZAÇÃO DE
CELULOSE E PROPRIEDADES DE INTERAÇÃO ENTRE XILANOS E CELULOSE**

**Dissertação de Mestrado apresentada ao Instituto de
Química da Universidade Estadual de Campinas como
parte dos requisitos exigidos para obtenção do título de
Mestra em Química na área de Físico-Química.**

Orientador: Prof. Dr. Munir Salomão Skaf

**ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELA
ALUNA CAROLINE SIMÕES PEREIRA, ORIENTADA PELO PROF. DR. MUNIR SALOMÃO SKAF**

**CAMPINAS
2016**

Agência(s) de fomento e nº(s) de processo(s): CNPq, 132423/2014-6

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Química
Camila Barteta Fullin - CRB 8462

P436d Pereira, Caroline Simões, 1991-
Dinâmica molecular de enzimas para despolimerização de celulose e propriedades de interação entre xilanos e celulose / Caroline Simões Pereira. – Campinas, SP : [s.n.], 2016.

Orientador: Munir Salomão Skaf.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Química.

1. Dinâmica molecular. 2. Celulose. 3. Parede celular. I. Skaf, Munir Salomão, 1963-. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Molecular dynamics of enzymes to cellulose depolymerization and properties of interactions between xylan and cellulose

Palavras-chave em inglês:

Molecular dynamics

Cellulose

Cell wall

Área de concentração: Físico-Química

Titulação: Mestra em Química na área de Físico-Química

Banca examinadora:

Munir Salomão Skaf [Orientador]

Carlos Eduardo Driemeier

Francisco Benedito Teixeira Pessine

Data de defesa: 26-02-2016

Programa de Pós-Graduação: Química

BANCA EXAMINADORA

Prof. Dr. Munir Salomão Skaf (Orientador)

Dr. Carlos Eduardo Driemeier (CTBE-CNPEM)

Prof. Dr. Francisco Benedito Teixeira Pessine (IQ-UNICAMP)

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica da aluna.

Este exemplar corresponde à redação final da
Dissertação de Mestrado defendida pela aluna
CAROLINE SIMÕES PEREIRA, aprovada pela
Comissão Julgadora em 26 de fevereiro de 2016.

*“Demore o tempo que for para decidir o que você quer da vida,
e depois que decidir não recue ante nenhum pretexto,
porque o mundo tentará te dissuadir.”*

Friedrich Nietzsche

Agradecimentos

“Nada é definitivo neste universo caótico e cambiante. Mas ciclos vêm e vão. Este é o fim de um e o começo de outro. Que permaneça sempre a boa vontade, o aguerrimento e a determinação.” Clóvis de Barros Filho

Para a conclusão deste ciclo devo meus agradecimentos

- Aos meus pais, Dionísio e Dora, pelo apoio e formação pessoal;
- Ao Prof. Munir, pela oportunidade e confiança em mim depositada durante este período;
- Ao Rodrigo, pelo companheirismo e parceria;
- Aos colegas de trabalho, pela amizade;
- Às colegas de república, principalmente, Elisa, Angélica, Pati e Rairís pela amizade e companheirismo;
- Ao Instituto de Química e à Unicamp, por todo suporte oferecido;
- Ao CNPq e Fapesp, pelo financiamento;
- Aos membros da comissão examinadora.

Resumo

Compreender a arquitetura da parede celular e suas modificações por enzimas é necessário para o desenvolvimento de novos materiais e biocombustíveis. Nesta Dissertação, utilizamos simulações computacionais de Dinâmica Molecular para entender as interações físicas da matriz polimérica da parede celular com a celulose e aspectos moleculares de enzimas que atuam sobre polissacarídeos da parede celular. Especificamente, abordamos dois tópicos: (1) adsorção, em celulose, de glucuronoarabinosilanos, com padrão de ramificação encontrado em coníferas, e (2) estudo da endoglucanase Cel45A do fungo *Phanerochaete chrysosporium* (PcCel45A), que possui alta similaridade com expansinas produzidas em plantas. No tópico (1), mostramos que glucuronoarabinosilanos adsorvem estavelmente na superfície hidrofílica (110) da celulose, enquanto as ramificações de α -1,2-ácido glucurônico estabilizam a adsorção do xilano, as ramificações de α -1,3-arabinose aumentam a mobilidade dos diedros glicosídicos e não contribuem para adsorção na superfície hidrofílica da celulose. Simulações com duas cadeias xilana em paralelo mostram que há efeitos cooperativos para adsorção na celulose, causados pelos efeitos de empacotamento e, no caso das cadeias ramificadas, ocorre reticulações entre ramificações α -1,2-ácido glucurônico e os íons cálcio. No tópico (2), observamos que a PcCel45A tem alta similaridade com expansinas de plantas e difere de outras enzimas da família GH45 pela falta de *loops* para ligação ao substrato e por ter uma asparagina, em vez de aspartato, atuando como base catalítica. As simulações mostraram que a PcCel45A compensa a falta de *loops* com os resíduos hidrofóbicos Trp161 e Tyr74, que reconhecem o substrato via interações de *stacking*, e pela presença de vários resíduos polares e carregados que fazem ligações de hidrogênio com os grupos hidroximetílicos exocíclicos do substrato. Mutações de tais resíduos causaram redução de afinidade pelo substrato. O mutante do resíduo base Asn99 para Asp não apresenta estabilidade conformacional, com isso o Asp99 não consegue atuar como base em uma catálise enzimática, mesmo possuindo maior basicidade que Asn. O *loop* da base Asn99 apresentou duas conformações distintas, a conformação da estrutura cristalina e uma conformação que permite uma ligação de hidrogênio entre base e ácido catalítico, que pode indicar um mecanismo da enzima para que ocorra a troca de prótons após uma reação catalítica e, assim, voltar à forma ativa. Em conjunto, estes estudos fornecem explicações para fenômenos que ocorrem na parede celular de plantas e para aspectos mecanísticos de enzimas que a modificam e degradam.

Abstract

Comprehending the plant cell wall architecture and its modification by enzymes is necessary for the development of novel materials and biofuels. In this dissertation, we have employed molecular dynamics computer simulations to understand physical interactions between the plant cell wall polymeric matrix and cellulose, and molecular aspects of cell wall modifying enzymes. In particular, we have studied two topics: (1) adsorption of glucuronoarabinoxylans onto the cellulose surface, with a substitution pattern xylans found in conifers, and (2) the study of the endoglucanase Cel45A from *Phanerochaete chrysosporium* (PcCel45A), which exhibits a high similarity with plant expansins. In the topic (1), we show that glucuronoarabinoxylans stably adsorb onto the hydrophilic surface (110) of the cellulose. Whereas glucuronic acid substitutions are able to stabilize the xylan adsorption, arabinose substitutions cause an increase of flexibility of the glycosidic torsion angles and do not contribute to stabilize the xylan chain on the cellulose surface. Simulations of two xylan chains revealed cooperativity for the adsorption on the cellulose, caused by packing effects and cross-linking between two glucuronic acids and calcium ions. In the topic (2), we observed that PcCel45A presents a high similarity with plant expansins and differs from other GH45 enzymes by the lack of loops that aid in the substrate binding, and for having an asparagine, instead of an aspartate, acting as catalytic base. The simulations showed that PcCel45A compensates the lack of loops with the presence of hydrophobic residues (Trp161 and Tyr74), which recognize the substrate through stacking interactions, and by the presence of several polar/charged residues that hydrogen bond hydroxymethyl groups of the substrates. Mutations of such residues induced severe decrease of the enzyme-substrate affinity. The catalytic base mutant Asn99Asp exhibited poor conformational stability, so that Asp99 cannot act as a general base in the enzymatic catalysis, even though Asp is more basic than Asn. The loop that contains Asn99 exhibited two distinct conformations: a crystallographic-like conformation and another conformation that allows a hydrogen bond between the catalytic acid and base, suggesting a mechanism of proton exchange from the base to the acid that would restore the enzyme to its active state. Together, these studies provide molecular-level explanations for events that occur within the plant cell wall structure and for mechanistic aspects of enzymes that modify plant cell walls.

Lista de Abreviaturas e Siglas

Ara	Arabinose
Cel45A	Celulase da família 45
CMC	Carboximetilcelulose
DM	Dinâmica molecular
EXLX1	Expansina de <i>Bacillus subtilis</i>
ExpB1	β -expansina de <i>Zea Mays</i> (milho)
GH45	Glicosil hidrolase da família 45
GlcA	Ácido glucurônico
HiCel45	Endoglucanase da família 45 de <i>Humicola insolens</i>
MeGlcA	Ácido O4-metil-glucurônico
NAMD	em inglês Nanoscale Molecular Dynamics (software)
NVE	Número de partículas, volume e energia constantes
PASC	em inglês Phosphoric Acid Swollen Celulose
PcCel45A	Endoglucanase da família 45 de <i>Phanerochaete chrysosporium</i>
PME	Malha de partículas de Ewald
pNPC	para-nitrofenil- β -glucopiranosídeo (sigla em inglês)
RMSD	Desvio da Raiz Média Quadrática (sigla em inglês)

Sumário

1. Introdução	12
1.1 Composição da biomassa lignocelulósica.....	13
1.2 Arquitetura da parede celular	15
1.3 Expansinas	15
1.4 Enzimas da família GH45	17
2. Dinâmica Molecular	20
2.1 Campo de força.....	21
2.2 Posições e velocidades iniciais	23
2.3 Algoritmos	24
2.3.1 Condições periódicas de contorno.....	24
2.3.2 Raio de corte e soma de Ewald	25
2.3.3 Método de Verlet.....	26
2.3.4 Temperatura e pressão.....	27
2.4 Cálculo de propriedades médias.....	27
3. Adsorção de cadeias xilanas à superfície hidrofílica da celulose	29
3.1 Introdução	29
3.2 Métodos	32
3.3 Resultados	35
3.3.1 Interação das cadeias xilana com a face hidrofílica da celulose.....	35
3.3.2 Efeito cooperativo xilano-xilano	41
3.3.3 Efeito das ramificações.....	44
3.3.4 Conformação diedral.....	48
3.4 Discussão	51
3.5 Conclusão	53
4. Endoglucanase Cel45A de <i>Phanerochaete chrysosporium</i>	54
4.1 Introdução	54
4.2 Métodos	57
4.3 Resultados	59
4.3.1 Ligação ao substrato.....	59
4.3.2 Mutante Asn99Asp.....	63

4.3.3 Restauração da forma ativa.....	64
4.4 Discussão	66
4.5 Conclusão	67
5. Considerações finais	68
Referências Bibliográficas.....	70
ANEXO: Nomenclatura e estrutura dos aminoácidos.....	75

Capítulo 1

Introdução

A biomassa lignocelulósica é o recurso renovável mais disponível para obtenção de biocombustíveis atualmente [1]. O desenvolvimento de novas tecnologias para sua conversão em açúcares fermentáveis com menores custos é essencial para viabilizar a produção de bioetanol do ponto de vista econômico. O maior empecilho para sua utilização é a recalcitrância (resistência mecânica, química e biológica) natural do material lignocelulósico, que faz necessária a utilização de severos tratamentos termoquímicos [2].

A recalcitrância é a causa da baixa velocidade das reações enzimáticas de despolimerização e dos elevados custos da produção de biocombustíveis. Portanto, o desenvolvimento de estratégias mais eficientes para aumentar a acessibilidade da celulose às enzimas é de suma importância [2,3]. Do ponto de vista molecular, a recalcitrância se origina das características químicas e da arquitetura do material lignocelulósico, esta ainda não inteiramente conhecida [4]. Compreender a arquitetura da parede celular e suas modificações por enzimas é necessário para o entendimento dos fundamentos da biologia das plantas e para a concepção de novos materiais, desconstrução da biomassa para produção de biocombustíveis e desenvolvimento de mutantes de plantas mais adequadas para a biorrefinaria.

Neste trabalho, utilizamos simulações computacionais de Dinâmica Molecular (DM) para entender interações físicas da matriz polimérica da parede celular com a celulose e aspectos mecanísticos de enzimas que atuam sobre polissacarídeos da parede celular. Especificamente, abordamos dois tópicos: (1) adsorção de glucuronoarabinosilanos de coníferas em celulose e (2) estudo da endoglucanase Cel45A do fungo *Phanerochaete chrysosporium* (PcCel45A), que possui alta similaridade com expansinas produzidas em plantas. Com base nestes tópicos, será feita uma breve introdução a seguir sobre a composição e arquitetura da parede celular, nas seções 1.1 e 1.2, respectivamente, e sobre expansinas e enzimas da família GH45, nas seções 1.3 e 1.4, respectivamente.

1.1 Composição da biomassa lignocelulósica

A biomassa lignocelulósica é constituída basicamente por celulose, hemicelulose e lignina, sendo a celulose a principal matéria-prima para produção de bioetanol. Estes componentes estão presentes na parede celular das plantas e se arranjam em uma arquitetura supramolecular complexa, onde fibras de celulose se estruturam dentro de uma matriz protetora de hemicelulose e lignina [1]. As ramificações da hemicelulose se ligam covalentemente à lignina ou a outras hemiceluloses, formando os complexos de lignina e carboidratos. Este arranjo molecular é um dos maiores responsáveis pela recalcitrância da parede celular [1,5].

A celulose é um polímero composto por cadeias lineares de monômeros de glicose ligados por β -1,4. As ligações de hidrogênio intramoleculares da cadeia glucana fazem dela uma cadeia rígida e favorecem a conformação diedral de dupla-hélice, mostrada na figura 1.1A, onde a soma dos ângulos diedrais glicosídicos φ e ψ é 120° . As cadeias glucana se arranjam formando complexos supramoleculares contendo domínios cristalinos e amorfos [6].

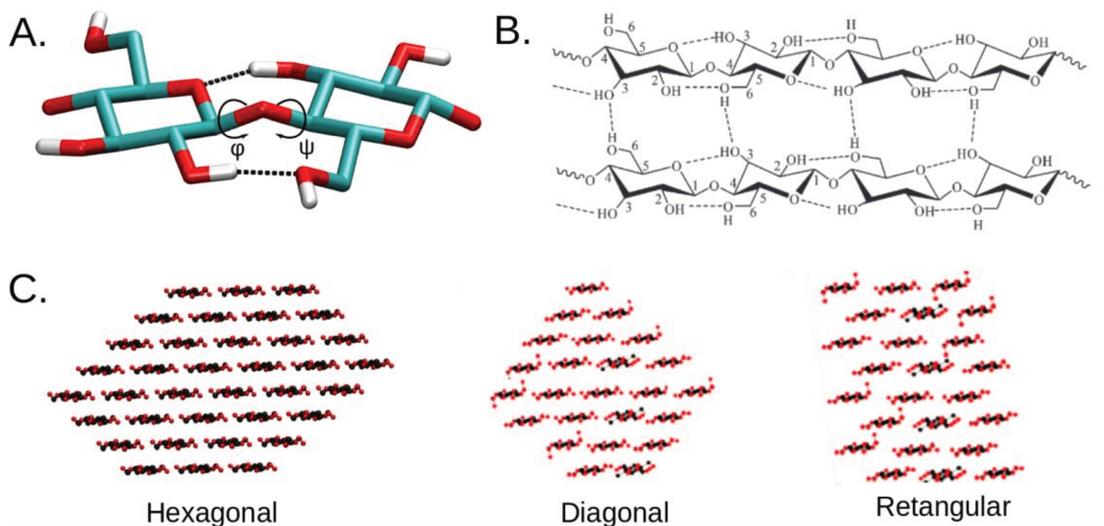


Figura 1.1 Representações das interações e arranjos moleculares da celulose. (A) Celobiose, dois monômeros de glicose ligados por ligação β -1,4. Nesta figura estão indicadas as ligações de hidrogênio intramoleculares da cadeia glucana e os ângulos diedrais φ e ψ . (B) Ligações de hidrogênio inter e intramoleculares de cadeias glucana no domínio cristalino. (C) Microfibrilas, da esquerda para direita, hexagonal com 36 cadeias, diagonal e retangular com 24 cadeias. As figuras 1.1C das microfibrilas diagonal e retangular foram adaptadas de Fernandes et al., 2011 [9].

Nos domínios cristalinos, as cadeias glucana de uma mesma camada interagem entre si por ligações de hidrogênio intermoleculares, mostradas na figura

1.1B, enquanto que as cadeias de camadas diferentes interagem entre si por interações hidrofóbicas, devido ao alto empacotamento [6]. A celulose cristalina encontrada na natureza apresenta os dois alomorfos $I\alpha$ e $I\beta$, que diferem no padrão de empacotamento [7].

O tamanho e a forma das microfibrilas de celulose cristalina encontradas na natureza ainda não são completamente conhecidos. A forma hexagonal, com 36 cadeias, é a mais utilizada para representar a microfibrila até o momento (figura 1.1C) [8], mas estudos sugerem outras formas para microfibrila, como a diamante e retangular (figura 1.1C) [9]. A forma da microfibrila altera as proporções de áreas das superfícies hidrofóbicas e hidrofílicas, com isso interfere na interação da celulose com outros políssacarídeos da parede celular [8].

Devido a seu arranjo na estrutura cristalina, a celulose tem baixa solubilidade na maior parte dos solventes e as microfibrilas são resistentes a hidrólises químicas e enzimáticas [6]. Entretanto, o acesso à celulose também é limitado pelo revestimento da matriz de hemicelulose e lignina, o que eleva o nível de recalcitrância da biomassa [10,11].

Hemiceluloses são heteropolissacarídeos caracterizados, com maior frequência, por cadeias de β -1,4 de resíduos de xilose, glucano ou manose, com diferentes tipos de ramificações [12]. Diferentemente da celulose, a hemicelulose varia sua composição química e abundância de acordo com a planta e com o tecido, pois possui diferentes funções na planta. Em paredes primárias, xiloglucano é a hemicelulose mais abundante, já em paredes secundárias as mais abundantes são o glucomanano e glucuronoarabinoxilano [13,14].

A principal função da hemicelulose é aumentar a resistência da parede celular através de sua associação com celulose e, em paredes secundárias, com a lignina [12]. O arranjo molecular das hemiceluloses com a celulose ainda é pouco descrito; sua interação com a face hidrofóbica da celulose é bem aceita, mas pouco se sabe como ocorre a interação com a face hidrofílica [15,16].

Assim como a hemicelulose, a composição de lignina varia de acordo com a planta e com o tipo de tecido, e é depositada sobre a parede celular apenas depois que cessou a etapa de crescimento da planta. A lignina é composta por polímeros derivados do fenilpropano, com monômeros de *p*-hidroxifenila, guaiacila e siringila [1]. Devido ao seu caráter hidrofóbico, ela expulsa a água da parede celular, criando uma barreira protetora à celulose cristalina e a outros polímeros, impedindo assim o ataque

enzimático e químico. Por se ligar irreversivelmente às enzimas, a lignina contribui para diminuição da atividade catalítica e exige maior quantidade de enzimas para degradar os substratos [4,11,17].

1.2 Arquitetura da parede celular

Durante o crescimento da planta, a parede celular é chamada de primária e é constituída principalmente de celulose, hemicelulose e pectina, com exceção das gramíneas, que não possuem quantidade relevante de pectina. A parede primária, apesar de resistente, é capaz de sofrer relaxação mecânica durante o crescimento. Este mecanismo é causado pela força de turgor da água e é dependente da arquitetura da matriz polimérica, da ação de expansinas e enzimas hidrolíticas [1,8].

Quando o processo de crescimento é cessado, algumas células criam uma parede secundária, onde tecidos de pectina são menos abundantes e as paredes celulares são reforçadas com lignina. A maior parte da composição da parede secundária é de celulose, hemicelulose e lignina, que proporcionam maior resistência mecânica à célula [1,18,19]. A parede secundária tem maior recalcitrância, necessitando de pré-tratamentos para degradação da biomassa em escala industrial, aumentando o custo da produção de bioetanol [4].

Apesar da composição polimérica das paredes celulares ser bem conhecida, os modelos da arquitetura tridimensional da parede celular ainda são muito limitados. Segundo o modelo mais recente de paredes primárias, representado na figura 1.2, xiloglucanos conectam pequenas regiões de contato entre as fibrilas de celulose [20]. Entretanto, este modelo não é aplicado para paredes secundárias e nem para paredes primárias de gramíneas, que possuem baixa composição de xiloglucano [5].

1.3 Expansinas

Expansinas são proteínas que induzem relaxação mecânica da parede celular em plantas em crescimento. Elas estão diretamente ligadas a processos naturais das plantas, como amadurecimento de frutos, germinação do pólen e enraizamento [21,22]. Expansinas também causam intumescimento de fibras de celulose, podendo, assim, potencialmente ser utilizadas como proteínas acessórias às enzimas celulolíticas para reduzir a recalcitrância da celulose. Embora os efeitos

macroscópicos das expansinas sejam bem conhecidos, pouco se sabe sobre o mecanismo destas proteínas do ponto de vista molecular. A razão disso é que, segundo os estudos realizados até o momento, expansinas não agem através de um mecanismo bioquímico de hidrólise, o que dificulta a realização de experimentos [3,23,24].

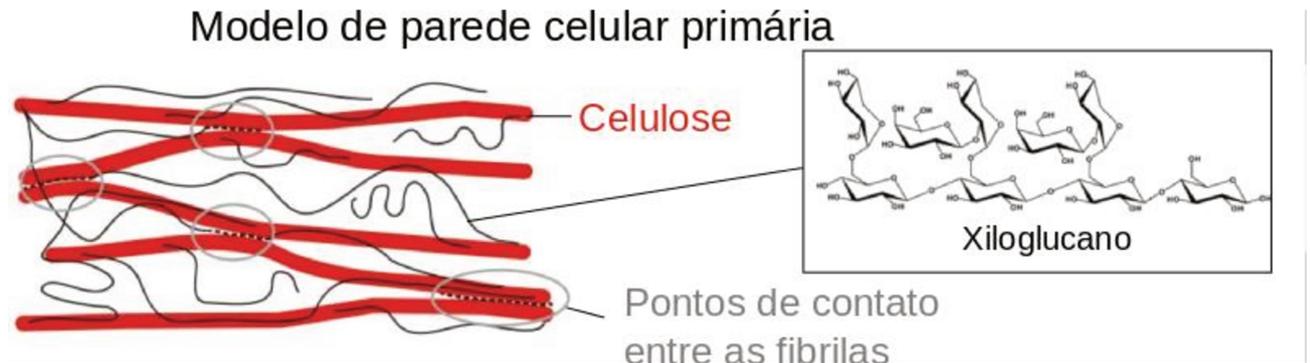


Figura 1.2 Representação do modelo proposto para a parede primária, onde há pequenos pontos de contato entre as microfibrilas de celulose, representadas em vermelho, mediados por xiloglucanos, representados pelas linhas pretas. A figura foi adaptada de Park and Cosgrove, 2011 [12].

Estudos funcionais com a expansina EXLX1 da bactéria *Bacillus subtilis* mostraram que um de seus dois domínios, o domínio D1, é essencial para sua atividade de ruptura [25]. Estudos computacionais da EXLX1, realizados em nosso grupo, indicam que o domínio D1 da EXLX1 pode induzir torções em cadeias de glucano [26]. Além disso, mutações experimentais dos resíduos responsáveis pela torção, identificados pelas simulações, causam inativação completa ou parcial da EXLX1, corroborando com a hipótese de que EXLX1 atua desestabilizando interações intermoleculares na parede celular [25,26].

As expansinas de plantas são classificadas em α e β -expansinas, que evoluíram para diferentes tipos de plantas. β -expansina de pólen de milho (ExpB1) atua apenas em gramíneas, o que faz com que ela seja de alto interesse para estudos voltados à bioenergia [22,27]. A estrutura cristalográfica da ExpB1 foi resolvida por Yennawar et al. em 2006 (Código PDB: 2HCZ), mas seu mecanismo molecular permanece elusivo e provavelmente envolve questões ainda não compreendidas da parede celular primária de plantas [28].

1.4 Enzimas da família GH45

Enzimas da família GH45 possuem alta similaridade com expansinas produzidas em plantas e foram pouco caracterizadas até o momento, se comparadas a outras enzimas produzidas por fungos [25,28]. As endoglucanases da família GH45 possuem atividade hidrolítica sobre β -glucanos, CMC (carboximetilcelulose) e PASC (do inglês, *Phosphoric Acid Swollen Celulose*). Entretanto, diferentemente das endoglucanases de outras famílias, enzimas da GH45 não possuem atividade hidrolítica sobre hemiceluloses, como xilanos e xiloglucanos e substratos artificiais, como pNPC (do inglês, *para-Nitrophenyl Cellobioside*) [29]. Além disso, possuem baixa atividade sobre cadeias menores que 5 monômeros [30].

Estas enzimas hidrolisam os substratos pelo mecanismo clássico de inversão de Koshland, esquematizado na figura 1.3A [31], o qual envolve a atuação de dois resíduos aspartato ou glutamato, sendo que um deles atua como ácido, doando um próton para o oxigênio glicosídico, e o outro atua como base, desprotonando uma molécula de água que irá fazer um ataque nucleofílico no carbono anomérico da cadeia glucana.

A enzima produzida no fungo *Humicola insolens* (HiCel45) foi a melhor caracterizada até o momento e a primeira enzima da família GH45 a ter sua estrutura resolvida por cristalografia de raio-X (código PDB: 4ENG) (figura 1.3B). Estudos realizados com a HiCel45 mostraram que os resíduos Asp121 e Asp10 atuam como ácido e nucleófilo catalíticos, respectivamente. Isso foi concluído com base em mutações destes resíduos para asparagina que anularam a atividade catalítica da enzima. Neste mesmo estudo, o mutante Asp114Asn, apresentou apenas uma diminuição da atividade catalítica e concluiu-se que ele atua como coadjuvante no processo da catálise. O ácido catalítico Asp121 da HiCel45 e os resíduos próximos a ele (Thr6, Tyr8, His119) são conservados nas outras enzimas da família GH45 e nas expansinas (figura 1.3B)[32]. Entretanto, não há conservação do resíduo nucleófilo Asp10 entre os membros da família GH45 [29].

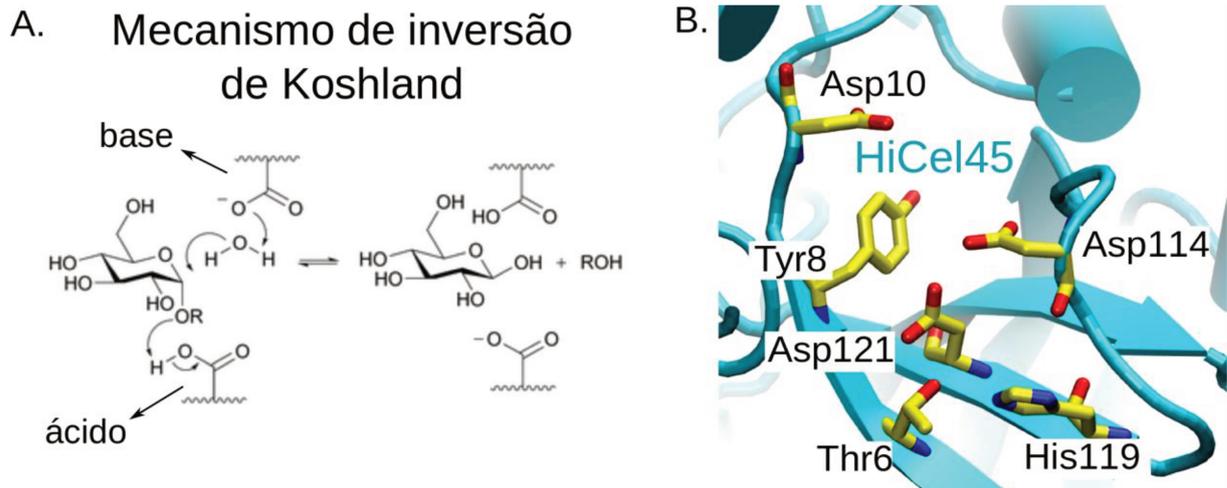


Figura 1.3 (A) Mecanismo de inversão de Koshland [31], mostrando que são necessário dois resíduos aspartatos para a reação catalítica. No final da catálise, o ácido está desprotonado e a base protonada, precisando então trocar os estados de protonação para voltar à forma ativa. (B) Estrutura cristalográfica da HiCel45 (código PDB:4ENG), indicando os resíduos catalíticos Asp121 e Asp10, o Asp114 que é coadjuvante no processo de catálise segunda Davies et al. 1995 [29] e os resíduos que estão em torno do ácido catalítico, Thr6, Tyr8 e His119, conservados nas outras enzimas da família GH45.

Comparando a estrutura cristalográfica da HiCel45 (código PDB: 4ENG) com a estrutura da ExpB1, mostrada na figura 1.4, o sítio catalítico da ExpB1 (código PDB: 2HCZ) é similar ao da HiCel45. Entretanto, a ExpB1 não possui o resíduo correspondente ao Asp10. Isso foi proposto como o motivo da ausência de atividade catalítica das expansinas. Além disso, a ExpB1 possui, em torno do sítio catalítico, loops muito menores que a HiCel45 (figura 1.4), fazendo com que sua fenda catalítica seja praticamente aberta e plana, resultando em uma interação menos efetiva com o substrato, quando comparada à HiCel45 [25,28,32].

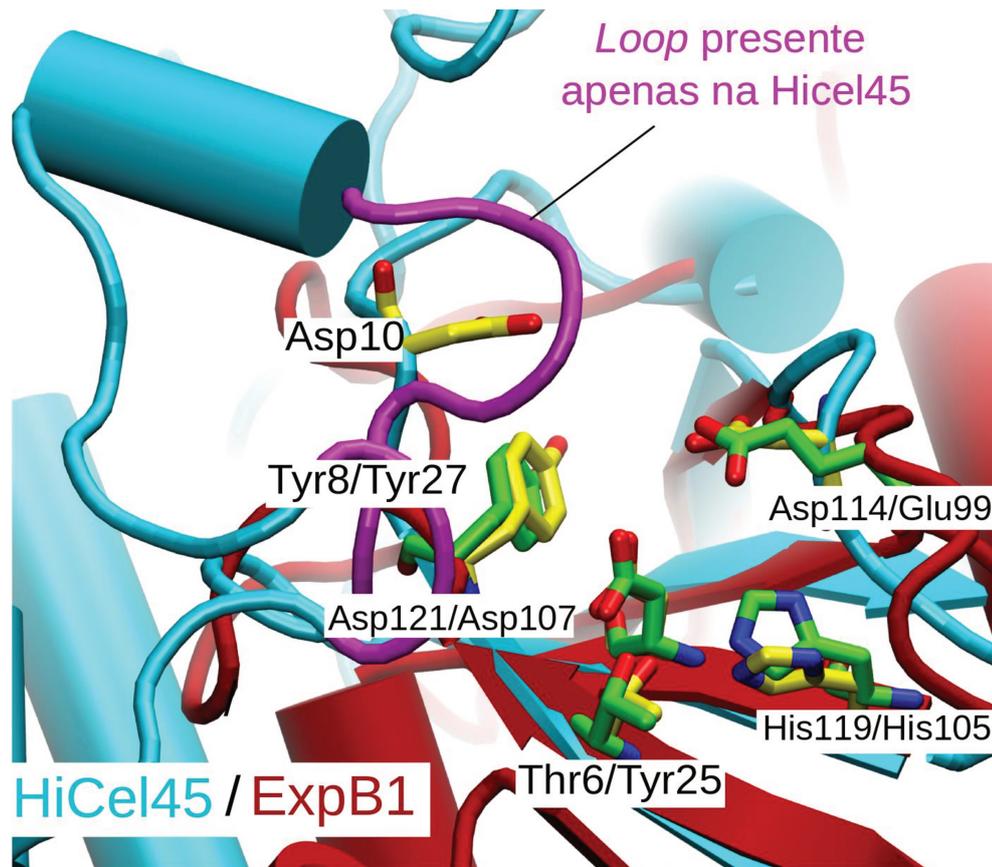


Figura 1.4. Alinhamento das estruturas cristalográficas da ExpB1 (código PDB: 2HCZ), com a cadeia principal mostrada na cor vermelho e cadeias laterais em cor verde, com a HiCel45 (código PDB: 4ENG), com a cadeia principal mostrada na cor azul e cadeias laterais em cor amarela. Mostrando que o sítio catalítico da ExpB1 é similar ao da HiCel45, entretanto, a ExpB1 não possui o resíduo correspondente ao resíduo base Asp10. O *loop* da HiCel45, em cor rosa, é responsável por manter o substrato ligado e a ExpB1 não possui este *loop*, apresentando assim uma fenda catalítica plana e rasa.

Capítulo 2

Dinâmica Molecular

A técnica de simulação de Dinâmica Molecular (DM) é um método computacional que estuda sistemas em nível atômico. Esta técnica permite a amostragem do espaço de fase e o desenvolvimento da dinâmica do sistema ao longo do tempo. Com isso, obtêm-se as conformações de menor energia potencial e parâmetros termodinâmicos do sistema, que muitas vezes não são detectáveis experimentalmente. Nos estudos de macromoléculas, esta técnica revela as estruturas e evolução temporal do sistema e, assim, pode ser vista como um microscópio molecular [33,34].

No modelo de DM mais utilizado, onde todos os átomos são explicitamente considerados (*“all-atom model”*), os átomos se movem obedecendo as leis da mecânica clássica e possuem carga e massa específica. Uma simulação de DM normalmente é realizada com número de partículas, volume e energia constantes, o que é chamado de ensemble NVE ou microcanônico e possui dimensão $6N$ para N átomos, pois são $3N$ posições e $3N$ velocidades. As configurações são geradas resolvendo a equação 2.0, da segunda lei de Newton, onde m_i é a massa da partícula e F_{x_i} a força ao longo da coordenada x_i , obtendo assim as posições e velocidades seguintes de cada partícula [35].

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i} \quad (2.0)$$

As forças que atuam sobre cada átomo são calculadas a partir das interações inter e intramoleculares, que serão descritas na seção 2.1. As simulações são geradas a partir das posições e velocidades iniciais, que devem entrar como dados para os cálculos das configurações seguintes. Na seção 2.2 é mostrado como é feita a escolha destas configurações iniciais. Para a geração das trajetórias, são utilizados alguns algoritmos e aproximações, introduzidos na seção 2.3, que diminuem custos computacionais e permitem estudar sistemas em diferentes condições, como temperatura e pressão constantes. Com a evolução do tempo, as simulações geram um número de configurações sucessivas, tornando possível a obtenção de

propriedades médias e de comportamentos dependentes do tempo, que é explicado na seção 2.4.

2.1 Campo de força

A técnica de DM requer como dado de entrada uma descrição das interações inter e intramoleculares. O modelo de “campo de força” é uma aproximação comum para tratamento destas interações. Neste modelo, as forças entre as partículas são determinadas empiricamente por princípios físicos, dados experimentais e parâmetros calculados por mecânica quântica. Por ser desenvolvido empiricamente, não é conveniente utilizar campos de força distintos em um mesmo sistema.

O potencial do campo de força pode ser considerado como a soma do potencial de ligação (V_{ligado}) e do potencial não-ligado ($V_{não-ligado}$), como é mostrado de forma simplificada nas equações 2.1-3 e representado na figura 2.1, onde $V_{total}(r^N)$ é a energia potencial em função das posições r dos N átomos do sistemas.

$$V_{total}(r^N) = V_{ligado}(r^N) + V_{não-ligado}(r^N) \quad (2.1)$$

$$V_{ligado}(r^N) = \sum_{lig} k_{ij}(r_{ij} - r_0)^2 + \sum_{âng} k_{ijk}(\theta_{ijk} - \theta_0)^2 + \sum_{died} \frac{v_n}{2}(1 + \cos(n\phi_{ijkl} - \phi_0)) \quad (2.2)$$

$$V_{não-ligado}(r^N) = \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \quad (2.3)$$

O $V_{ligado}(r^N)$ modela as interações intramoleculares, como mostrado na equação 2.2, onde primeiro, segundo e terceiro termos são referentes, respectivamente, ao potencial harmônico da ligação de dois átomos (figura 2.1A), deformação angular (figura 2.1B) e torsões diedrais (figura 2.1C). A modelagem das interações é feita através da atribuição das penalidades energéticas causadas por desvios dos valores de equilíbrio. Com isso é possível descrever estiramentos e efeitos estéricos, por exemplo.

O $V_{não-ligado}(r^N)$, definido pela equação 2.3, considera as interações entre os átomos que não estão ligados covalentemente entre si. O primeiro termo da equação 2.3 modela as interações de van der Waals (figura 2.1D), utilizando o potencial de Lennard-Jones, onde σ é a distância interatômica em que a energia é zero e ϵ é a

profundidade do potencial, mostrados na figura 2.2. O segundo termo da equação 2.3 modela as interações eletrostáticas, utilizando o potencial de Coulomb, onde ϵ_0 corresponde à permissividade do vácuo, q_i e q_j às cargas das partículas i e j , como mostrado na figura 2.1E.

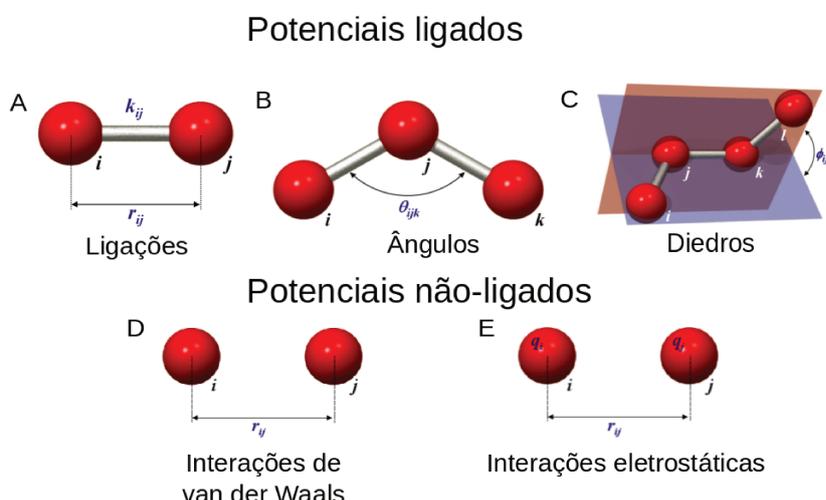


Figura 2.1. Representação das principais interações utilizadas para modelagem de campos de força, sendo os potenciais, referentes às interações intramoleculares, de ligações, ângulos e diedros mostrados, respectivamente, por (A), (B) e (C). Os potenciais não-ligados das interações van der Waals e eletrostáticas são mostrados em (D) e (E), respectivamente.

O potencial de Lennard-Jones possui um termo referente às forças repulsivas e outro às atrativas e variam com r^{-12} e r^{-6} , respectivamente. As forças repulsivas são mais pronunciadas quando a distância entre os átomos é igual ou ligeiramente menor à soma de seus raios. Este potencial modela bem as interações de van der Waals, que ocorrem a curta distância e rapidamente decaem conforme aumenta a distância entre os átomos [35].

Diversos potenciais efetivos clássicos foram desenvolvidos e aplicados no tratamento das interações intermoleculares. Dentre eles, destacam-se os campos de força CHARMM [36] e AMBER [37], que podem ser utilizados para cálculos em grande escala de biomoléculas utilizando tratamento explícito das moléculas do solvente (e.g., água, compostos orgânicos, etc). A transferibilidade destes campos de força é a propriedade que permite que parâmetros desenvolvidos para moléculas pequenas, como aminoácidos, possam ser aplicadas para moléculas maiores, como proteínas [35].

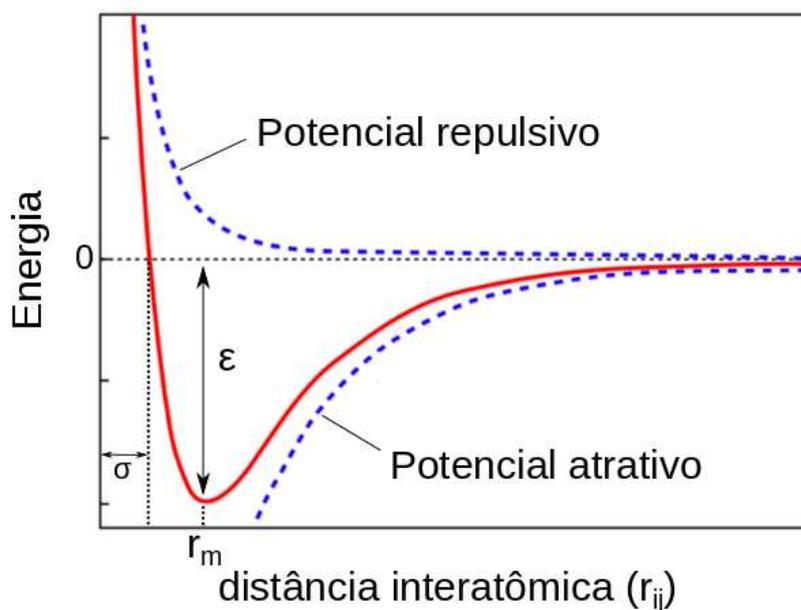


Figura 2.2. Representação do potencial de Lennard-Jones, indicando os potenciais atrativos, potenciais repulsivos e os coeficientes da equação do potencial de Lennard-Jones, σ e ε , que representam a distância interatômica em que a energia é zero e a profundidade do potencial, respectivamente.

2.2 Posições e velocidades iniciais

As posições e velocidades iniciais são atribuídas para todos os átomos do sistema no início da simulação e serão utilizados para determinação das configurações seguintes. As posições dos átomos iniciais devem ser compatíveis com a estrutura que se pretende estudar, como estruturas determinadas experimentalmente por difração de raios-X. As partículas também não devem ter sobreposição entre elas e para isso existem alguns softwares, como o *packmol* [38], que permitem geração da configuração inicial.

Já a velocidade inicial do sistema é escolhida de forma aleatória para a temperatura de interesse. Para isto é utilizada a distribuição de Maxwell-Boltzmann, descrita na equação 2.4, que fornece a probabilidade (p) de uma partícula de massa m_i estar com velocidade v_{ix} na direção x . Depois da distribuição aleatória, as velocidades são ajustadas de tal forma que o momento total do sistema seja 0 em todas as direções [35].

$$p(v_{ix}) = \left(\frac{m_i}{2\pi k_B T} \right)^{\frac{1}{2}} \exp \left[-\frac{1}{2} \frac{m_i v_{ix}^2}{k_B T} \right] \quad (2.4)$$

2.3 Algoritmos

2.3.1 Condições periódicas de contorno

O uso de condições periódicas de contorno faz com que as forças aplicadas sobre as partículas simulem um sistema infinito (i.e., infinitamente periódico), que se aproxima das condições reais e elimina em parte os efeitos de superfícies. Além disso, permite a utilização de um número menor de moléculas, diminuindo custos computacionais. A figura 2.3 representa um modelo 2D de condição periódica de contorno. Na prática, este modelo é 3D e composto de caixas retangulares. Sendo todas as caixas iguais, elas permanecem com número constante de partículas, quando uma partícula sai da caixa, ela é substituída por uma partícula no lado oposto [35].

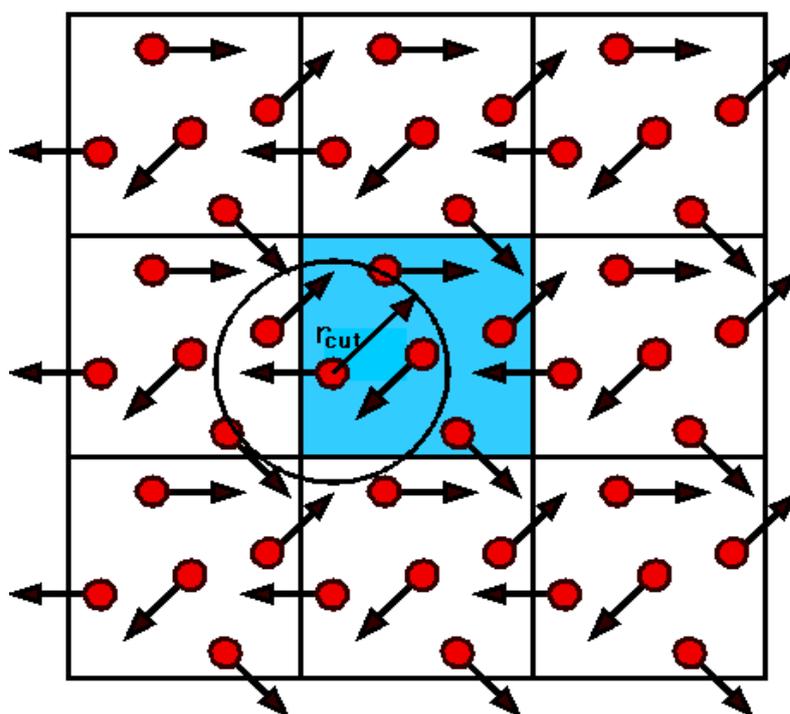


Figura 2.3. Representação 2D do modelo de condição periódica de contorno, onde as nove caixas quadradas são iguais. O r_{cut} indica o raio de corte aplicado para a interação entre dois átomos, mostrando que o átomo pode interagir com átomos da caixa vizinha.

2.3.2 Raio de corte e soma de Ewald

O cálculo das interações entre os átomos não-ligados é responsável pela maior parte do custo computacional em simulações por DM. Se as interações forem calculadas para todos os pares de átomos do sistema, o número de interações de van der Waals e eletrostáticas aumentam proporcionalmente ao quadrado do número de átomos. Com isto, o uso de métodos para diminuir o número dos cálculos necessários para expressar estas interações nas simulações, faz com que os cálculos sejam muito mais rápidos.

As interações de Lennard-Jones são de curto alcance, pois possuem uma dependência de r^{-6} , sendo r a distância interatômica. Para distâncias altas, este tipo de interação é insignificante, o que viabiliza o emprego de um raio de corte, como mostrado na figura 2.3, tal que os átomos que estão para fora deste raio não são considerados nos cálculos. Para simulações de proteínas é comum utilizar um raio de corte de 12 Å. Entretanto, este método não pode ser utilizado indiscriminadamente para as interações eletrostáticas [35].

As interações eletrostáticas são consideradas de longo alcance, pois variam diretamente com o inverso da distância interatômica e decaem lentamente. Para diminuir os custos computacionais, o método de soma de Ewald é utilizado para o cálculo destas interações. Este método considera que os átomos são cargas pontuais no sistema e faz uso das condições de periódicas de contorno e da eletroneutralidade do sistema.

O cálculo da contribuição coulômbica total do sistema de N partículas é feito pelas equações 2.5 e 2.6, em que $\Phi(r_i)$ é o potencial eletrostático na posição r_i , L é o diâmetro da caixa e r_{ij} é a distância entre as partículas i e j na mesma caixa. A soma na equação 2.6 é feita sobre todas as n imagens periódicas, exceto para $j = i$ se $n = 0$. Com isso, assumimos que a partícula i interage com todas as outras partículas do sistema e suas imagens periódicas, menos com si mesma. A descrição completa deste método está na referência [35].

$$V = 1/2 \sum_{i=1}^N q_i \Phi(r_i) \quad (2.5)$$

$$\Phi(r_i) = \sum_{j,n} \frac{q_j}{4\pi\epsilon_0 |r_{ij} + nL|} \quad (2.6)$$

2.3.3 Método de Verlet

Para gerar as trajetórias é utilizado um método numérico, em que as configurações, velocidades e acelerações são calculadas a cada intervalo de tempo fixo (δt), durante o qual a força é considerada constante. Em simulações de biomoléculas, geralmente, é considerado δt de 2 femtosegundos.

O algoritmo de Verlet utiliza as posições e acelerações no tempo t e no tempo $t - \delta t$, que é referente ao passo anterior, para calcular a posição e aceleração do passo seguinte utilizando as equações 2.7 e 2.8.

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + \dots \quad (2.7)$$

$$r(t - \delta t) = r(t) - \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + \dots \quad (2.8)$$

Somando as equações 2.7 e 2.8, obtem-se a equação 2.9, que é utilizada para calcular o passo seguinte da trajetória, referente à $r(t + \delta t)$.

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \quad (2.9)$$

O maior problema deste algoritmo, para simulações de DM, é que as velocidades não aparecem explicitamente e a temperatura, por exemplo, é dependente diretamente da velocidade das partículas. Além disso, considerando que no início da trajetória, em $t = 0$, não se tem os dados do passo anterior $r(t - \delta t)$, este método então é suficiente para dar início à trajetória, necessitando de outros recursos para isso.

O método *velocity Verlet* foi desenvolvido para solucionar estes empecilhos. Nesse método, as velocidades passam a ser explícitas, com as equações 2.10 e 2.11, sem que a precisão seja comprometida. É um dos métodos mais amplamente utilizados para geração de trajetórias de DM.

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) \quad (2.10)$$

$$v(t + \delta t) = v(t) + \frac{1}{2} \delta t [a(t) + a(t + \delta t)] \quad (2.11)$$

Para solucionar estas equações são necessários 3 passos. Primeiramente se utiliza a equação 2.10 para calcular as posições em $t + \delta t$, a partir das velocidades e acelerações do tempo t . Em segundo, as forças e acelerações são calculadas no tempo $t + \delta t$. Por último, as velocidades no tempo $t + \delta t$ são calculadas utilizando a equação 2.11. Neste algoritmo, as posições anteriores não são necessárias para se calcular a próxima configuração [35].

2.3.4 Temperatura e pressão

Ao longo da simulação de um sistema microcanônico NVE as velocidades flutuam e com elas também flutuam a energia cinética e a temperatura. A temperatura instantânea do sistema pode ser determinada pela equação 2.12 [38].

$$T(t) = \sum_{i=1}^N \frac{m_i v_i^2(t)}{2k_B N_f} \quad (2.12)$$

Entretanto, em simulações de biomoléculas é interessante que a temperatura e pressão se mantenham constantes, buscando aproximar-se de condições experimentais. Para manter a temperatura constante foi utilizada neste trabalho a dinâmica de Langevin, que é uma equação diferencial estocástica que permite controlar a energia cinética do sistemas, e assim a temperatura, como um termostato. A pressão, por sua vez, foi controlada em nossas simulações utilizando o método do pistão de Langevin. Este método atua como um pistão que flutua alterando as dimensões da caixa e mantendo a pressão constante. Estes métodos estão descritos nas referências [40-42].

2.4 Cálculo de propriedades médias

Um dos principais interesses de simulações de DM é estudar propriedades médias do sistemas através da evolução natural do tempo. Considerando a propriedade termodinâmica A de um sistema de N átomos, seu valor instantâneo irá depender dos valores da posição r e do momento p de cada átomo do sistema. Ao longo do tempo, ela flutua dependendo da interação entre as partículas. Para experimentos laboratoriais, é considerado que o valor de uma propriedade A é a média de ensemble $\langle A \rangle_{ensemble}$, dada pela equação 2.13, integrada sobre todas as possíveis posições e momentos.

$$\langle A \rangle_{ensemble} = \int \int A(p^N(t), r^N(t)) dr^N dp^N \quad (2.13)$$

Devido a impossibilidade de determinar todas as configurações iniciais, sistemas estudados com DM são considerados ergódicos. A hipótese da ergodicidade estabelece que uma única simulação de DM, suficientemente longa, pode ser utilizada para determinar uma propriedade termodinâmica macroscópica do sistema. Com isso, mostrado na equação 2.14, se tem que a média de um ensemble microcanônico corresponde à média temporal $\langle A \rangle_{temporal}$ de um sistema ergódico [38].

$$\langle A \rangle_{ensemble} = \langle A \rangle_{temporal} = \lim_{\tau \rightarrow \infty} 1/\tau \int_{t=0}^{\tau} A(p^N(t), r^N(t)) dt \quad (2.14)$$

O cálculo de $\langle A \rangle_{temporal}$ é feito considerando um tempo infinito, onde se considera que todos os microestados possíveis são visitados. Evidentemente isto não é válido se existirem barreiras energéticas que separem os mínimos de energia. Neste caso, é necessário recorrer a outros métodos de amostragem. No caso de uma simulação de DM, o número de configurações obtidas é finito e igual ao número de passos M da simulação, mas pode ser considerado para o cálculo de propriedades médias do sistema $\langle A \rangle_{DM}$, descrita na equação 2.15, se a simulação for longa [35].

$$\langle A \rangle_{DM} = 1/M \sum_{i=1}^M A(p^N(t), r^N(t)) \quad (2.15)$$

Várias análises com dependência temporal também podem ser realizadas, como a variação das distâncias, ligações de hidrogênio e energia de interação entre dois grupos moleculares, para citar alguns poucos exemplos. Além disso, podem ser realizadas análises estruturais, como a análise de Desvio da Raiz Média Quadrática (RMSD, sigla em inglês), definida pela equação 2.16, em que se calcula a variação média da distância ao longo do tempo para cada átomo do conjunto n , com relação à sua posição em uma configuração definida $r_{i,o}$, que pode ser a configuração inicial.

$$RMSD(t) = \left[\frac{1}{n} \sum_{i=1}^n (r_i(t) - r_{i,o})^2 \right]^{1/2} \quad (2.16)$$

Essas análises são dependentes do sistema estudado e dos objetivos do trabalho. Normalmente são feitas utilizando programas desenvolvidos dentro do próprio grupo.

Capítulo 3

Adsorção de cadeias xilanas à superfície hidrofílica da celulose

3.1 Introdução

Xilano é uma das hemiceluloses mais abundantes em paredes secundárias de plantas. Sua estrutura é constituída por unidades de D-xilose ligadas por ligações β -1,4 (figura 3.1) [12]. A xilose se diferencia da glicose pela falta do grupo hidroximetílico exocíclico, o que a faz apresentar um número menor de ligações de hidrogênio intramoleculares e maior mobilidade dos ângulos diedrais glicosídicos ψ e ϕ , quando comparada à cadeia glucana (figura 3.1) [43]. Os xilanos encontrados na natureza possuem acetilações ou ramificações de ácido glucurônico (GlcA), ácido 4-O-metilglucurônico (MeGlcA) e arabinose (Ara) (figura 3.1). Estas ramificações ou acetilações são dependentes da espécie e do tipo de tecido da planta e influenciam na interação do xilano com água e polímeros da parede celular, modulando, assim, as propriedades da parede celular [12].

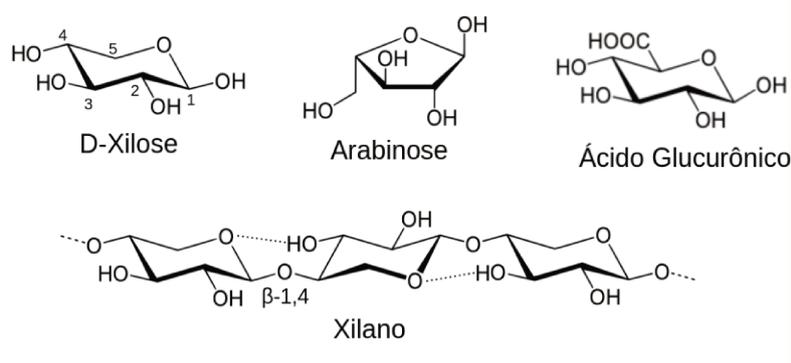


Figura 3.1. Representação molecular da cadeia xilana. D-xilose, monômero que constitui a cadeia xilana, com a numeração de seus carbonos; Arabinose (Ara) e ácido glucurônico (GlcA), monômeros encontrados como ramificações de cadeias xilana e plantas. Xilano, mostrando as ligações de hidrogênio intramoleculares e as ligações β -1,4.

As ramificações de Ara são relacionadas com propriedades de hidratação da planta e para prevenir agregação dos polímeros da parede celular com a perda de

água. Experimentos *in vitro* por medidas de isotermas de adsorção mostraram que ramificações de Ara diminuem interações do xilano com a superfície da celulose [44]. Já plantas mutadas para produzir xilanos sem ramificações de GlcA apresentaram paredes celulares do caule ligeiramente mais fracas que as da nativa. A extração de xilanos nestas mutantes também ocorreu mais facilmente. Isto indica que ramificações de Ara e GlcA estão relacionadas com a interação do xilano com os outros polímeros da parede celular, inclusive com a própria celulose [45].

O arranjo molecular de xilanos na parede celular é pouco descrito: sabe-se apenas que eles são um dos responsáveis pela recalcitrância. Seu modo de interação com outros polímeros da parede celular é uma questão ainda em aberto. Uma hipótese é que xilanos recobrem as microfibrilas da celulose e interagem com outros polímeros, através das ramificações, por ligações de hidrogênio e ligações covalentes [19,46]. Devido às interações hidrofóbicas entre anéis glicosídicos, a interação do xilano com a face hidrofóbica da celulose é evidente. Estudos de DM indicaram que a adsorção do xilano nas faces (100) e (200) de microfibrilas de celulose é estável (figura 3.2) [16].

Acredita-se que uma cadeia xilana com ramificações em todos os resíduos não poderia se ligar na face hidrofílica da celulose, devido a efeitos estéricos e à falta de hidroxilas para fazer ligações de hidrogênio com a celulose. Mas, se o padrão de ramificação for alternado, apenas com os resíduos pares ramificados, uma cadeia xilana com conformação dupla-hélice ficaria com as ramificações em apenas uma superfície e a outra permaneceria disponível para interagir com a celulose, sendo compatível com sua adsorção na superfície hidrofílica de celulose [16].

Corroborando com a hipótese da necessidade de ramificações em apenas um lado da cadeia xilana para adsorção na face hidrofílica, a determinação do padrão de ramificação de xilanos em paredes secundárias de angiospermas mostram que a maior parte da cadeia xilana é formada por domínios onde resíduos alternados são acetilados e as cadeias laterais MeGlcA são distribuídas com o padrão de espaçamento par [16, 47,48].

Estudos de DM realizados anteriormente com xilanos com padrão de ramificação e acetilação alternadas mostraram que estas cadeias podem ligar na face hidrofílica da celulose (figura 3.2). Entretanto, estas simulações foram modeladas nas superfícies (020) e (010) da microfibrila retangular (figura 3.2), onde as duas faces da cadeia xilana estão sujeitas às fortes interações de empilhamento estabelecidas com

as cadeias de glucano [16]. Com isso, uma questão permaneceu em aberto, se nas faces hidrofílicas (110) e (110) de uma fibrila hexagonal, onde não existe sítio de ligação específico, o xilano poderia se ligar de forma estável.

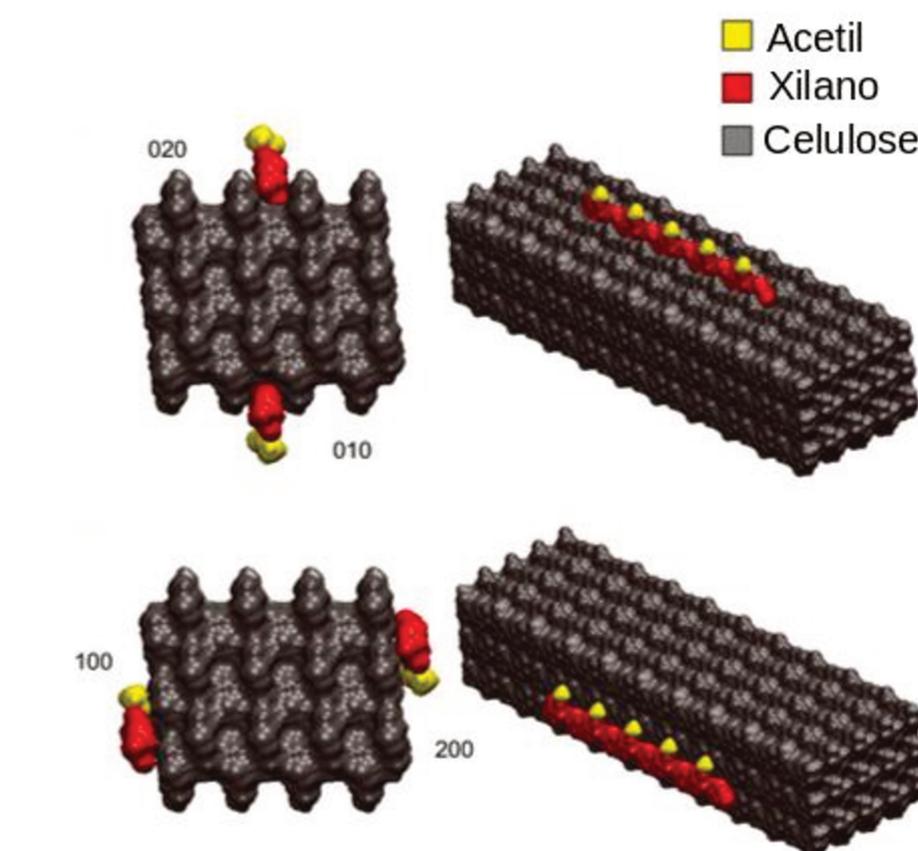


Figura 3.2. (A) Modelo utilizado nos estudos anteriores de DM da adsorção da cadeia xilana acetilada nas faces hidrofílicas (010) e (020) e hidrofóbicas (100) e (200) da celulose. [16]

Recentemente, o padrão de ramificação de xilanos em paredes secundárias de coníferas foi determinado por colaboradores (Paul Dupree, University of Cambridge). Esses xilanos seguem um padrão de ramificações de α -1,2-ácido 4-O-metilglucurônico a cada seis resíduos e α -1,3-arabinose precisamente a dois resíduos do ácido (dados não publicados) (figura 3.3). Na conformação diedral dupla-hélice, a cadeia xilana com este padrão de ramificação possui todas as ramificações em apenas um lado, permitindo assim, segundo a hipótese proposta, sua adsorção na face hidrofílica da celulose.

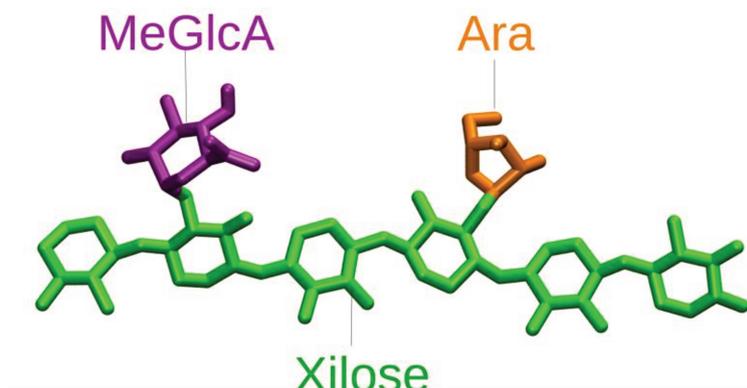


Figura 3.3. (A) Representação do padrão de ramificação encontrado em cadeias xilanas de coníferas, com ramificação α -1,2-ácido 4-O-metil-glucurônico a cada seis resíduos e α -1,3-arabinose a dois resíduos. A cadeia xilana está na conformação dupla-hélice para mostrar que todas as ramificações estão de um mesmo lado.

Neste trabalho, estudamos por simulações de DM, a compatibilidade da cadeia xilana com ramificações α -1,3-Ara e α -1,2-GlcA (glucuronoarabinoxilano), no padrão encontrado para coníferas, de se ligar na face hidrofílica (110) da celulose. Para avaliar os efeitos das ramificações de α -1,3-Arab e α -1,2-GlcA conduzimos simulações de DM da adsorção de xilano sem ramificação, xilano com ramificações apenas α -1,2-GlcA e com ramificações apenas de α -1,3-Ara (arabinoxilano). Com isso, investigamos também se xilanos adsorvem na face hidrofílica (110) da celulose. Para analisar a possibilidade da adsorção de duas cadeias xilana em paralelo na face hidrofílica da celulose e os possíveis efeitos estéricos das ramificações, estudamos dois sistemas diferentes: um deles com duas cadeias de xilano sem ramificação e o outro com duas cadeias de glucuronoarabinoxilano.

3.2 Métodos

Neste trabalho conduzimos simulações de DM, pelo método descrito a seguir, da adsorção na face hidrofílica da celulose (110) (figura 3.4A) das cadeias xilana: xilano, glucuronoarabinoxilano, arabinoxilano e glucuronoxilano (figura 3.5) para estudar a interação destas cadeias com a celulose e os efeitos das ramificações. A cadeia de glucuronoarabinoxilano possui padrão representativo de xilano de coníferas, faltando apenas a metilação da ramificação GlcA que desconsideramos, visto que não traria diferenças significativas para o estudo realizado e adicionaria

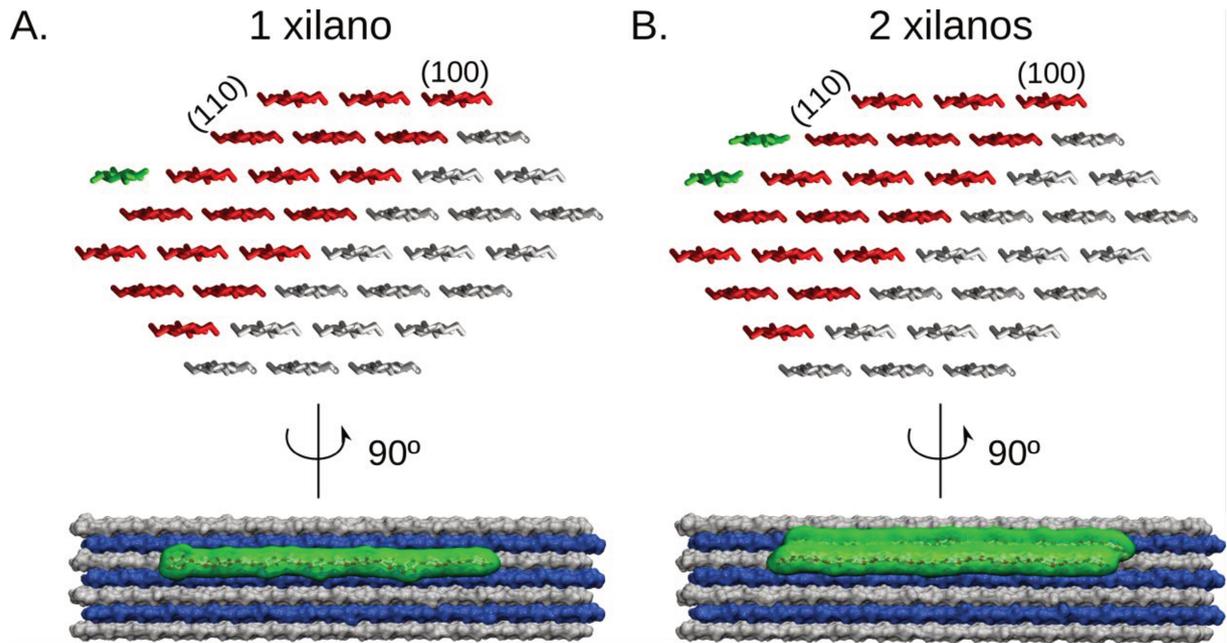


Figura 3.4. Representação da modelagem para simulações de DM dos sistemas com (A) 1 xilano e (B) 2 xilanos. Figura acima, em vermelho mostramos a parte da fibrila utilizada nas simulações, em branco a parte da fibrila desconsiderada nas simulações e em verde estão indicadas as cadeias de xilana nas posições iniciais e que foram obtidas a partir da extensão da própria celulose. Figura abaixo, representação da face hidrofílica da celulose com xilano: em azul e cinza mostramos as diferentes camadas de celulose e em verde as cadeias xilana.

etapas de parametrização do MeGlcA para realização do trabalho (ver seção 3.3.3). Também conduzimos simulações de DM de duas cadeias de xilano e glucuronoarabinoxilano sobre a face hidrofílica da celulose (110) (figura 3.4B), para avaliar os efeitos cooperativo xilano-xilano e estéricos das ramificações.

Na construção dos complexos de celulose-xilano, usamos Cellulose-Builder [49] para obter a estrutura de uma fibrila elementar hexagonal de celulose contendo 36 cadeias de glucano de DP = 22 (grau de polimerização) no polimorfo I β [50,51]. Em seguida, replicamos uma cadeia glucana e a transpomos para a superfície hidrofílica (110) da celulose. Removemos os grupos hidroximetílicos desta cadeia glucana para transformá-la em uma cadeia xilana. Consideramos apenas os 14 resíduos centrais da cadeia xilana para reduzir os efeitos de borda provenientes das fibrilas de celulose. Em seguida, cortamos a fibrila de celulose ao meio, paralelamente ao plano (110), de modo a considerar apenas as cadeias próximas à face onde o xilano está adsorvido (figura 3.4A). O sistema foi imerso numa caixa contendo moléculas explícitas de água, tal que uma camada de solvente de pelo 16 Å de espessura tenha se formado ao redor do complexo em todos os lados. Este mesmo procedimento foi utilizado para construir

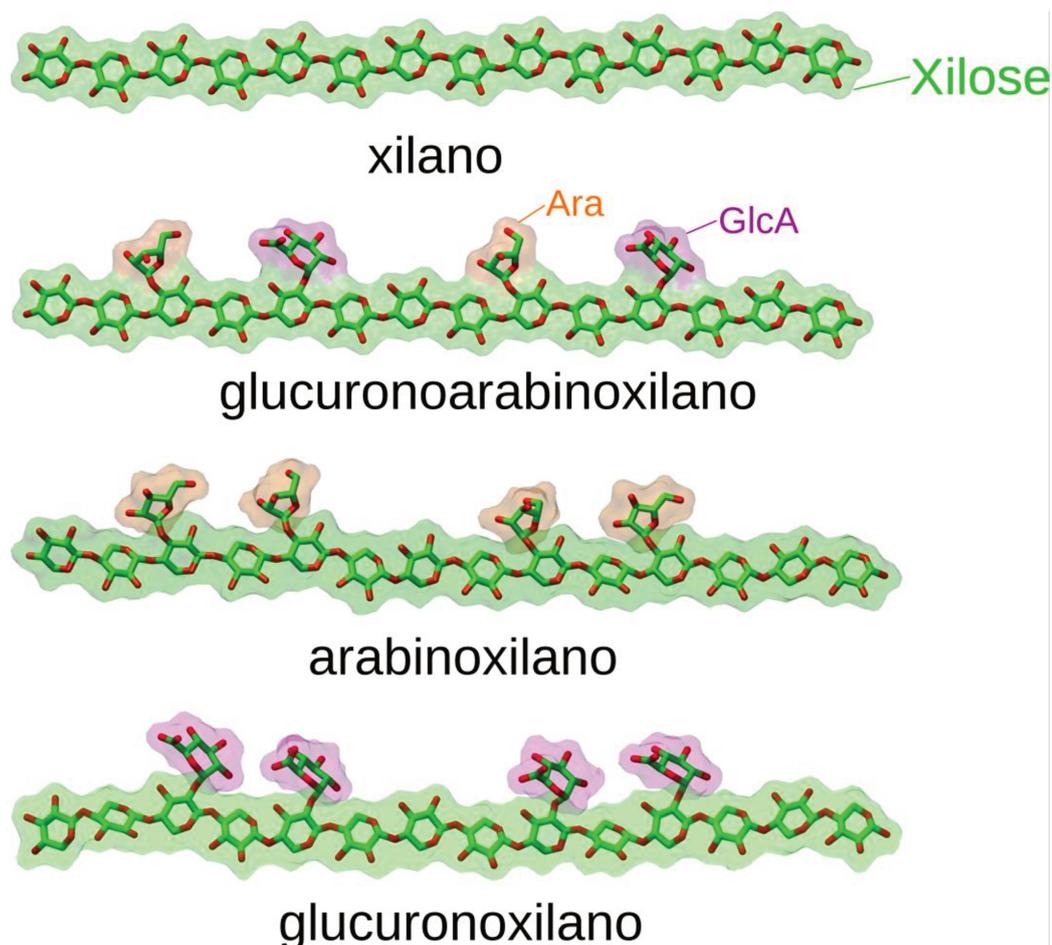


Figura 3.5. Representação das cadeias xilana utilizadas neste trabalho, na ordem de cima para baixo estão xilano (sem ramificações), glucuronoarabinoxilano (ramificada com padrão encontrado em coníferas com ligações α -1,3-Ara e α -1,2-GlcA), arabinoxilano (ramificada apenas com ligações α -1,3-Ara) e glucuronoxilano (ramificada apenas com α -1,2-GlcA).

os complexos de glucuronoarabinoxilano, arabinoxilano e glucuronoxilano com celulose. Para obtenção do glucuronoarabinoxilano introduzimos as substituições de GlcA, no estado desprotonado, e Ara de acordo com o seguinte padrão: α -1,2-GlcA (4), α -1,3-Ara (6), α -1,2-GlcA (10) e α -1,3-Ara (12) (o número entre parênteses representa a posição ao longo da cadeia de xilano). Para o glucuronoxilano e arabinoxilano, introduzimos substituições apenas de α -1,2-GlcA e α -1,3-Ara, respectivamente, nas posições 4, 6, 10 e 12 (figura 3.5). Adicionamos um íon cálcio no sistema de glucuronoarabinoxilano-celulose e dois íons cálcio no sistema glucuronoxilano, para manter o sistema eletricamente neutro. Utilizamos as coordenadas internas de campo de força CHARMM36 [52,53] para adicionar as substituições. Utilizamos este mesmo procedimento para os sistemas de dois xilanos, sem ramificações e sem íon cálcio, e dois glucuronoarabinoxilanos, com o mesmo

padrão de ramificação descrito anteriormente e com dois íons cálcio.

Executamos as simulações empregando condições de contorno periódicas, tratamento com malha de partículas de Ewald (PME - sigla em inglês) [54] para as interações electrostáticas e raio de corte de 12 Å para as interações de curto alcance. Utilizamos passo de integração de 2 fs, usado para resolver numericamente as equações do movimento, e truncamos todas as ligações químicas envolvendo átomos de hidrogênio em seus comprimentos de equilíbrio. Mantivemos a temperatura em 300 K e a pressão em 1 atm utilizando termostato e pistão de Langevin, respectivamente [40]. Para evitar distorções da celulose I β , mantivemos os átomos C1, C2, C3, C4, C5 e O5 dos anéis de glicose da celulose harmonicamente presos com uma constante de força de 50 kcal/mol.Å². Não adicionamos potencial sobre os grupos hidroximetílicos (C6).

Antes da produção das trajetórias, minimizamos e equilibramos o sistema pelas seguintes etapas: (1), com apenas os átomos das substituições e solventes livres, 100 passos de minimização de energia, utilizando gradientes conjugados seguido de 500 ps da DM; (2) o mesmo que (1), mas apenas com os átomos de celulose fixos (como descrito acima). Após estas etapas, produzimos simulações de DM de 200 ns de tempo real, sobre as quais fizemos as análises. Produzimos três simulações independentes a partir do passo (1) utilizando NAMD [40] com o campo de força CHARMM36 [52,53] e o modelo de água TIP3P [55]. As análises foram realizadas com VMD [56] e por programas criados pelo próprio grupo.

3.3 Resultados

3.3.1 Interação das cadeias xilana com a face hidrofílica da celulose

Para analisar a compatibilidade de uma cadeia xilana, com o padrão de ramificação encontrado em coníferas, de ligar na superfície hidrofílica (110) da celulose e também os efeitos das ramificações na interação das cadeias xilana com a celulose, geramos simulações em triplicata de quatro sistemas diferentes: xilano-celulose, arabinóxilano-celulose, glucuronóxilano-celulose e glucuronoarabinóxilano-celulose. Analisamos estas simulações segundo a mobilidade das cadeias xilana, a força das interações xilose-celulose e quanto ao modo de ligação destas cadeias sobre a celulose.

Através do monitoramento do RMSD das xiloses das cadeias xilana, com relação à posição dessas no início da trajetória (figura 3.4A), analisamos a mobilidade global em função do tempo para as três simulações de cada sistema (figura 3.6A). O xilano e o arabinoxilano apresentaram RMSD máximo em torno de 25 Å e não se estabilizaram, pois mantiveram alta variação do RMSD durante toda a simulação. O glucuronoarabinoxilano apresentou elevação do RMSD nos primeiros 50ns, chegando ao máximo de 10 Å em uma das simulações. Após este período a cadeia se estabilizou e o RMSD permaneceu flutuando em torno desse valor. O glucuronoxilano apresentou a menor mobilidade entre os quatro xilanos estudados, com o RMSD variando em torno de 3 Å na maior parte das simulações.

Esta análise mostrou que as cadeias xilana ramificadas com GlcA (glucuronoxilano e glucuronoarabinoxilano) adsorvem na face hidrofílica da celulose com maior estabilidade, quando comparadas com o xilano e arabinoxilano. Inclusive, durante uma das simulações do complexo xilano-celulose, o xilano se desloca completamente para a face hidrofóbica (100) da celulose, como mostra a figura 3.6B, sugerindo baixa estabilidade desta cadeia na interação com a face hidrofílica.

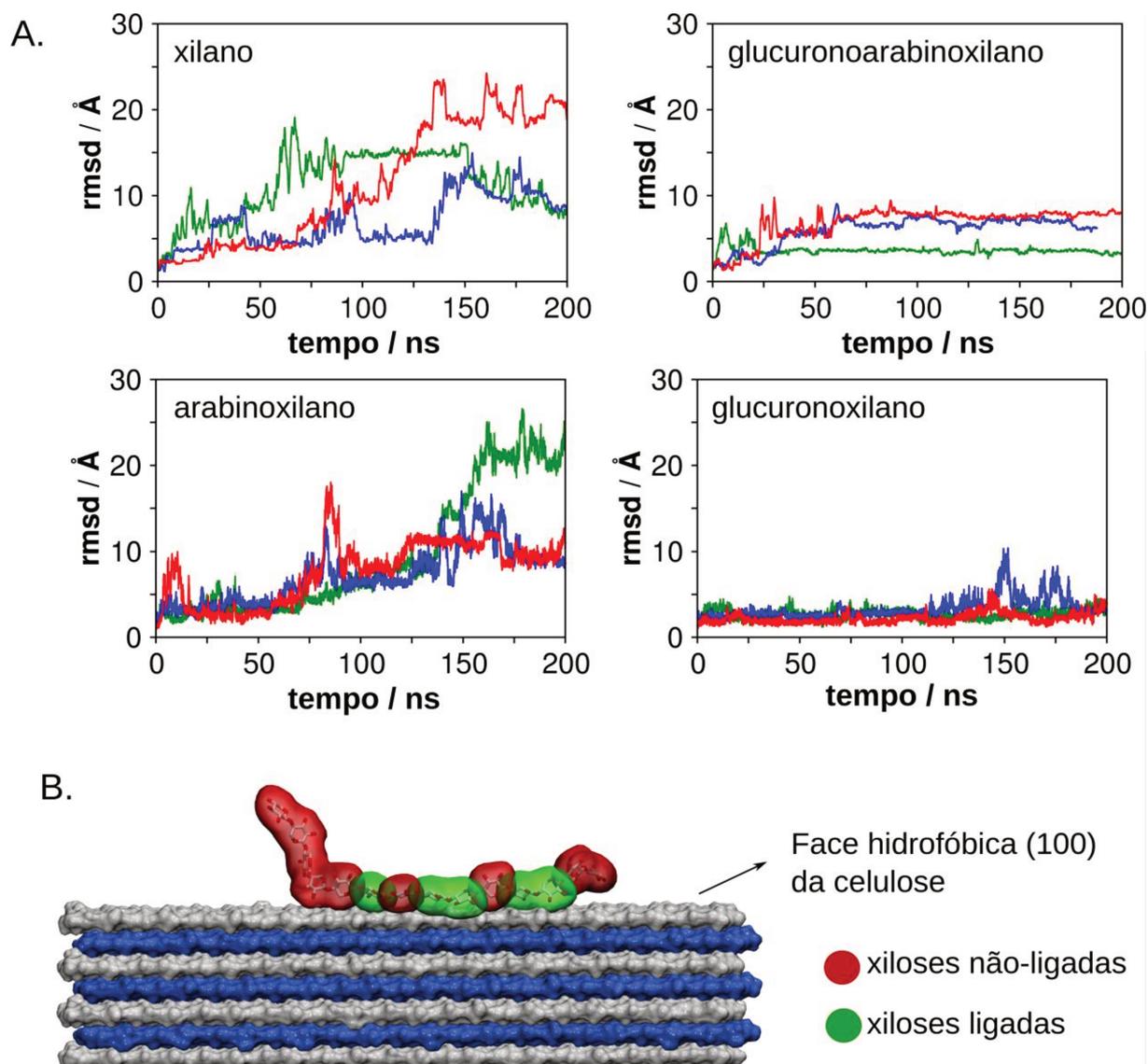


Figura 3.6. Análise da mobilidade das cadeias xilanas. (A) RMSD das cadeias xilanas. Cores diferentes indicam diferentes simulações. (B) Pose da simulação do complexo xilano-celulose em que o xilano se desloca completamente para a face hidrofóbica (100) da celulose.

A força da interação da cadeia xilana com a celulose está relacionada à flutuação do número de xiloses adsorvidas na superfície da celulose. Assim, para quantificar esta força de interação, computamos o número de resíduos de xilose que, em um dado momento, contacta a superfície da celulose. Consideramos que a xilose está adsorvida quando a distância mínima xilose-celulose é $2,5 \text{ \AA}$, como exemplificado na figura 3.7A, que mostra uma pose da simulação do complexo xilano-celulose com apenas uma xilose adsorvida sobre a celulose. O critério de distância de $2,5 \text{ \AA}$ foi escolhido considerando que esta é uma distância em que não há espaço para moléculas de água entre a celulose e a cadeia xilana, dado que a entrada de

água entre a celulose e a cadeia xilana caracteriza um início do processo de dessorção. A porcentagem de xiloses adsorvidas, computada em função do tempo para os quatro sistemas estudados está mostrada na figura 3.7B.

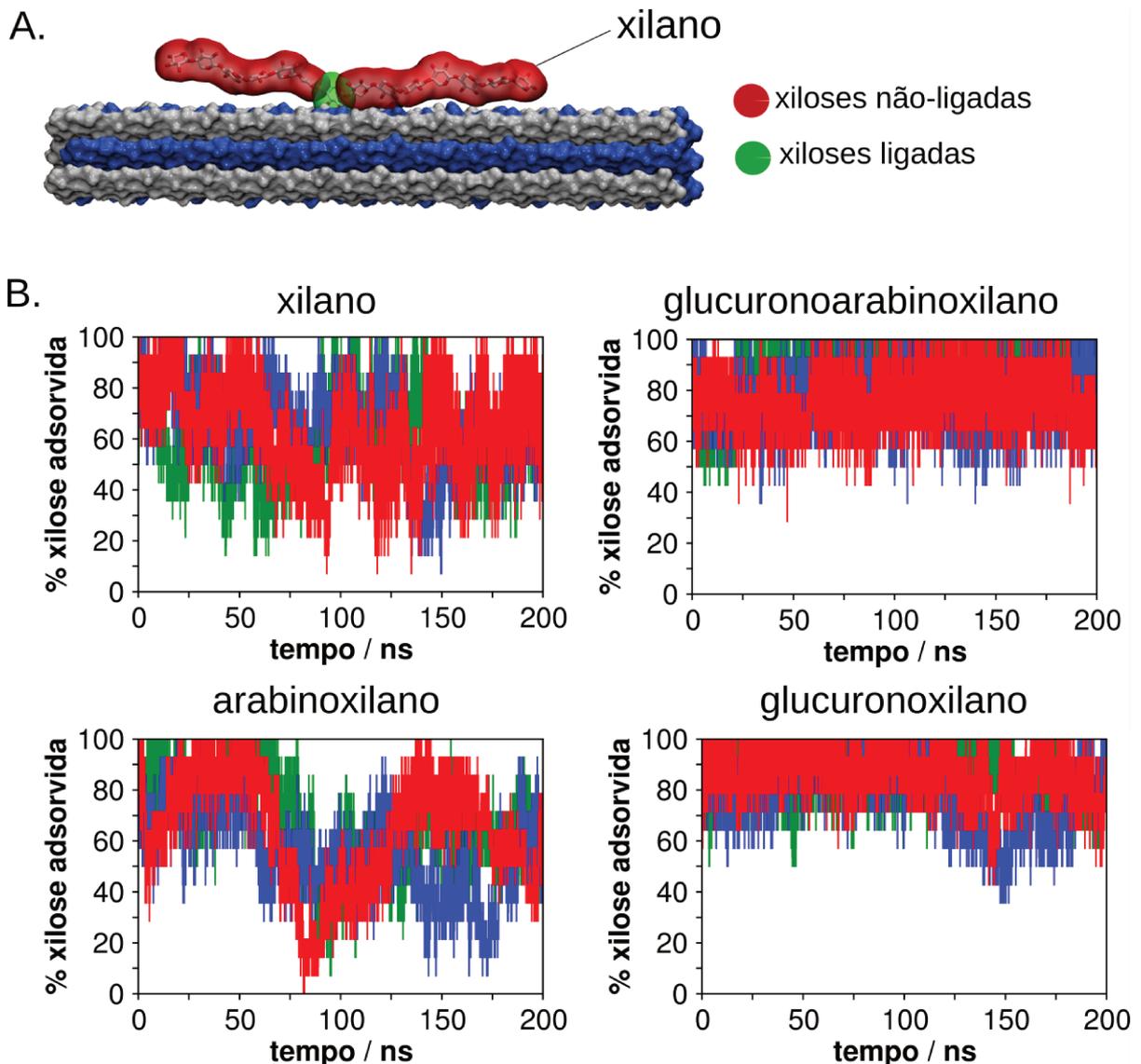


Figura 3.7. Análise da porcentagem de xiloses adsorvidas das cadeias xilana com a celulose. (A) Pose da simulação do complexo xilano-celulose, exemplificando uma situação em que apenas um resíduo de xilose está adsorvido na celulose. (B) Porcentagem de xiloses adsorvidas na celulose em função do tempo. Cores diferentes indicam diferentes simulações.

A porcentagem de xiloses adsorvidas do xilano e do arabinoxilano foi em torno de 65% (figura 3.7B), sendo que esta porcentagem teve alta variação durante toda a simulação, o menor valor observado foi de 7% para o xilano, equivalente a um único resíduo de xilose, e 0% para o arabinoxilano (ou seja, dessorção completa de toda a

cadeia). O glucuronoarabinoxilano e o glucuronoxilano apresentaram porcentagem de xiloses adsorvidas variando em torno de 85% (figura 3.7B), com porcentagem mínima em torno de 30% e 40%, para o glucuronoarabinoxilano e glucuronoxilano, respectivamente, mostrando uma variação menor que os descritos primeiramente.

Assim, apesar das flutuações térmicas, os resíduos das cadeias ramificadas com GlcA possuem maior probabilidade de estarem adsorvidos quando comparado com as cadeias sem este tipo de ramificação. Isso mostra que as cadeias xilana ramificadas com GlcA têm maior força de interação com a celulose.

A figura 3.8A mostra a vista lateral de uma pose da simulação do complexo glucuronoarabinoxilano-celulose, com a cadeia xilana adsorvida sobre uma fissura da celulose e de forma estendida. Os resultados indicam que ela não se adsorve paralelamente às cadeias glucana da celulose cristalina. Isso ocorre porque a cadeia xilana estabelece ligações de hidrogênio com as cadeias glucana localizadas diretamente à direita e abaixo da cadeia xilana (figura 3.8A), de forma que a cadeia xilana consegue se ligar a 2 cadeias glucana da celulose, simultaneamente. Com a cadeia glucana localizada à direita da xilana, as ligações de hidrogênio são formadas entre as hidroxilas ligadas aos carbonos C2 e C3 da cadeia xilana e a hidroxila ligada ao carbono C6 da cadeia glucana. Com a cadeia glucana localizada abaixo da xilana, as ligações de hidrogênio são formadas entre a hidroxila ligada ao carbono C6 da cadeia glucana e o oxigênio ligado ao carbono C5 e o oxigênio glicosídico da cadeia xilana.

Portanto, as simulações mostram que a cadeia xilana completamente estendida e alinhada com uma fissura da superfície da celulose, interage com duas cadeias glucana diferentes, ao mesmo tempo. Para analisar se a cadeia xilana está alinhada com a celulose ou se está deslocando sobre sua superfície de forma não-específica, calculamos o número de cadeias glucana com as quais a cadeia xilana faz contato em cada instante das simulações. Consideramos que uma dada cadeia glucana contacta a cadeia xilana se a distância entre a cadeia glucana e alguma xilose for menor que 2,5 Å. Esta quantidade foi computada, em função do tempo, para os quatro sistemas estudados (figura 3.8B).

O xilano e o arabinoxilano se deslocaram não-especificamente sobre as fissuras da superfície da celulose e interagiram com até 7 cadeias glucana ao mesmo

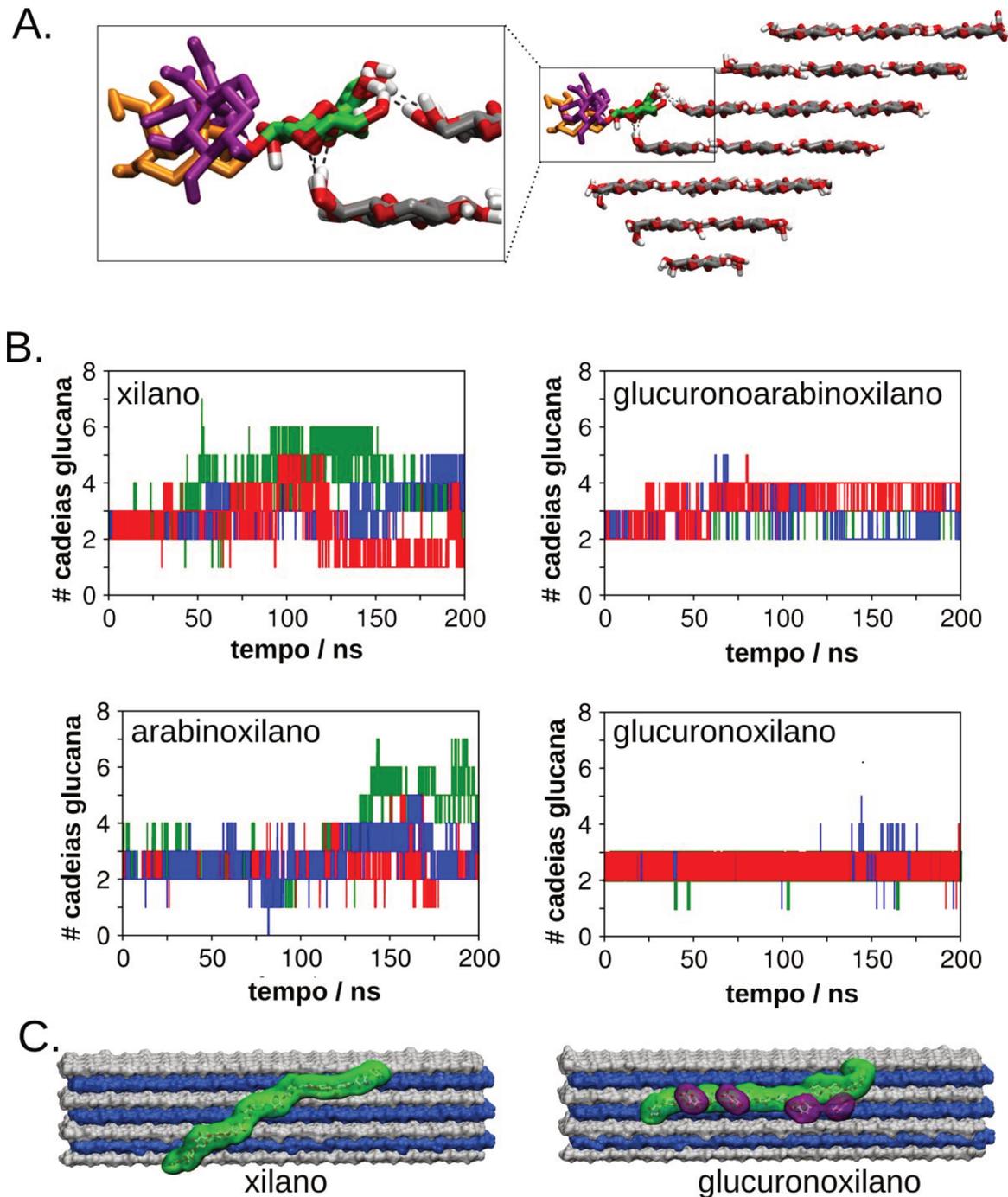


Figura 3.8. Análise da adsorção da cadeia xilana sobre as fissuras da celulose e seu alinhamento com as cadeias glucana da superfície hidrofílica (110) da celulose. (A) Representação da vista lateral do modo de adsorção da cadeia xilana sobre uma fissura da superfície da celulose, mostrando que a cadeia xilana não se alinha a uma única cadeia glucana da celulose, interagindo simultaneamente com duas cadeias glucana. (B) Número de cadeias glucana que contactam as cadeias xilanas adsorvidas em função do tempo. Cores diferentes indicam diferentes simulações. (C) À esquerda, uma pose da simulação do complexo xilano-celulose, exemplificando uma situação em que o xilano está adsorvido sobre 7 cadeias glucana. À direita, pose da simulação do complexo glucuronoxilano-celulose, exemplificando uma situação em que a cadeia xilana está adsorvida sobre 5 cadeias glucana, mas ainda mantém especificidade sobre as fissuras da superfície da celulose, como visto pelo alto grau de alinhamento da cadeia xilana.

tempo (figura 3.8B). A figura 3.8C mostra uma pose da simulação do complexo xilano-celulose, onde o xilano interage com 7 cadeias glucana. O número de cadeias glucana em que o glucuronoarabinoxilano adsorveu variou, na maior parte da simulação, entre 2 e 4 (figura 3.8B). Apesar de não permanecer completamente estendido sobre a celulose, o glucuronoarabinoxilano manteve certa especificidade com relação às fissuras da celulose. O glucuronoxilano permaneceu, na maior parte da simulação, com o centro da cadeia na mesma fissura da superfície da celulose e apenas com as pontas interagindo com até 5 cadeias glucana (figura 3.8B). A figura 3.8C mostra uma pose da simulação do complexo glucuronoxilano-celulose em que a cadeia xilana está interagindo com 5 cadeias glucana, mas a parte central permaneceu na fissura inicial durante toda a simulação. Esses resultados mostram que as ramificações de GlcA contribuíram para que a cadeia xilana não se deslocasse pela superfície de celulose, o que se correlaciona com a maior força de interação com a celulose que as cadeias xilana com estas ramificações apresentaram.

3.3.2 Efeito cooperativo xilano-xilano

Para analisar a compatibilidade das cadeias xilana para adsorção em paralelo na superfície hidrofílica (110) da celulose e também dos possíveis efeitos estéricos entre as ramificações, geramos simulações de dois sistemas diferentes, sendo um deles com dois xilanos (sem ramificações) e o outro com dois glucuronoarabinoxilanos. Fizemos as análises descritas anteriormente para comparar com os sistemas com apenas uma cadeia xilana adsorvida.

Os sistemas com duas cadeias xilana apresentaram menor mobilidade que os modelados com apenas uma cadeia, conforme revelam as análises de RMSD mostradas na figura 3.9. O RMSD máximo foi em torno de 10 Å e 5 Å para os xilanos e glucuronoarabinoxilanos, respectivamente. Conforme descrito anteriormente, o sistema com apenas um xilano adsorvido apresentou RMSD máximo de 25 Å e, com um glucuronoarabinoxilano, 10 Å.

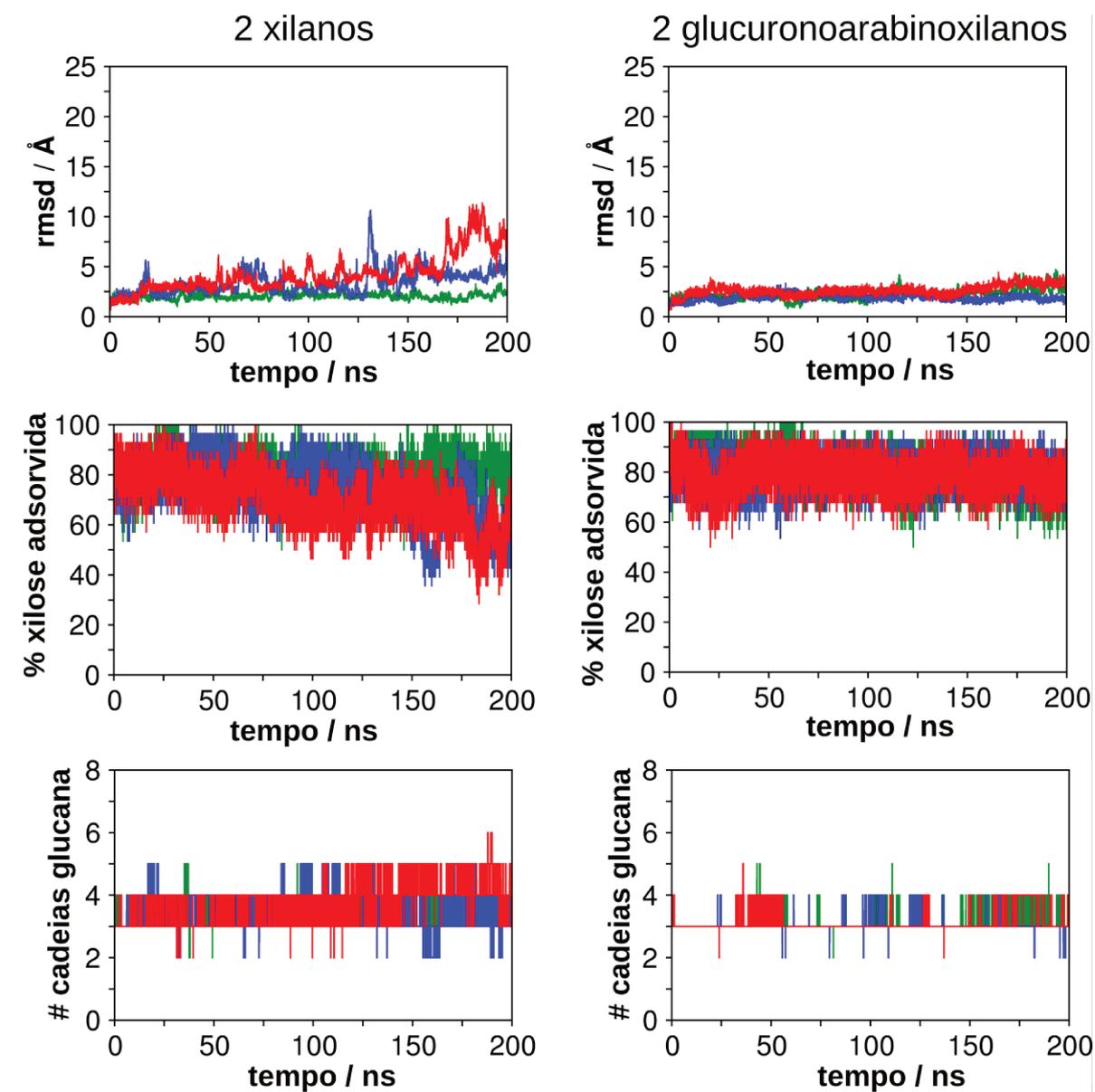


Figura 3.9. Análises da adsorção de duas cadeias xilana em paralelo. De cima para baixo, RMSD, porcentagem de xiloses adsorvidas na celulose, número de cadeias glucana adsorvidas pelas cadeias xilana. Cores diferentes indicam diferentes simulações.

As cadeias modeladas em paralelo apresentaram maior força de ligação com a celulose. As porcentagens mínimas de xiloses adsorvidas foram de 30 e 50%, para os sistemas com dois xilanos e dois glucuronoarabinosilanos, respectivamente (figura 3.9). Para estes mesmos sistemas com apenas uma cadeia xilana adsorvida, as porcentagens mínimas foram de 7% e 30%, respectivamente, mostrando forte estabilização quando duas cadeias estão adsorvidas.

A figura 3.10A mostra a vista lateral de uma pose da simulação do complexo glucuronoarabinosilano-celulose, indicando que duas cadeias xilana estão esticadas

e alinhadas sobre a superfície da celulose quando elas fazem contato com 3 cadeias glucanas. A figura 3.9 mostra que o máximo de cadeias glucana que adsorvem os 2 xilanos e 2 glucuronoarabinoxilanos foi 6 e 5, respectivamente. As poses representando estes máximos são mostradas na figura 3.10B. As cadeias xilana dos sistemas modelados com duas cadeias em paralelo permaneceram mais estendidas que as modeladas com apenas uma cadeia. As xiloses, localizadas no centro das cadeias adsorvidas em paralelo, permaneceram interagindo com as mesmas cadeias glucana da configuração inicial, durante toda a simulação (figura 3.10B).

A adsorção em paralelo de xilanos apresentou efeitos cooperativos para interação com a celulose, que ocorre devido ao empacotamento da cadeia xilana. Estes efeitos foram ainda mais pronunciados para o glucuronoarabinoxilano, indicando que não há efeitos estéricos entre as ramificações na adsorção em paralelo. As interações entre as cargas negativas das ramificações de GlcA serão discutidas na seção 3.3.3.

Analizamos as interações xilano-xilano e xilano-celulose nas simulações de dois glucuronoarabinoxilanos. As interações de van der Waals constituem cerca de 60% das interações xilose-xilose, refletindo a natureza preponderantemente hidrofóbica destas interações. As ligações de hidrogênio entre as cadeias xilanas são bem raras, devido à geometria da adsorção das cadeias. As interações celulose-xilano incluem ligações de hidrogênio de uma mesma cadeia xilana com duas cadeias glucana diferentes da celulose, ao mesmo tempo. Isso resultou em uma média de 9.7 ± 2.6 de ligações de hidrogênio entre as 14 xiloses do glucuronoarabinoxilano e a celulose.

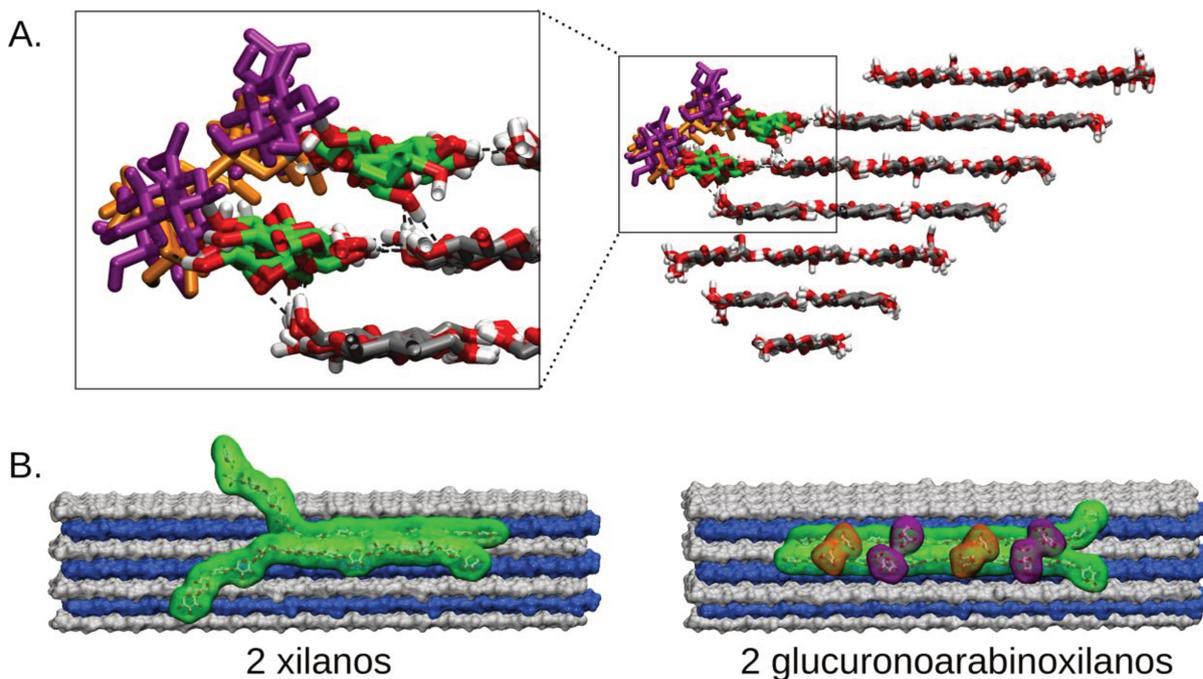


Figura 3.10. Representação do modo de ligação de duas cadeias xilana adsorvidas em paralelo sobre a superfície hidrofílica (110) da celulose. (A) Vista lateral de uma pose da simulação do complexo 2 glucuronoarabinoxilanos-celulose, mostrando a adsorção de duas cadeias xilana sobre duas fissuras da superfície da celulose e que as cadeias xilana interagem com três cadeias glucana ao mesmo tempo. (B) Poses das simulações dos complexos 2 xilanos-celulose e 2 glucuronoarabinoxilanos-celulose, onde observamos maior número de cadeias glucana adsorvidas.

3.3.3 Efeito das ramificações

As análises apresentadas anteriormente mostram que enquanto as ramificações de GlcA cooperaram para adsorção da cadeia xilana na face hidrofílica (110) da celulose, as ramificações de Ara não apresentaram efeitos significativos quando comparadas com o xilano sem ramificações. Com base nisso, fizemos algumas análises para entender os efeitos das ramificações na interação das cadeias xilana com celulose.

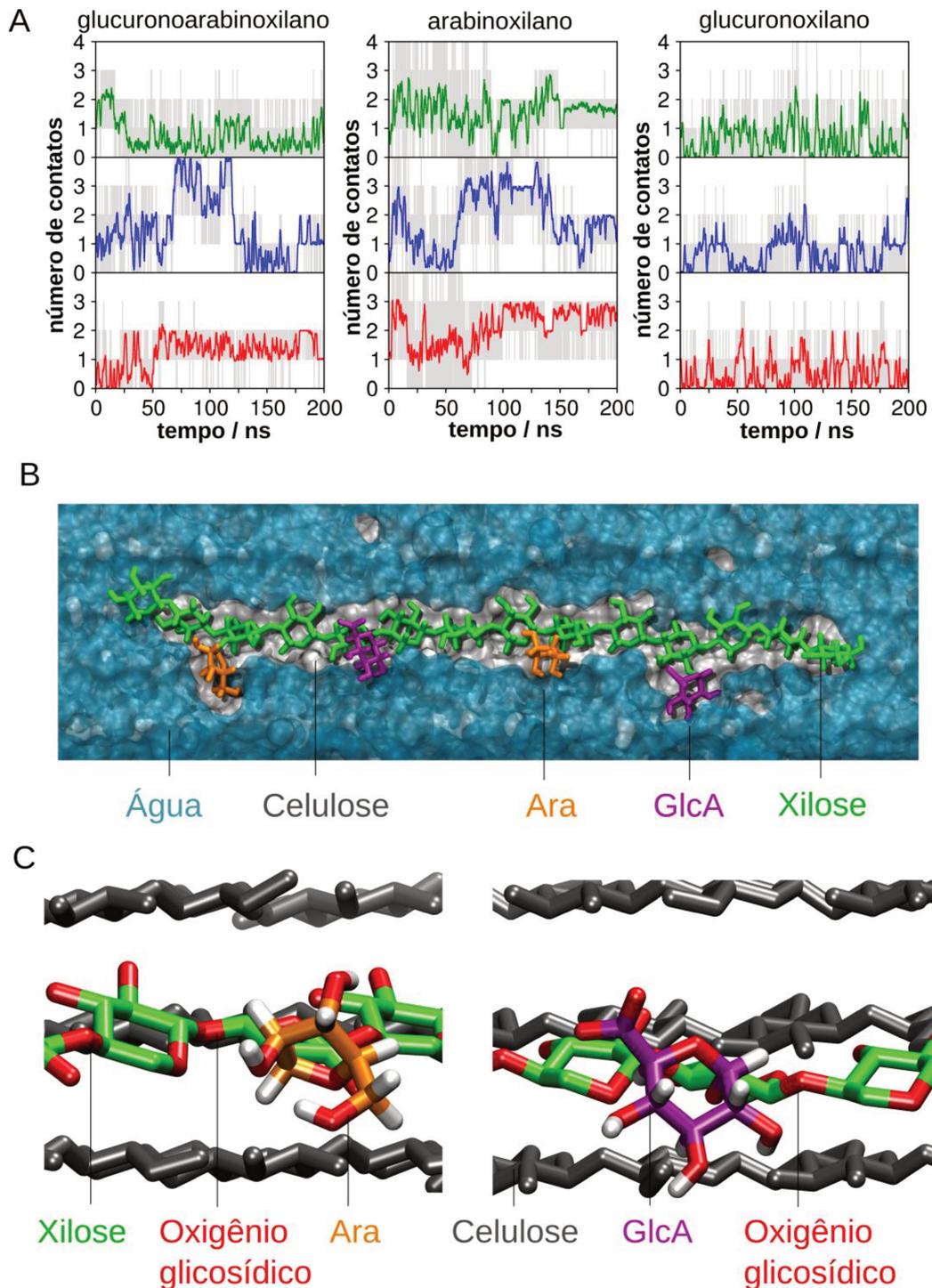


Figura 3.11. Interação das ramificações com a celulose. (A) Número de contatos das ramificações com a celulose, seguindo a ordem da esquerda para a direita: glucuronoarabinoxilano, arabinoxilano e glucuronoxilano. Cores diferentes indicam simulações diferentes. (B) Pose da simulação do complexo glucuronoarabinoxilano-celulose, em que duas das quatro ramificações contatam a superfície de celulose, mostrando a desidratação local da interface xilano-celulose em torno das substituições adsorvidas. (C) Pose da simulação do complexo glucuronoarabinoxilano-celulose mostrando a interação dos oxigênios glicosídicos com as ramificações Ara (à esquerda) e GlcA (à direita).

Primeiramente, investigamos se as ramificações contactam a celulose e se podem estabilizar a adsorção da cadeia xilana. As ramificações contactaram a celulose de forma intermitente (figura 3.11A). A figura 3.11B mostra pose da simulação do complexo glucuronoarabinoxilano-celulose, juntamente com uma camada de 5Å de água, mostrando duas ramificações contatando a celulose e promovendo a desidratação local da interface xilano-celulose em torno das ramificações adsorvidas. O arabinoxilano apresentou maior frequência de contatos das ramificações com a celulose, enquanto que o glucuronoxilano apresentou menor frequência (figura 3.11A). Entretanto, estes contatos aparentemente não constituem a razão da maior estabilização do glucuronoxilano em relação ao xilano. Nas análises de mobilidade e força de interação, o arabinoxilano mostrou comportamento semelhante ao xilano não ramificado, indicando que a interação das ramificações com a superfície da celulose não é o fator que aumenta a estabilização da interação xilano-celulose.

A ramificação Ara está na posição α -1,3 e impede a única ligação de hidrogênio intramolecular da cadeia xilana, pois esta ligação é formada entre a hidroxila ligada ao carbono C3 e o oxigênio O5 da xilose (figura 3.1B). Isso aumenta a mobilidade dos ângulos ϕ e ψ e favorece a conformação de tripla-hélice, cooperando negativamente para adsorção na celulose. O oxigênio glicosídico mais próximo da Ara está voltado para a celulose e tem pouca interação com a ramificação (figura 3.11C).

O GlcA ramificado na posição α -1,2 interage mais com o oxigênio glicosídico (figura 3.11C), que está voltado para a solução e sofre efeitos estéricos de repulsão a ramificação. Isso pode limitar os ângulos permitidos para rotação dos diedros glicosídicos ϕ e ψ e favorecer a adsorção da cadeia xilana na celulose.

Nas simulações de glucuronoarabinoxilanos adsorvidos em paralelo, os GlcA poderiam se repelir devido às cargas negativas. Entretanto, observamos que o Ca^{2+} , cátion mais abundante em paredes celulares e que foi adicionado para neutralizar o sistema, atuou neutralizando a repulsão e também criando reticulação entre os ácidos. A figura 3.12A mostra as distâncias entre os dois GlcA e o Ca^{2+} nas três simulações. A distância flutua em torno de 2.2 Å quando ocorre formação da reticulação GlcA- Ca^{2+} -GlcA, que é representada na figura 3.12B. Este complexo possui carga total zero e permaneceu estável pelo restante da simulação. A figura 3.12C, obtida de uma pose da simulação do complexo 2 glucuronoarabinoxilanos-celulose, mostra formação de

dois complexos GlcA-Ca²⁺-GlcA. Com isso, as xiloses que estão entre os pontos de reticulação permanecem adsorvidas na celulose, enquanto as pontas se dissociam. Portanto, a formação desta reticulação cooperou com a interação da cadeia xilana com a celulose.

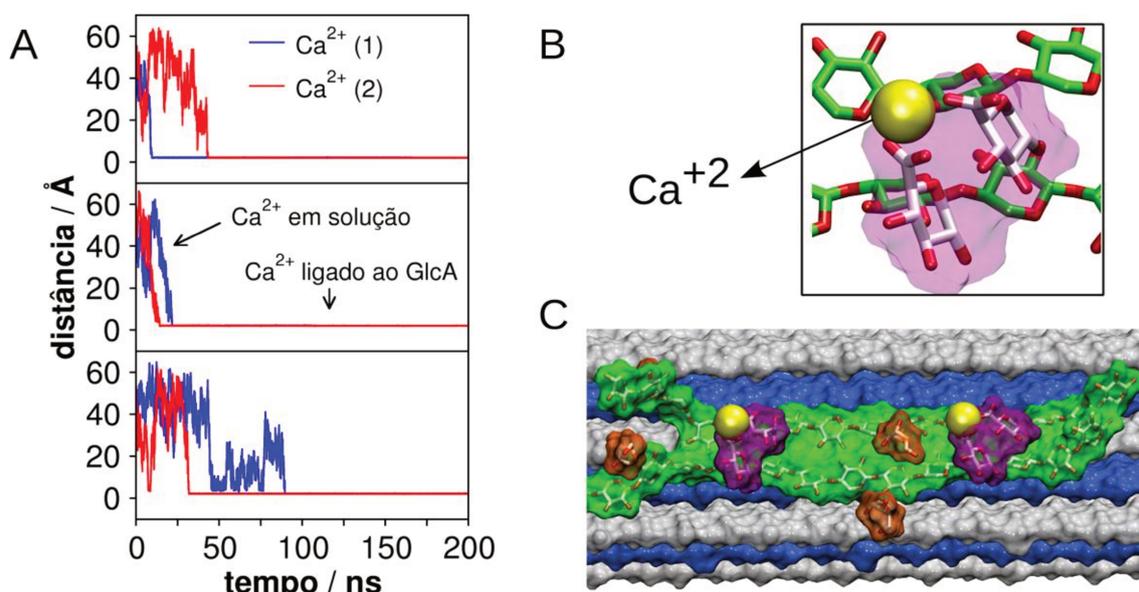


Figura 3.12. Formação do complexo GlcA-Ca²⁺-GlcA. (A) Distância entre o Ca²⁺ e os GlcAs. A distância flutua em torno de 2.2 Å quando a reticulação é formada; distâncias maiores que este valor indica que o Ca²⁺ está em solução. Os dois íons cálcio (1) e (2) de uma mesma simulação estão representados em cores diferentes e o três gráficos representam simulações diferentes. (B) Representação do complexo GlcA-Ca²⁺-GlcA. (C) Pose da simulação do complexo 2 glucuronoarabinoxilanos-celulose, em que há a formação de dois complexos GlcA-Ca²⁺-GlcA, mostrando a estabilidade da adsorção das xiloses que estão entre os complexos.

Como observação final, nos modelos utilizamos ramificações de GlcA ao invés de MeGlcA. Para certificar que nossos resultados são independentes da falta de metilação na hidroxila HO4, analisamos as interações desta hidroxila com o restante do sistema. Observamos que nas simulações com uma cadeia xilana a HO4 é solvatada pela água e não faz interações significativas com nenhum grupo da celulose ou da cadeia xilana. Nas simulações com duas cadeias xilana, HO4 faz ligações de hidrogênio apenas com o grupo carboxilato do GlcA da cadeia xilana adsorvida em paralelo, entretanto essas representam apenas 26% das ligações de hidrogênio entre os GlcA em paralelo (envolvendo grupos diferentes do HO4). Com isso, nossas conclusões são independentes da metilação na hidroxila HO4.

3.3.4 Conformação diedral

Xilanos sem ramificação e em solução permanecem com conformação majoritariamente tripla-hélice, com a soma dos diedros glicosídicos ψ e ϕ variando em torno de 190° (figura 3.13A). Para analisar as possíveis alterações causadas pelas ramificações, calculamos a conformação diedral do xilano e do glucuronoarabinóxilano em solução. As cadeias glucana na celulose cristalina permanecem com conformação dupla-hélice, onde a soma dos diedros glicosídicos ψ e ϕ variam em torno de 120° (figura 3.13B). Para verificar a conformação da cadeia xilana adsorvida na celulose calculamos as somas dos diedros glicosídicos com relação à distância da celulose para os sistemas estudados.

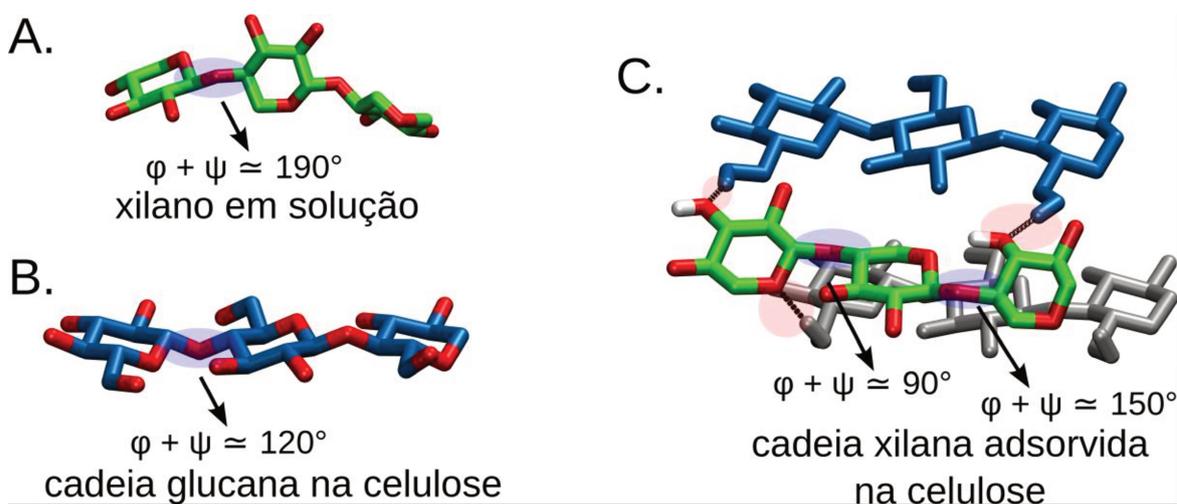


Figura 3.13. Representação de polissacarídeos em diferentes conformações diedrais. (A) Xilano na conformação tripla-hélice, $\phi + \psi = 190^\circ$. (B) Cadeia glucana na conformação dupla-hélice, $\phi + \psi = 120^\circ$. (C) Cadeia xilana adsorvida na superfície hidrofílica da celulose em conformação aproximada de dupla-hélice, onde os diedros adjacentes anulam as variações.

Nos sistemas em solução, a distribuição das somas dos ângulos diedrais ($\phi + \psi$) de todos os diedros das cadeias do xilano e do glucuronoarabinóxilano apresentaram um pico em 190° (figura 3.14). No glucuronoarabinóxilano, a distribuição de $\phi + \psi$ dos diedros diretamente ao lado das ramificações de Ara apresentaram um pico em 190° com maior frequência que dos outros diedros, mostrados em vermelho na figura 3.14, causada pela ramificação α -1,3-Ara, que favorece a conformação tripla-hélice.

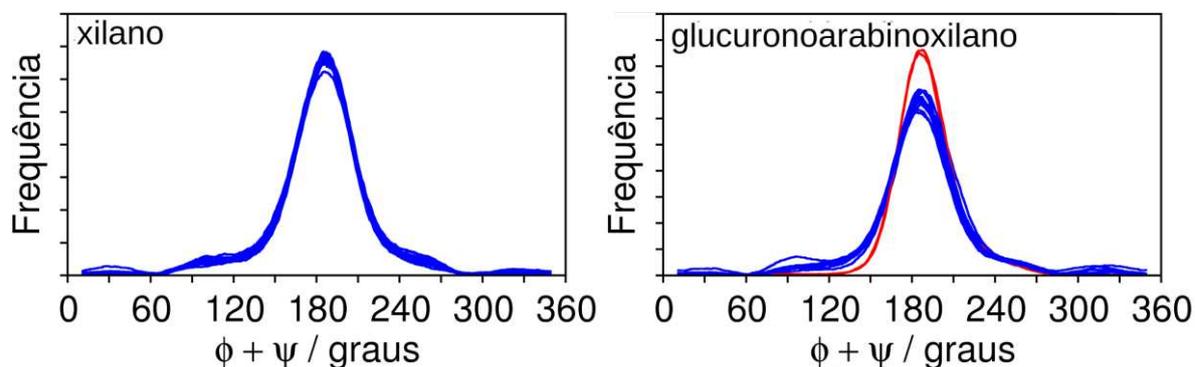


Figura 3.14. Distribuição das somas dos ângulos diedrais ($\phi+\psi$) para todos os diedros das cadeias xilana em solução.

A figura 3.15 mostra distribuições bi-dimensionais de $\phi+\psi$ para todos o diedros das cadeias xilanas e das distâncias entre as xiloses e a superfície da celulose (os primeiros 2 resíduos de xilose de cada uma extremidades da cadeia xilana não foram considerados nesta análise por serem mais móveis). Para distâncias maiores que 2,5 Å, as xiloses referentes ao diedro computado estão dissociadas da celulose, e em distâncias menores que 2,5 Å, as xiloses estão adsorvidas na celulose. Assim, pode-se analisar a conformação diedral apenas das cadeias adsorvidas na celulose.

As distribuições de $\phi+\psi$ para o xilano e arabinoxilano apresentaram um máximo de frequência acentuado em torno de 190° (figura 3.15A), independente da distância da celulose, indicando que estas cadeias estão na conformação tripla-hélice mesmo quando próximas à celulose. Isso mostra ligação não-específica e com a mesma conformação do xilano em solução. Para o glucuronoarabinoxilano e glucuronoxilano, quando as xiloses estão perto da superfície (110) da celulose, a distribuição $\phi+\psi$ oscila em torno de 120° e entre os picos em 90° e 180°, com um pico de maior em 90° (figura 3.15A). Durante os poucos momentos em que algumas xiloses se distanciam da superfície da celulose, a cadeia xilana assume a conformação diedral tripla-hélice, com $\phi+\psi$ variando em torno de 190° (valor observado para a cadeia xilana em solução). Nos sistemas com as cadeias xilana adsorvidas em paralelo, 2 xilanos e 2 glucuronoarabinoxilanos, a distribuição $\phi+\psi$ também oscila em torno de 120°, mas com picos em 90° e 150° (figura 3.15A).

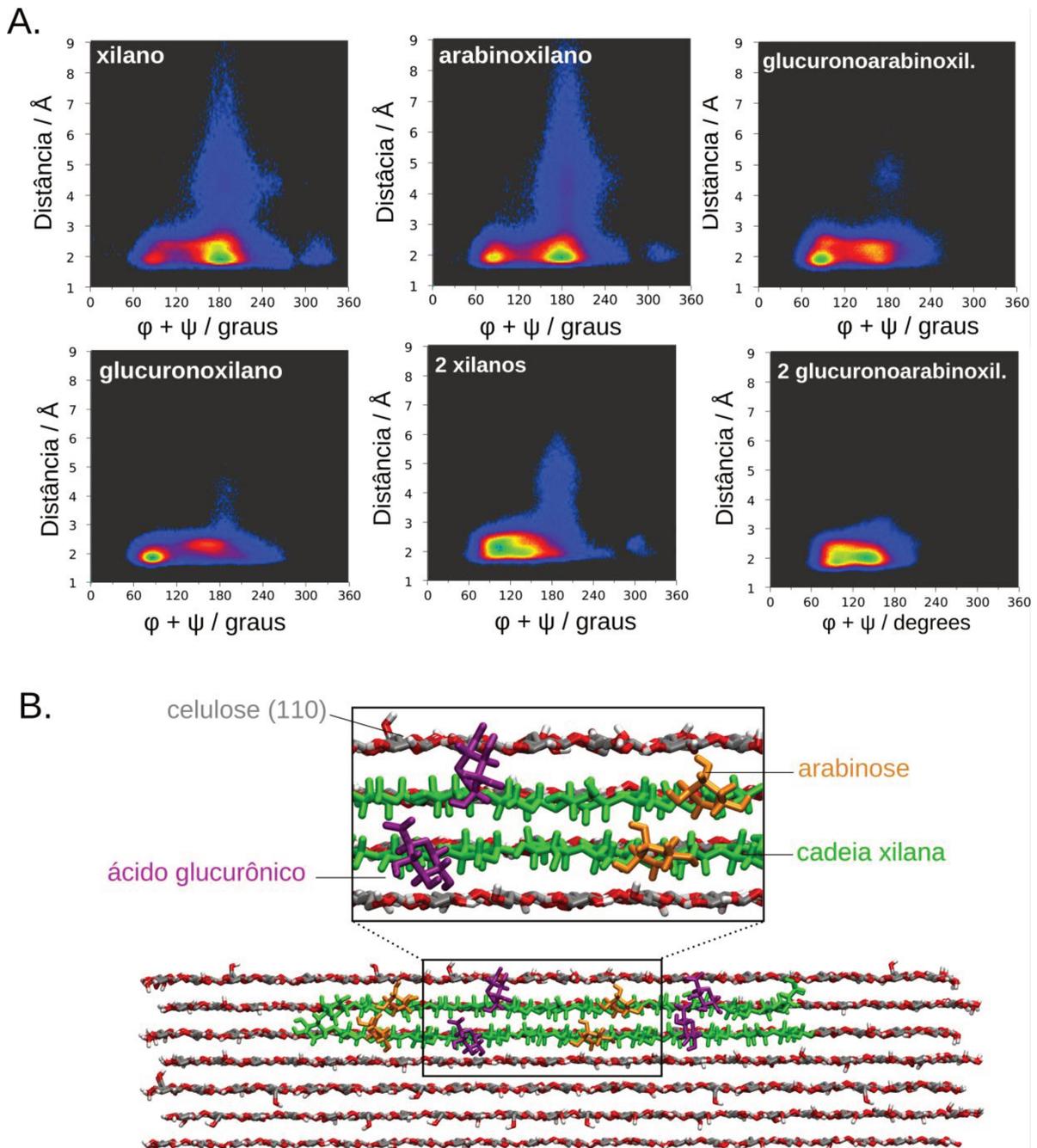


Figura 3.15. Conformação diedral das cadeias xilana adsorvidas na superfície hidrofílica (110) da celulose. (A) Distribuição bi-dimensão de $\phi+\psi$ para todos os diedros das cadeias xilana e das distâncias das xiloses da superfície da celulose. (B) Pose obtida da simulação de 2 glucuronoarabinoxilanos adsorvidos na face hidrofílica (110) da celulose, mostrando a configuração estendida da cadeia.

A figura 3.13C mostra a conformação diedral da cadeia xilana adsorvida na celulose em uma pose da simulação do complexo glucuronoarabinoxilano, mostrando que os valores de $\phi+\psi$ alternam entre maior e menor que 120° para os resíduos

vizinhos. Com isso a cadeia xilana não possui valores semelhantes para todos os diedros glicosídicos. Em vez disso, um diedro glicosídico com $\phi+\psi$ em torno de 90° é seguido por outro com $\phi+\psi$ em torno de 150° , de tal forma que a diferença de um anula a diferença do próximo com relação a 120° . Como resultado, observa-se então uma conformação dupla-hélice aproximada, com torções alternadas. Isso permite que a cadeia xilana permaneça estendida sobre a celulose, como está ilustrada na figura 3.15B, para 2 glucuronarabinoxilanos adsorvidos na celulose. Estas distorções diedrais ocorrem devido às ligações de hidrogênio com as cadeias glucana que estão abaixo da cadeia xilana (3.13C).

3.4 Discussão

As simulações mostraram que xilanos ramificados com o padrão encontrado em coníferas possuem estabilidade na adsorção da superfície hidrofílica (110) da celulose e maior força de interação com a celulose que cadeias xilanas sem ramificações, mostrando que podem adsorver na face hidrofílica da celulose presente na parede celular.

No trabalho de DM realizado anteriormente da adsorção de xilanos na superfície hidrofílica da celulose, o xilano foi modelado nas superfícies (010) e (020) da celulose, que possuem fendas para a ligação do substrato com alta interação hidrofóbica [16]. No trabalho aqui apresentado, modelamos o xilano na superfície hidrofílica (110) da microfibrila hexagonal, que tem geometria diagonal e não possui fendas para ligação. Com isso, mostramos que a cadeia xilana não necessita de fendas com efeitos de empacotamento para se ligar à celulose.

A cadeia xilana possui menos ligações de hidrogênio intramoleculares que a cadeia glucana e isso aumenta a mobilidade dos ângulos diedrais glicosídicos do xilano. Essa mobilidade possibilita a interação da cadeia xilana com duas cadeias glucana ao mesmo tempo e faz com que a cadeia xilana não adsorva na conformação de dupla-hélice perfeita, como se pensava anteriormente. Isso coloca em questão a necessidade da ramificação em apenas um lado da cadeia xilana, para que ocorra adsorção na superfície hidrofílica. Na conformação adotada pelas cadeias xilanas nas simulações deste trabalho, seria possível a ramificação nos dois lados da cadeia xilana, sem que houvesse repulsão com a celulose. Entretanto, as ramificações nas xiloses ímpares iriam impossibilitar a ligação de hidrogênio entre a hidroxila ramificada

e a celulose e estas ligações de hidrogênio mostraram-se importantes para adsorção das cadeias xilanas na celulose. Ainda não foi estudado se as ramificações poderiam criar novas interações que estabilizariam o complexo ou se a falta das ligações de hidrogênio destas xiloses com a celulose fariam a cadeia xilana desmoronar.

Para adsorção de cadeias xilanas em paralelo é evidente a necessidade de ramificações em apenas um lado da cadeia xilana, pois haveria repulsões entre as ramificações e as xiloses da cadeia em paralelo. Com isso, pode-se sugerir que nos domínios onde as cadeias xilanas possuem ramificações em apenas de um lado, elas estão aptas a adsorver paralelamente às outras cadeias xilanas da parede celular.

Corroborando com a hipótese de que as cadeias xilanas recobrem a microfibrila de celulose, os glucuronoarabinoxilanos adsorvidos em paralelo apresentaram alta estabilidade e força de interação com a celulose, devido aos efeitos de empacotamento e à energia de interação hidrofóbica entre as cadeias xilanas. Não houve efeitos de repulsão entre as ramificações de GlcA, devido à reticulação com o íon cálcio, indicando que a adsorção em paralelo é preferencial para os GlcA pareados, de tal forma que ocorra reticulações entre todas as cadeias xilanas adsorvidas. Mostramos então que é possível que as cadeias xilanas recubram toda a superfície das microfibrilas de celulose, desde que as ramificações estejam em apenas um lado da cadeia xilana.

Já foi demonstrado que as interações estéricas entre a cadeia principal de polissacarídeos e suas ramificações podem restringir a faixa de ângulos permitidos para rotação dos diedros glicosídicos ϕ e ψ [57]. Os efeitos estéricos dos GlcA ramificados na posição α -1,2, juntamente com a própria superfície da celulose, provavelmente limitam a faixa permitida para rotação dos ângulos diedrais glicosídicos, de tal forma que a cadeia xilana fica restrita às conformações adsorvidas na celulose. Para cadeias xilanas adsorvidas em paralelo, a importância do GlcA é clara devido à reticulação com o íon cálcio. Isso corrobora com dados experimentais, que mostram que xilanos sem ramificação de GlcA são facilmente extraídos [45]. As ramificações de Ara nas posições α -1,3 diminuem as interações de hidrogênio intramoleculares, favorecendo a conformação tripla-hélice e não contribuindo com a adsorção da cadeia xilana na celulose. Isso sugere que as ramificações nestas posições estão relacionadas a outras funções da parede celular, como interação com outros polissacarídeos, água e lignina e ferulização, por exemplo.

3.5 Conclusão

Xilano é um dos polissacarídeos mais comuns em paredes celulares secundárias. Entretanto, sua interação com a celulose e com os outros polímeros presentes na parede celular ainda não é bem compreendida. Neste trabalho, mostramos que uma cadeia xilana isolada com o padrão de ramificação encontrado em coníferas se liga estavelmente na superfície hidrofílica (110) da celulose, mas não se adsorve de forma perfeitamente alinhada às cadeias glucana e nem com conformação diedral de dupla-hélice perfeita. As ramificações de GlcA cooperam para a adsorção na superfície da celulose e que as ramificações de Ara aumentam a mobilidade dos diedros glicosídicos devido à diminuição das ligações de hidrogênio intramoleculares e não apresentam contribuições para adsorção na superfície hidrofílica da celulose. Na adsorção de duas cadeias xilanas em paralelo há efeito cooperativo xilose-xilose, que ocorre devido às interações de van der Waals entre os resíduos de xilose, aumentando assim a estabilidade das cadeias. Os glucuronoarabinoxilanos adsorvidos em paralelo apresentaram ainda mais estabilidade de adsorção sobre a celulose, que os xilanos sem ramificação, devido às reticulações das ramificações GlcA com os íons cálcio.

Capítulo 4

Endoglucanase Cel45A de *Phanerochaete chrysosporium*

4.1 Introdução

A endoglucanase Cel45A do fungo *Phanerochaete chrysosporium* (PcCel45A) é uma enzima monomodular pertencente à família GH45 e que hidrolisa os substratos pelo mecanismo de inversão, como outras enzimas desta família. Entretanto, ela é mais similar às β -expansinas, como a ExpB1, que às outras enzimas da família GH45. Isto faz da PcCel45A uma ferramenta para o entendimento do mecanismo de ação pouco compreendido das expansinas, cujo estudo é limitado por conta de dificuldades de expressar expansinas de plantas em bactérias [58].

A estrutura tridimensional da PcCel45A, utilizada neste trabalho, foi cristalizada sem substrato e resolvida por cristalografia de raio-X, com resolução de 1,7 Å, pelo grupo do Prof. Igor Polikarpov (IFSC/USP). A figura 4.1 mostra a estrutura da enzima complexada a uma cadeia glucana (com procedimentos computacionais). O Asp121 é o resíduo que atua como ácido geral na reação catalítica. Os subsítios de ligação ao carboidrato foram numerados a partir do sítio catalítico, sendo quatro subsítios negativos na direção da extremidade não-redutora do substrato e três subsítios positivos na direção da extremidade redutora.

Segundo dados experimentais, a PcCel45A possui atividade de clivagem sobre PASC, CMC e β -glucanos, e não possui atividade sobre xilanos. Na determinação experimental dos oligômeros produzidos pela clivagem de polissacarídeos, tetraose e pentose são os produtos mais abundantes, mas celohexaose também é um produto de clivagem [58]. Portanto, a enzima tem limitações para clivar cadeias com 6 ou menos monômeros.

Dados experimentais mostraram que os mutantes Tyr25Ala, Asp92Ala e Trp161Ala apresentam menor atividade catalítica que a enzima nativa (dados ainda

não publicados). Os resíduos Tyr25 e Trp161, que estão nos subsítios -1 e -4, respectivamente, são conservados nas outras enzimas da família GH45 e na ExpB1 e não participam diretamente da reação catalítica. Devido ao caráter hidrofóbico e planar destes resíduos, eles podem estar associados ao processo de ligação ao substrato através de interações do tipo *stacking*. O resíduo Asp92 é conservado apenas na ExpB1 e está no subsítio -2, mas não está numa posição que permitiria sua atuação como base na reação catalítica. Isto sugere que o papel desse resíduo pode estar mais relacionado ao processo de ligação do substrato.

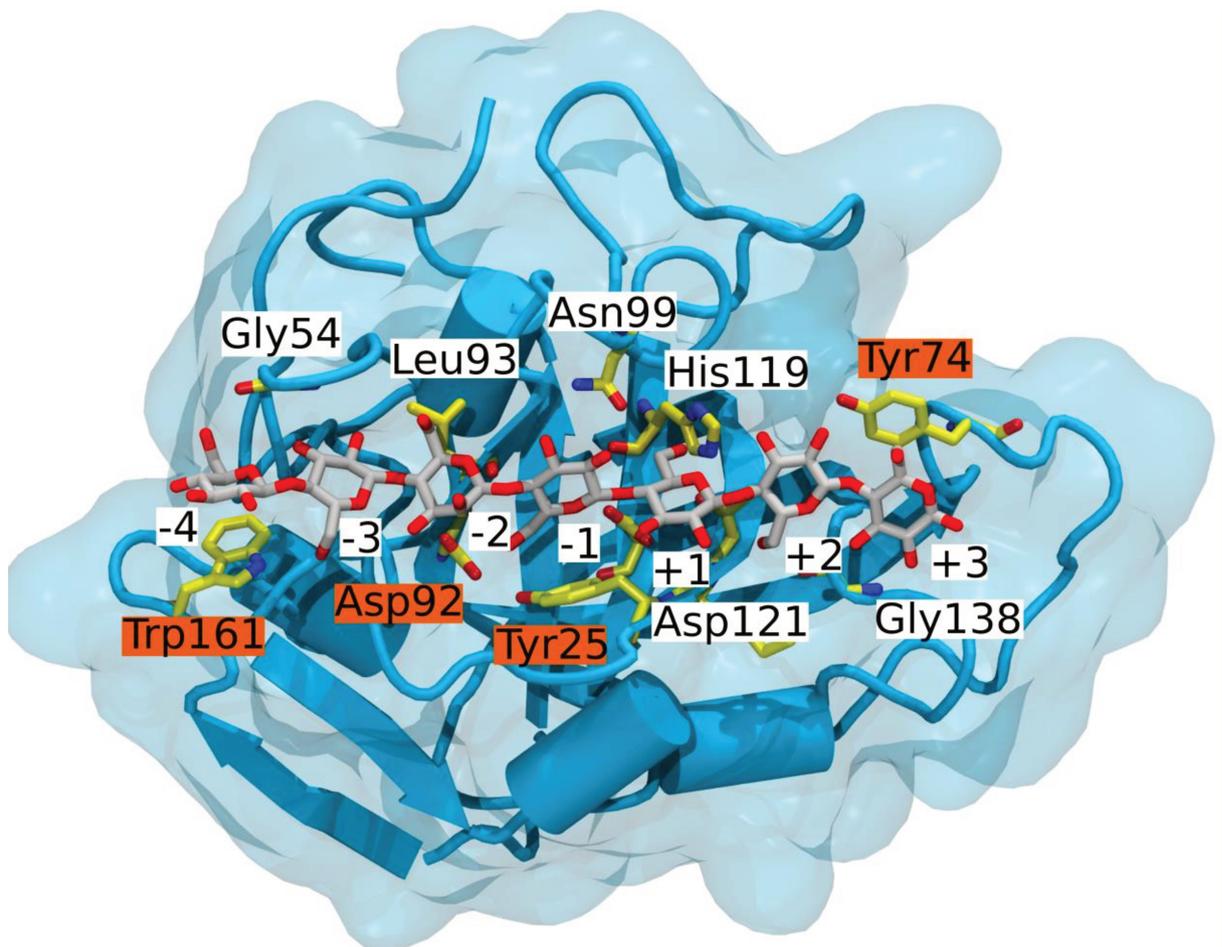


Figura 4.1. PcCel45A complexada com celotriose. Os resíduos responsáveis pela ligação do substrato estão mostrados em amarelo e os subsítios de ligação ao substrato estão numerados de -4 a +3.

A comparação das estruturas da PcCel45A (Fig 4.2A) e da HiCel45 (figura 4.2B) (código PDB:4ENG) mostra que a primeira possui *loops* menores em torno do sítio catalítico, enquanto que os *loops* da segunda formam um túnel para ligar e

prender o substrato para a hidrólise. Com isso, a PcCel45A possui uma fenda catalítica rasa e muito semelhante à ExpB1 (Fig 4.2C) (código PDB:22HCZ) [28,32]. A falta dos *loops* sugere que a PcCel45A possui resíduos com função de ligar ao substrato (para compensar a ausência de *loops*), para que este permaneça ligado à enzima por tempo suficiente para que ocorra a catálise. Os resíduos aromáticos Trp161 e Tyr74 são candidatos potenciais para essa função, devido sua interação hidrofóbica com o substrato.

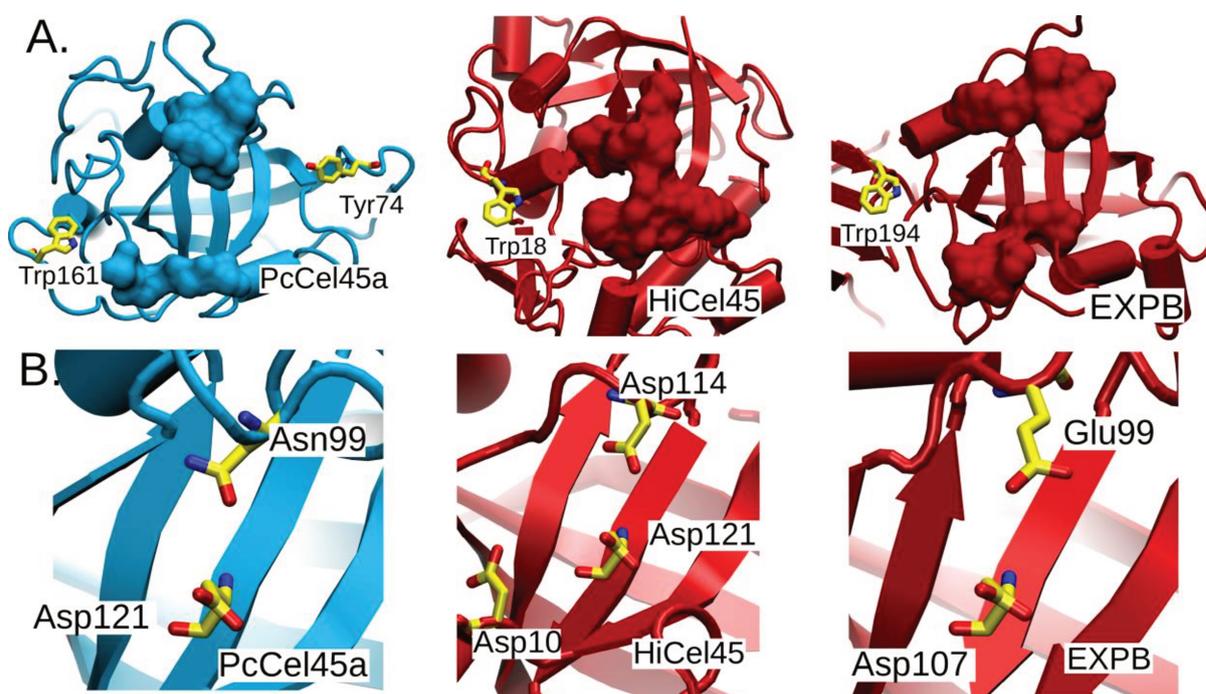


Figura 4.2. (A) Representação dos *loops* em torno do sítio catalítico e dos resíduos hidrofóbicos, que podem ter papel na ligação ao substrato. (B) Comparação dos resíduos alinhados aos aspartatos responsáveis pela catálise na HiCel45. A PcCel45A e a EXPB não possuem resíduo básico referente ao Asp10 (base) da HiCel45, e a PcCel45A possui uma asparagina (Asn99) que se alinha estruturalmente ao resíduo Asp114 da HiCel45.

A HiCel45, que é a enzima melhor caracterizada da família GH45, possui dois aspartatos atuando como resíduos catalíticos, o Asp121 como ácido e o Asp10 como nucleófilo [32]. A PcCel45A não possui o resíduo base conservado, assim como a ExpB1 (figura 4.2B). Recentemente, Nakamura et al. [59] mostraram que a Asn99 atua como nucleófilo durante a reação catalítica da PcCel45A (figura 4.1). Isto justifica o pH ótimo da PcCel45A ser mais baixo que das outras enzimas da família GH45, visto que a asparagina tem pKa menor que o aspartato.

A Asn99 corresponde, na HiCel45 e na ExpB1, aos resíduos Asp114 e Glu99, respectivamente (figura 4.1B). O resíduo Asp114 foi caracterizado apenas como um auxiliador na reação de catálise da HiCel45, e não como o resíduo base. Já a ExpB1, até o momento, não foi comprovado que ela atue utilizando um mecanismo de clivagem. Portanto, não há evidências que os resíduos Asp114 e Glu99 atuem como base.

O mutante Asn99Asp da PcCel45A apresentou atividade muito menor que a enzima nativa [59]. Isso não era esperado, pois o aspartato é um resíduo característico de atuar como base e possui caráter mais básico que a asparagina. Uma possibilidade é que alterações estruturais no mutante possam justificar sua baixa atividade.

Para entender esses dados experimentais apresentados para a PcCel45A, executamos e analisamos simulações de DM da enzima na forma nativa, sem substrato, e na forma complexada com celoheptaose (C7), xiloheptaose (X7), celohexaose (C6) e celoheptaose após clivagem (C4+C3). Simulamos também a enzima ligada à celoheptaose com os resíduos Tyr74, Trp161, Tyr25 e Asp92 mutados para alanina, para analisar a importância destes resíduos na ligação do substrato. Além disso, simulamos o mutante Asn99Asp ligado também à celoheptaose, para verificar possíveis alterações na dinâmica da proteína que expliquem a baixa atividade deste mutante.

4.2 Métodos

As simulações de DM foram realizadas para PcCel45A (nativa e mutantes) na ausência e na presença de vários substratos. Coordenadas inicial do complexo PcCel45A livre de substrato foram tomadas a partir da estrutura cristalográfica (resolução de 1,7 Å) resolvida por raio-X pelo grupo do Prof. Igor Polikarpov (IFSC-USP) (dados ainda não publicados), e os outros complexos PcCel45A-substrato foram obtidos de acordo com o procedimento descrito abaixo.

Após a remoção de heteroátomos presentes na estrutura cristalina, átomos de hidrogênio foram adicionados de acordo com as previsões de estados de protonação de pH = 4,0 obtidas com o servidor H++ [60,61]. Todas as moléculas de água presentes na estrutura de cristal foram mantidas e todo o sistema foi inserido numa caixa de pelo menos 16 Å de espessura. Íons sódio e cloreto foram adicionados na concentração de 0,1 M, com contra-íons em excesso para tornar o sistema

eletricamente neutro.

O complexo PcCel45A-celoheptaose (PcCel45A-C7) foi construído a partir das duas cellotriosos (C3) presentes na estrutura cristalográfica da HiCel45 [32] (código PDB: 4ENG) após o alinhamento estrutural com PcCel45A, e completando a unidade de glicose central com vários ciclos de minimização de energia (descritos abaixo). Os mutantes e complexos com os diferentes substratos foram obtidos a partir da estrutura do complexo PcCel45A-C7 equilibrado em solução. O complexo de PcCel45A com C4 + C3 foi obtido através da remoção da ligação glicosídica entre os subsítios -1 e +1 da cadeia C7. Dois complexos PcCel45A-C6 foram obtidos por remoção de um resíduo da cadeia C7, sendo eles o resíduo glicosídico ligado ao subsítio -4 (C6-Tyr74) e o ligado ao subsítio +3 (C6-Trp161). O complexo PcCel45A-X7 foi obtido pela remoção dos grupos hidroximetílicos da cadeia C7. As mutações Tyr25Ala, Tyr74Ala, Asp92Ala, Asn99Asp e Trp161Ala foram realizadas com o módulo *psfgen* do programa de simulação NAMD [40], utilizando as coordenadas do complexo PcCel45A-C7 equilibrado.

Depois de adicionar as duas moléculas de C3, fizemos minimização de energia de 2000 passos utilizando o método de gradientes conjugados (GC), seguido por 50 ps de DM, mantendo todos os átomos pesados da proteína fixos. Depois disso, uma glícose foi inserida no espaço entre as duas moléculas de C3 e os passos a seguir foram realizados com um potencial harmônico entre o ácido catalítico Asp121 e o átomo de oxigênio glicosídico que seria protonado durante a catálise: (1) 1000 passos de CG + 25 ps de MD com a proteína e as unidades glicosídicas ligadas nos subsítios -4 a -2 e +1 a +3 fixas; (2) passo (1), mas apenas com as unidades glicosídicas ligadas nos subsítios -4, -3 e +3, +2 fixas; (3) passo (1), mas com as unidades glicosídicas ligadas aos subsítios -4 e +3 fixas; (4) passo (1), mas com toda a cadeia de C7 livre; (5) 1000 passos de CG + 50 ps de MD apenas com os carbonos alfa fixos; (6) 50 ns de MD, mantendo apenas a restrição entre Asp121 e do oxigênio glicosídica. Após estas etapas, a cadeia C7 estava modelada no sítio catalítico da PcCel45A, o complexo PcCel45A-C7 foi simulado, sem nenhuma restrição entre a cadeia C7 e a enzima, até que a cadeia C7 assumisse um modo de ligação não produtivo. A partir da estrutura equilibrada do complexo PcCel45A-C7, foram obtidos os demais sistemas estudados neste trabalho. Os mutantes foram minimizados por mais alguns passos, pelo procedimento a seguir: (1) 1000 passos de CG + 50 ps de MD com a proteína fixa; (2) passo (1), mas apenas com os carbonos alfa fixos; (3) 10 ns de MD sem

restrições. Os passos (1) a (3) foram realizadas com potencial harmônico da distância Asp121-oxigênio glicosídico. Três simulações independentes foram realizadas para cada sistema, sem restrição entre a cadeia C7 e a enzima. As trajetórias duraram de dezenas a centenas de nanossegundos, de acordo com o tempo em que o substrato se manteve no modo de ligação produtiva.

As simulações foram realizadas com o programa NAMD [40] utilizando o campo de força CHARMM [52,62,63] e modelo para água TIP3P [55]. A temperatura e a pressão foram mantidas constantes a 300 K e 1 atm utilizando termostato e pistão de Langevin [40], respectivamente. Interações de longo alcance foram calculadas com particle mesh Ewald (PME) [54] e as interações de curto alcance foram truncadas com um raio de corte de 12 Å. Ligações químicas envolvendo átomos de hidrogênio foram restringidas aos seus comprimentos de equilíbrio e um passo de integração de 2 fs foi usado para integrar as equações de movimento. As análises foram realizadas com VMD [56] e programas desenvolvidos pelo grupo.

Em todos estes sistemas, foi avaliado o tempo de ligação na configuração produtiva (configuração suscetível de ocorrer a catálise) através do monitoramento do RMSD do substrato, com relação à sua posição no início da trajetória. Também foram calculadas as energias de interação entre o substrato e a enzima nativa.

4.3 Resultados

4.3.1 Ligação ao substrato

Para analisar o modo de ligação e os resíduos que mais interagem com o substrato, fizemos simulações de DM da enzima nativa complexada à celoheptaose (C7), cadeia glucana com 7 monômeros. Nestas simulações foram observadas duas configurações distintas do substrato, mostradas na figura 4.3A. Em uma delas o substrato estaria suscetível de ser clivado pela enzima e é mostrada com mais detalhes na figura 4.3B, com a sobreposição das configurações da cadeia glucana em diferentes instantes da simulação. Denominamos esta configuração de “configuração produtiva” e o RMSD da cadeia glucana C7 permanece em torno de 2 Å nesta configuração (figura 4.3C). A figura 4.3A mostra o tempo em que os substratos permaneceram ligados na enzima nativa e mutantes na configuração produtiva (com

RMSD em torno de 2 Å). As cores diferentes indicam as três simulações independentes de cada sistema.

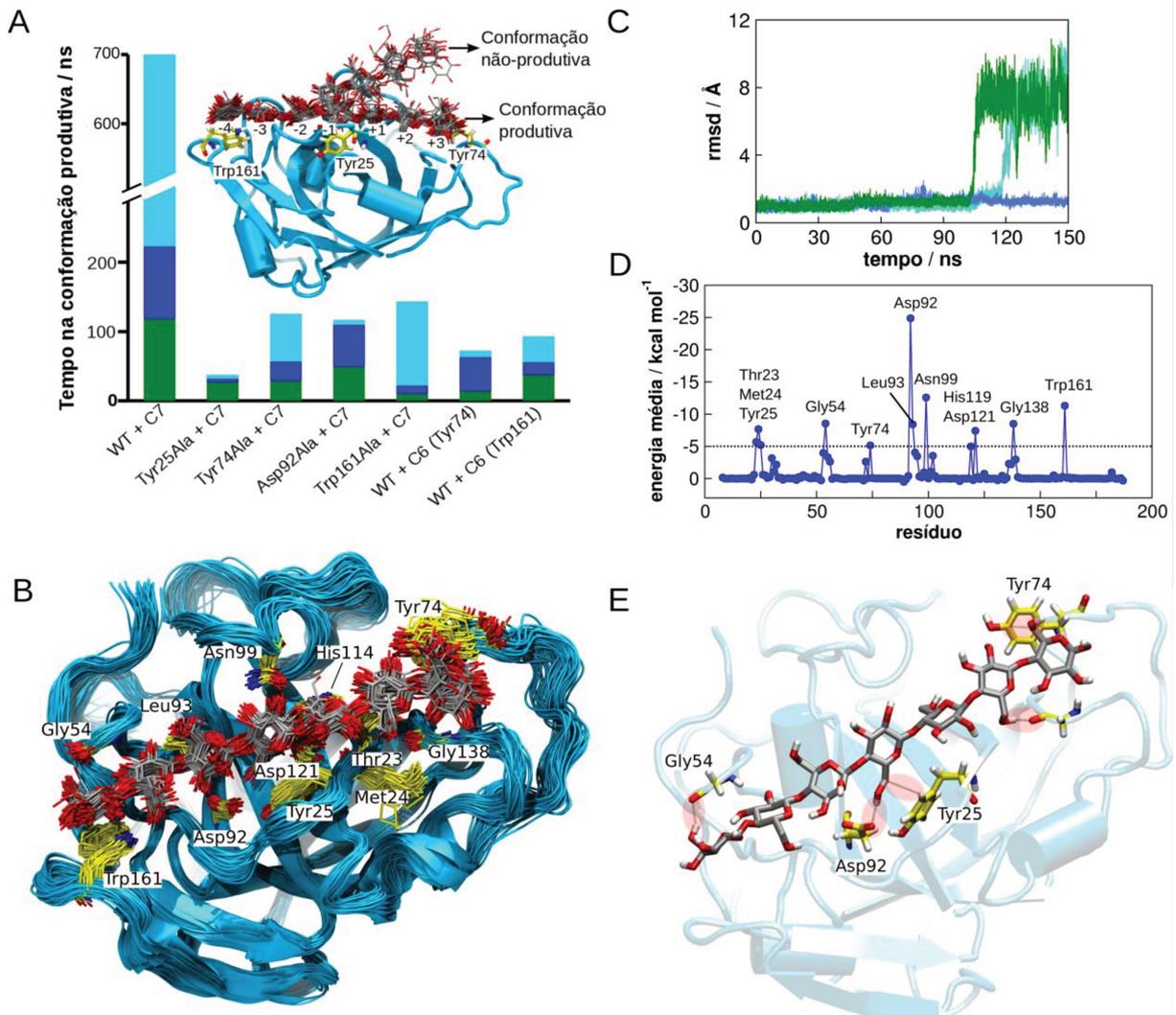


Figura 4.3. (A) Tempo em que os substratos permaneceram ligados na enzima em configuração produtiva, para diferentes mutantes e substratos (C7 e C6) e em diferentes simulações (indicadas por cores diferentes). Sobreposição das configurações do complexo PcCel45A-C7 representando as configurações produtivas e não produtivas do substrato ligado na enzima. (B) Sobreposição das configurações do complexo PcCel45A-C7, mostrando a mobilidade do substrato e os resíduos que mais interagem com o substrato. (C) RMSD da cadeia C7 em relação à configuração inicial, quando o substrato está ligado com uma configuração produtiva o RMSD varia em torno de 2 Å. Cores diferentes representam diferentes simulações do mesmo sistema. (D) Energia de interação média de cada resíduo da PcCel45A com a cadeia C7, sendo indicados os resíduos com energia de interação menor que -5 kcal/mol. (E) Interações entre a PcCel45A e os grupos exocíclicos da cadeia C7.

Na outra configuração observada, o substrato faz interações apenas com os subsítios negativos e o RMSD aumenta para 8 Å quando a cadeia glucana está nesta configuração (figura 4.3C). Apesar do substrato não dissociar completamente para a solução, nesta configuração ele não estaria suscetível de ser clivado pela enzima, pois não está interagindo de forma efetiva com os resíduos catalíticos. Esta configuração foi denominada de “configuração não-produtiva” (figura 4.3A).

Nas simulações com a enzima nativa do complexo PcCel45A-C7, o substrato se manteve ligado na configuração produtiva durante centenas de nanossegundos (figura 4.3A). Uma das simulações foi interrompida em 500 ns sem que o substrato se desviasse da configuração produtiva. Em todos os subsítios, o substrato permaneceu estruturalmente estável, com maior mobilidade apenas no subsítio +3 (figura 4.3A,B).

Os resíduos com maior energia (mais negativa) de interação com o substrato foram Asp92 (subsítio -2), Asn99 (subsítio -1) e Trp161 (subsítio -4) (figura 4.3A,B,D), que estão localizados nos subsítios negativos da fenda de ligação ao substrato (figura 4.1). Os resíduos nos subsítios positivos apresentaram menor energia de ligação com substrato e isto explica a alta mobilidade do subsítio +3 e o modo de ligação adotado na configuração não-produtiva (figura 4.3A).

As simulações dos substrato após a clivagem (C4+C3) mostraram que o RMSD do C3, com relação à posição inicial, variou para mais de 12 Å nos primeiros 30 ns (figura 4.4A), enquanto que o RMSD do C4 permaneceu em torno de 2 Å durante toda simulação (figura 4.4B). O C3 estava ligado inicialmente aos subsítios positivos da enzima e se dissociou rapidamente para a solução. Já o C4 estava ligado nos subsítios negativos e permaneceu ligado à enzima durante todo o tempo de simulação. Isto mostra a maior energia de interação com os subsítios negativos da enzima com o substrato.

As simulações com a enzima ligada à cadeia C7 mostraram que cerca $\frac{1}{3}$ da energia total de interação (-157 kcal/mol) do complexo enzima-substrato está relacionada com as ligações de hidrogênio dos grupos hidroximetílicos exocíclicos da cadeia glucana com a enzima, através dos resíduos Tyr25, Gly54, Tyr74 e Asp92 (figura 4.3E). As simulações com o substrato xiloheptaose (X7) (a xilose difere da glicose pela ausência dos grupos hidroximetílicos) confirmaram a importância destas ligações de hidrogênio para ligação do substrato, pois na falta dos grupos hidroximetílicos, o substrato migrou para solução nos primeiros 5 ns, revelando pouca ou nenhuma afinidade enzima-substrato (figura 4.4C).

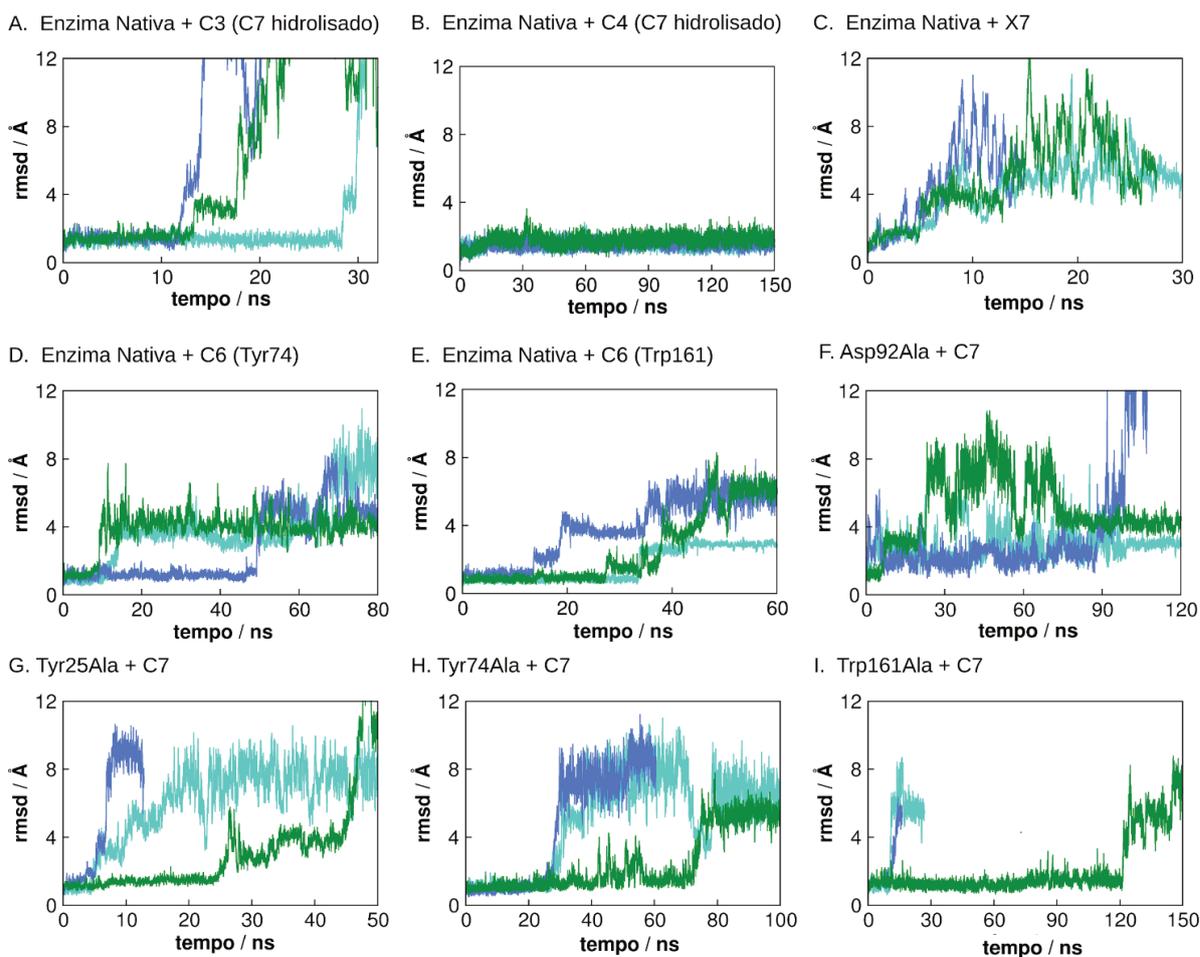


Figura 4.4. RMSD da cadeia glucana em relação à configuração inicial para os diferentes sistemas estudados. Cores diferentes representam diferentes simulações do mesmo sistema.

Nas simulações com substrato celohexaose (C6) faltando o resíduo no subsítio -4 (C6-Tyr74) e a outra faltando no subsítio +3 (C6-Trp161), com os RMSDs das cadeias glucana mostrados nas figuras 4.4D e 4.4E, analisamos o tempo de ligação do substrato na configuração produtiva (figura 4.3A), para avaliar a importância dos resíduos nos subsítios -4 e +3 para ligação do substrato. Nestas simulações, as cadeias glucana chegaram na configuração não-produtiva em um tempo menor que 50ns. Isso indica a necessidade de uma cadeia de pelo menos 7 monômeros para que a enzima possa se ligar, na configuração produtiva, pelo tempo necessário da catálise.

Nas simulações dos mutantes Asp92Ala, Tyr25Ala, Tyr74Ala e Trp161Ala, com os RMSDs das cadeias glucana mostrados nas figuras 4.4F,G,H, analisamos o tempo de ligação do substrato na configuração produtiva, pois dados experimentais indicaram que estes resíduos atuam no processo de ligação ao substrato (figura 4.3A).

Todos os mutantes assumiram configuração não-produtiva numa escala de tempo muito menor que a forma nativa (figura 4.3A), mostrando que tais resíduos são importantes para que o substrato permaneça ligado à enzima na configuração produtiva.

No mutante Trp161Ala (subsítio negativo), o substrato assumiu configuração não produtiva pela perda de interação com os subsítios positivos, o que não era esperado. Em uma investigação mais detalhada, observamos pequeno deslocamento do substrato na direção dos sítios negativos, causado pela falta do Trp161, fazendo com que o substrato deixasse de fazer ligação de hidrogênio com o resíduo Gly138 e diminuísse a interação com a Tyr74, que estão nos subsítios positivos, justificando assim a perda de interação com os subsítios positivos.

4.3.2 Mutante Asn99Asp

Comparamos as simulações do mutante Asn99Asp (base catalítica) com a forma nativa, para verificar possíveis alterações estruturais da enzima que explicassem a baixa atividade deste mutante. As simulações do complexo PcCel45A-C7 nativo mostraram que o resíduo Asn99 faz ligação de hidrogênio, através de seu grupo NH₂, com o carbonil da cadeia principal do resíduo Pro95. A figura 4.5A mostra a conformação da Asn99 em uma pose da simulação do complexo PcCel45A-C7 (em azul claro). Com isso, a Asn99 permanece estável, na conformação da estrutura cristalográfica, durante toda a simulação.

Já no mutante Asn99Asp, o resíduo Asp99 sofre repulsão eletrostática com o grupo carbonil da Pro95. Com isso, o Asp99 se move para fora do sítio catalítico e faz ligações de hidrogênio com o grupo NH₂ dos resíduos Glu101 e Phe102, conforme mostrado na figura 4.5A (em vermelho). Através do RMSD também é possível observar que, na forma nativa, Asn99 adota uma conformação bem definida, como esperado de um resíduo catalítico, enquanto que o mutante é extremamente móvel (figura 4.5B). Portanto, isto impede que o aspartato atue como base em uma reação catalítica pelo mecanismo de inversão e fornece explicação para os dados experimentais.

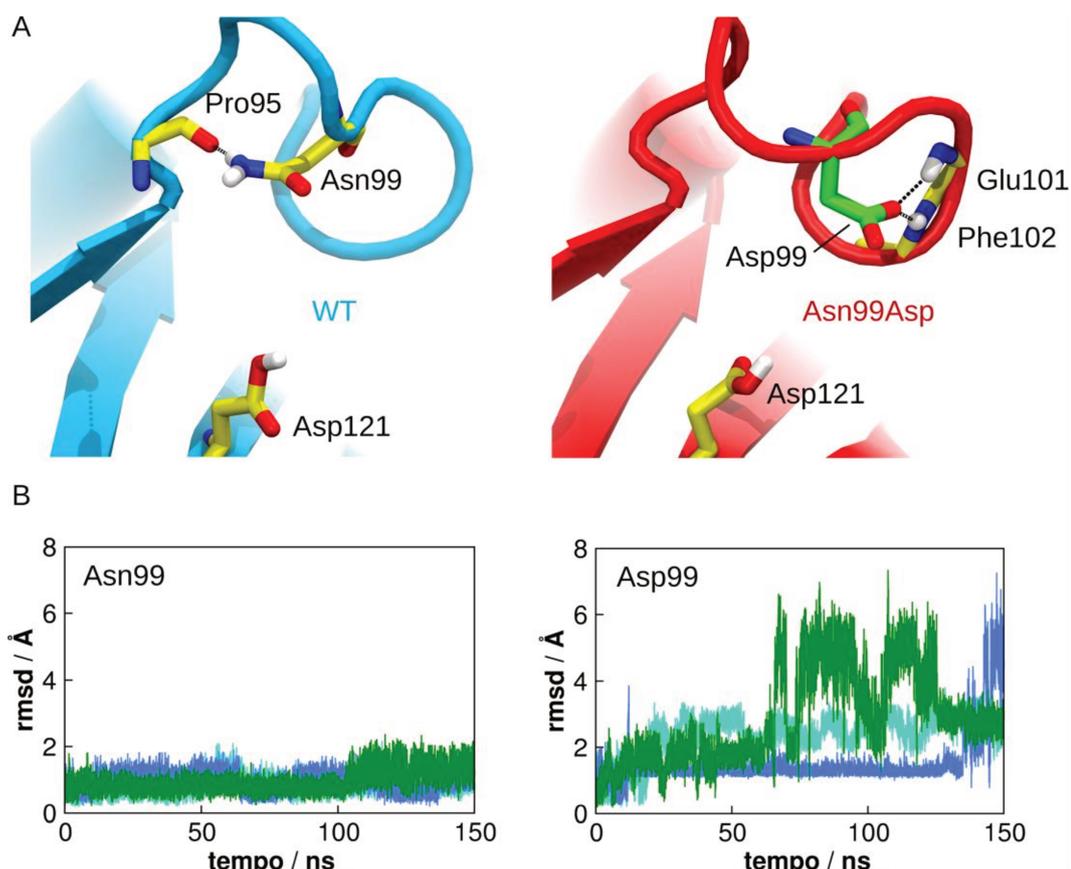


Figura 4.5. (A) Pose obtida das simulações dos complexos PcCel45A-C7 e Ans99Asp-C7, mostrando a conformação do resíduo Asn99 (à esquerda) e do Asp99 (à direita). (B) RMSD dos resíduos Asn99 (à esquerda) e Asp99 (à direita). Cores diferentes correspondem a simulações independentes do mesmo sistema.

4.3.3 Restauração da forma ativa

Para observar as possíveis alterações na dinâmica conformacional dos *loops*, fizemos simulações da enzima na forma nativa e sem nenhum substrato. Observamos que o *loop* da Asn99 exibiu duas conformações bem definidas. A figura 4.6A mostra a sobreposição desse *loop* nas duas configurações observadas durante as simulações e a figura 4.6B mostra dois picos definidos na distribuição das distâncias entre os resíduos Asp121 e Asn99. Uma das configurações observadas é equivalente à estrutura cristalográfica (em azul claro), onde a enzima estaria pronta para ligar ao substrato (figura 4.6A). Na segunda configuração (em vermelho), menos frequente, a conformação do *loop* permite que os resíduos Asn99 e Asp121 (base e ácido catalíticos) interajam entre si por meio de uma ligação de hidrogênio e a distância entre eles é em torno de 2 Å (figura 4.6A).

O histograma da figura 4.6B mostra que as duas conformações são muito específicas e que a frequência na segunda conformação é menor que na conformação cristalográfica. O tempo de permanência na segunda conformação é de alguns nanossegundos, como mostrado da figura 4.6B. Este tempo é suficiente para ocorrer transferência de prótons entre os resíduos [64]. Como após a catálise a enzima fica inativa até que ocorra a protonação do ácido e a desprotonação da base, este pode ser o mecanismo utilizado pela enzima para restituir a forma catalítica ativa.

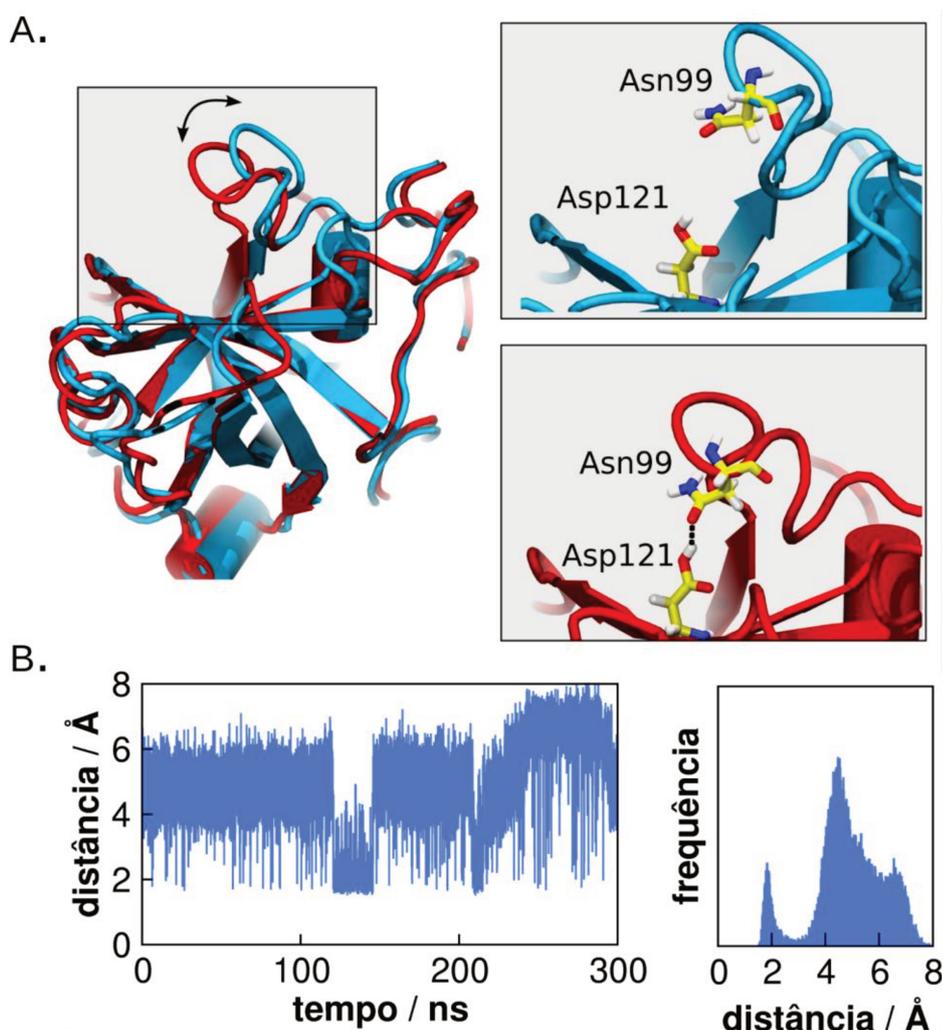


Figura 4.6. (A) Duas conformações do *loop*, em que se encontra o resíduo catalítico, observadas durante as simulações de DM e a interação entre o ácido catalítico Asp121 e o nucléofíco Asn99 nas diferentes conformações. (B) Distância entre Asn99 e Asp121 em função do tempo e sua distribuição.

4.4 Discussão

As simulações mostram que a enzima interage com o substrato principalmente por ligações de hidrogênio entre os grupos hidroximetílicos exocíclicos do substrato e os resíduos Gly54, Asp92 e Gly138, mas também com interações de *stacking* com os resíduos Tyr25, Tyr74 e Trp161. Com isso, cadeias xilana não se ligam à enzima por tempo suficiente para que ocorra clivagem, corroborando com os dados experimentais que mostram a falta e atividade catalítica da enzima sobre xilanos.

Na falta de *loops* maiores, os resíduos hidrofóbicos Trp161 e Tyr74 são essenciais para a ligação do substrato à PcCel45A. O Trp161 é conservado nas outras enzimas da família GH45, mas a Tyr74 é observada apenas na PcCel45A. Isto explica o comportamento das outras enzimas da família GH45, que possuem *loops* maiores, e então clivam cadeias com apenas 6 monômeros [30]. Já a PcCel45 necessita de uma cadeia com pelo menos 7 monômeros, para que o substrato se ligue nos subsítios -4 e +3. Isso corrobora com os resultados experimentais que mostram que a PcCel45A tem baixa atividade catalítica sobre celohexaose (C6), que é um dos produtos de clivagem da PcCel45A sobre polissacarídeos maiores [58].

Comparando a ExpB1 com a PcCel45A, a ausência de atividade catalítica detectável, até o momento, da β -expansina não está relacionada à falta do resíduo nucleófilo, pois o Glu99 está presente na ExpB1 na posição correspondente ao nucleófilo, Asn99, da PcCel45A. A falta de *loops* maiores e do resíduo Tyr74 na ExpB1 pode indicar limitação da enzima na ligação de β -1,4-glucano. Isso sugere que a enzima se ligue a outro tipo de substrato ou que ela necessite de outras substâncias, presentes na parede celular, para se ligar ao substrato.

A baixa atividade do mutante Asn99Asp não é justificada quimicamente, já que o aspartato possui maior basicidade que a asparagina. As simulações de DM mostraram que a enzima tem estrutura necessária para estabilizar a asparagina na posição catalítica através de ligações de hidrogênio muito específicas. Estas interações não são aplicadas ao aspartato e fazem com que ele apresente mudanças conformacionais que o impedem de atuar, efetivamente, como base na reação catalítica. Isto mostra que a enzima é complexa e que sua atividade catalítica depende da interação dos resíduos com todo o ambiente químico da enzima, e não apenas dos resíduos isolados.

A enzima fica inativa após a reação pelo mecanismo de inversão, devido ao fato da base estar protonada e o ácido catalítico desprotonado. Se o substrato se ligar na enzima na forma inativa, a reação de catálise não ocorre e o substrato leva um tempo maior para se desligar, tempo ao qual a enzima está inibida. Nas simulações de DM observamos que a enzima tem um mecanismo próprio para voltar à forma ativa, sem depender da interação com a água. O *loop* do resíduo base possui uma conformação muito específica, onde a base e o ácido catalítico fazem ligação de hidrogênio e podem realizar a troca de prótons.

4.5 Conclusão

A PcCel45A é a enzima da família GH45 com maior similaridade com expansinas e que se diferencia das outras enzimas desta família pela falta de *loops* e pelo fato de seu resíduo catalítico que atua como base ser uma asparagina. Neste trabalho mostramos que a falta de *loops* é compensada pelos resíduos hidrofóbicos Trp161 e Tyr74 no processo de ligação ao substrato e também que as ligações de hidrogênio entre a enzima e os grupos hidroximetílicos do substrato são essenciais para que o substrato se ligue por tempo suficiente para que ocorra a catálise. Este tipo de interação enzima-substrato justifica as limitações da enzima na clivagem de cadeias xilanas e glucana com 6 ou menos monômeros. A baixa atividade do mutante Asn99Asp é causada pela perda de ligações de hidrogênio do resíduo Asp99 com a cadeia principal da enzima, fazendo com que o mutante apresente estabilidade conformacional para que o resíduo 99 atue como base em uma catálise enzimática. O *loop* do base Asn99 apresentou duas conformações distintas, uma equivalente à estrutura cristalina e outra que permite uma ligação de hidrogênio entre a base e o ácido catalítico, indicando que a enzima tem um mecanismo para voltar à forma ativa, sem depender de interações com a água.

Capítulo 5

Considerações finais

Para a concepção de novos materiais e desconstrução da biomassa para produção de biocombustíveis é necessário compreender a arquitetura da parede celular e as modificações causadas por enzimas na estrutura da parede. Nesta dissertação de mestrado, utilizamos simulações computacionais de dinâmica molecular para estudar, do ponto de vista atômico-molecular, as interações físicas da matriz polimérica da parede celular com a celulose e aspectos mecanísticos de enzimas que atuam sobre polissacarídeos da parede celular. Especificamente, estudamos a adsorção da hemicelulose de coníferas, composta por glucuronoarabinoxilanos, em celulose, e a endoglucanase Cel45A do fungo *Phanerochaete chrysosporium*.

As simulações de dinâmica molecular da adsorção de xilanos com padrão de ramificação encontrado em coníferas mostraram que glucuronoarabinoxilanos se adsorvem estavelmente na superfície hidrofílica (110) da celulose. As ramificações de α -1,2-ácido glucurônico estabilizam a adsorção e as ramificações de α -1,3-arabinose aumentam a mobilidade dos diedros glicosídicos e não contribuíram para adsorção na superfície hidrofílica da celulose. Nas simulações com duas cadeias xilanas em paralelo há efeitos cooperativos para a adsorção na celulose, causados pelos efeitos de empacotamento e, no caso das cadeias ramificadas, também pela formação de reticulações entre ramificações α -1,2-ácido glucurônico e os íons cálcio.

Os estudos da PcCel45A mostraram que sua ligação às cadeias glucana ocorre principalmente por ligações de hidrogênio com os grupos hidroximetílicos exocíclicos interações hidrofóbicas do tipo *stacking*. Na falta de *loops* para ligação do substrato, a interação do substrato com os resíduos hidrofóbicos Trp161 e Tyr74 é essencial para que o substrato permaneça ligado por tempo suficiente para que ocorra a catálise. Isso justifica as limitações da enzima na clivagem de alguns substratos. O mutante Asn99Asp não apresenta estabilidade conformacional do resíduo 99 para

atuar como nucleófilo em uma catálise enzimática. O *loop* do nucleófilo Asn99 apresentou duas conformações distintas, a conformação da estrutura cristalina e uma conformação que permite uma ligação de hidrogênio entre o nucleófilo e o ácido catalítico, que pode indicar um mecanismo da enzima para voltar à forma ativa.

Perspectivas

Os estudos apresentados neste trabalho abriram novas possibilidades para uma série de novos estudos que proporcionariam um aumento no entedimento dos fundamentos parede celular de plantas e sua modificação por enzimas e proteínas acessórias. Em relação aos estudos dos xilanos, duas novas questões seriam de grande interesse de serem estudadas. Uma delas se refere à possibilidade de xilanos com ramificações em posições ímpares se ligarem à celulose. Uma outra seria abordar efeitos relacionados aos diferentes polimorfos da celulose, como celulose I α vs. I β , e a presença de falhas de empilhamento (*stacking faults*) na adsorção de xilanos. Em relação às enzimas da família GH45 e expansinas, um próximo passo seria realizar estudos para entender a ligação do substrato às expansinas de plantas e outras GH45, como a HiCel45 que, apesar de ter sido caracterizada experimentalmente, ainda não foram feitos estudos em nível molecular para explicar questão da interação dos *loops* com o substrato.

Referências Bibliográficas

1. Chundawat, S.P.S., Beckham, G.T., Himmel, M.E., Dale, B.E. (2011) Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annu. Rev. Chem. Bio Eng* 2:121-145.
2. Zhang Y. -H.P., Himmel M.E., Mielenz J.R. (2006) Outlook for cellulase improvement: screening and selection strategies. *Biotech. Adv.* 24:452-481.
3. Arantes V., Saddler J.M. (2010) Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol. Biofuels* 3:4.
4. Zhang, Y.H., Lynd, L.R. (2004) Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnol. Bioeng.* 88:797-824.
5. Mccann, M., Carpita, N. (2008) Designing the deconstruction of plant cell walls. *Curr. Opin. Plant Biol.* 11:314–320.
6. Heinze, T. (2015) Cellulose: structure and properties. *Adv. Polym. Sci.* 271:1-52.
7. Atalla, R.H., Vanderhart, D.L. (1984) Native cellulose: A composite of two distinct crystalline forms. *Science* 223: 283-85.
8. Cosgrove, D.J. (2014) Reconstructing our models of cellulose and primary cell wall assembly. *Curr. Opin. Plant Biol.* 22:122–131.
9. Fernandes, A.N., Thomas, L.H., Altaner, C.M., Callow, P., Forsyth, V.T., Apperley, D.C., Kennedy, C.J., Jarvis, M.C. (2011) Nanostructure of cellulose microfibrils in spruce wood. *Proc. Natl. Acad. Sci. USA* 108:E1195-E1203.
10. Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W., Foust, T.D. (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315:804–807.
11. Hu, F., Ragauskas, A. (2012) Pretreatment and lignocellulosic chemistry. *Bioenerg. Res.* 5:1043–1066.
12. Scheller, H.V., Ulvskov, P. (2010) Hemicelluloses. *Annu. Rev. Plant Biol.* 61: 263–289.
13. Fry, S. C. (2001) Plant Cell Walls. eLS.
14. Burton, R.A., Gidley M.J., Fincher, G.B. (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat. Chem. Biol.* 6:724–732.
15. Zhao, Z. Zhao, Z., Crespi, V., Kubicki, J., Cosgrove, D., Zhong, L. (2013) Molecular dynamics simulation study of xyloglucan adsorption on cellulose surfaces: effects of surface hydrophobicity and side-chain variation. *Cellulose* 21: 1025-039.
16. Busse-Wicher, M., Gomes, T.C.F., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N.J., Bolam, D.N., Skaf, M.S., Dupree, P. (2014) The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of *Arabidopsis thaliana*. *Plant J.* 79:492–506.
17. Balan, V. Sousa, L. da C., Chundawat, S.P., Marshall, D., Sharma, L.N., Chambliss,

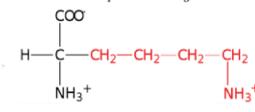
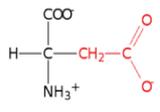
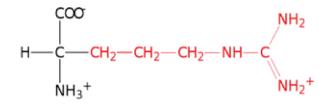
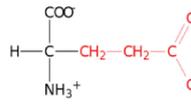
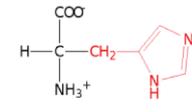
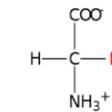
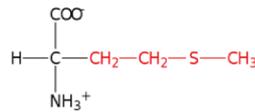
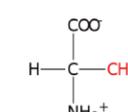
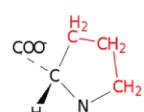
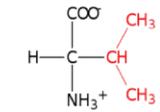
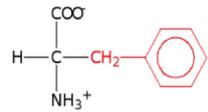
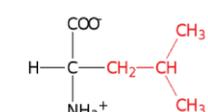
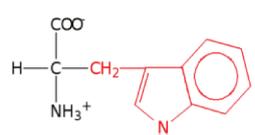
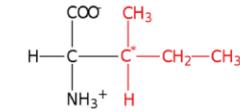
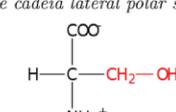
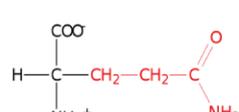
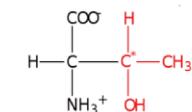
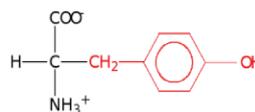
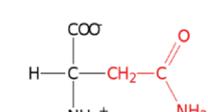
- C.K., Dale, B.E. (2009) Enzymatic digestibility and pretreatment degradation products of AFEX treated hardwoods (*Populus Nigra*). *Biotechnol. Progress* 25:365–375.
18. Cosgrove, D.J., Jarvis, M.C. (2012) Comparative structure and biomechanics of plant primary and secondary cell walls. *Front. Plant Sci.* 3:204
 19. Malinovsky, F.G., Fangel, J.U., Willats, W.G.T. (2014) The role of the cell wall in plant immunity. *Front. Plant Sci.* 5:178.
 20. Park, Y.B., Cosgrove, D.J. (2012) A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant Physiol.* 158:1933-943.
 21. Wolf, S., Kian, H., Höfte, H. (2012) Growth control and cell wall signaling in plants. *Annu. Rev. Plant Biol.* 63:381–407
 22. Cho, H.-T., Cosgrove, D.J. (2000) Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 97:9783-788.
 23. Cosgrove, D.J. (2015) Plant expansins: diversity and interactions with plant cell walls. *Curr. Opin. Plant Biol.* 25:162–172
 24. Gourlay, K., Arantes, V., Saddler, J.N. (2012) Use of substructure specific carbohydrate binding modules to track changes in cellulose accessibility and surface morphology during the amorphogenesis step of enzymatic hydrolysis. *Biotechnol. Biofuels* 5:1
 25. Georgelis, N., Tabuchi, A., Nikolaidis, N., Cosgrove, D.J. (2011) Structure-function analysis of the bacterial expansin EXLX1. *J. Biol. Chem.* 286:16814-6823.
 26. Silveira, R. L., Skaf, M. S. (2016) Molecular dynamics of the *Bacillus subtilis* expansin EXLX1: interaction with substrates and structural basis of the lack of activity of mutants. *Phys. Chem. Chem. Phys.* 18:3510-521.
 27. Sampedro, J., Guttman, M., Li, L.C., Cosgrove, D.J. (2015) Evolutionary divergence of β -expansin structure and function in grasses parallels emergence of distinctive primary cell wall traits. *Plant. J.* 81:108–120
 28. Yennawar, N. H., Li, L.C., Dudzinski, D.M., Tabuchi, A., Cosgrove, D.J. (2006) Crystal structure and activities of EXPB1 (*Zea m 1*), a β -expansin and Group1 Pollen Allergen from maize. *Proc. Natl. Acad. Sci.* 103:14664–14671.
 29. Payne, C.M., Knott, B.C., Mayes, H.B., Hansson, H., Himmel, M.E., Sandgren, M., Ståhlberg, J., Beckham, G.T. (2015) Fungal cellulases. *Chem. Rev.* 115:1308–1448.
 30. Xu, B., Hellman, U., Ersson, B., Janson, J.C. (2000) Purification, characterization and amino acid sequence analysis of a thermostable, low molecular mass endo β 1,4-glucanase from Blue mussel, *Mytilus edulis*. *Eur. J. Biochem.* 267:4970–4977
 31. Koshland, D.E. (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* 28:416–436
 32. Davies, G.J., Tolley, S.P., Henrissat, B., Hjort, C., Schülein, M. (1995) Structures of oligosaccharide bound forms of the endoglucanase V from *Humicola insolens* at 1.9Å resolution. *Biochemistry* 34:16210–16220
 33. Dror, R. O., Dirks, R.M., Grossman, J.P., Xu, H., Shaw, D.E. (2012) Biomolecular

- simulation: a computational microscope for molecular biology. *Annu. Rev. Biophys.* 41: 429-452.
34. Karplus, M., Mccammon, J.A. (2002) Molecular dynamics simulations of biomolecules. *Nat. Struct Biol.* 9: 646-652.
 35. Leach, A. R. (2001) *Molecular modelling: principles and applications*. Harlow: Pearson Prentice Hall.
 36. Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., Mackerell, A.D.Jr. (2010) CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* 31: 671-690.
 37. Hornak, V. Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C. (2006) Comparison of multiple Amber Force Fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* 65:712-25.
 38. Martínez, L., Andrade, R., Birgin, E.G., Martínez, J.M. (2009) Packmol: a package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.*, 30:2157-2164.
 39. Frenkel, D., Smit, B. (2002) *Understanding molecular simulation: from algorithms to applications*. San Diego: Academic Press.
 40. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26: 1781-802.
 41. Feller, S. E., Zhang, Y., Pastor, R.W., Brooks, B.R. (1995) Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* 103:4613-4621.
 42. Hoover, W.G. (1986) Constant-pressure equations of motion. *Phys. Rev. A*34: 2499.
 43. Sundararajan, P.R., Rao, V.S.R. (1969) Conformational studies of β -D-1,4'-xylan. *Biopolymers* 8: 305–312
 44. Kabel, M.A., van den Borne, H., Vincken, J.P., Voragen, A.G.J., Schols, H.A. (2007) Structural differences of xylans affect their interaction with cellulose. *Carb. Polym.* 69:94–105
 45. Mortimer, J. C. Miles, G.P., Brown, D.M., Zhang, Z., Segura, M.P., Weimar, T., Yu, X., Seffen, K.A., Stephens, E., Turner, S.R., Dupree, P. (2010) Absence of branches from xylan in *Arabidopsis* Gux mutants reveals potential for simplification of lignocellulosic biomass. *Proc. Natl. Acad. Sci.* 107:17409–17414
 46. Rennie, E. A., Scheller, H. V. (2014) Xylan biosynthesis. *Curr. Opin. Biotechnol.* 26:100-107
 47. Bromley, J.R., Busse-Wicher, M., Tryfona, T., Mortimer, J.C., Zhang, Z., Brown, D.M. and Dupree, P. (2013) GUX1 and GUX2 Glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *Plant J.* 74: 423-434.
 48. Chong, S.-L. (2014) O-acetylation of glucuronoxylan in *Arabidopsis thaliana* wild type and its change in xylan biosynthesis mutants. *Glycobiology* 3:1–13.
 49. Gomes, T.C.F., Skaf, M.S. (2012) Cellulose-Builder: a toolkit for building crystalline

- structures of cellulose. *J. Comput. Chem.* 33:1338-1346.
50. Ding, S.-Y., Himmel, M.E. (2006) The maize primary cell wall microfibril: a new model derived from direct visualization. *J. Agric. Food Chem.* 54:597- 606.
 51. Nishiyama, Y., Langan, P., Chanzy, H. (2002) Crystal structure and hydrogen-bonding system in cellulose I β from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* 124:9074-9082.
 52. Guvench, O., Hatcher, E., Venable, R. M., Pastor, R. W., MacKerell, Jr., A. (2009) D. CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses. *J. Chem. Theory Comput.* 5:2353-2370.
 53. Raman, E. P., Guvench, O., MacKerell, Jr., A. D. (2010) CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J. Phys. Chem. B* 114:12981-12994.
 54. Darden, P., York, D., Pedersen, L. (1993) Particle mesh Ewald: an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089-10092.
 55. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926-935.
 56. Humphrey, W., Dalke, A., Schulten, K. (1996) VMD – Visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
 57. Rees, D. A., Scott, W. E. (1971) Polysaccharide conformation. Part VI. Computer model-building for linear and branched pyranoglycans. Correlations with biological function. Preliminary assessment of inter-residue forces in aqueous solution. Further interpretation of optical rotation in terms of chain conformation. *J. Chem. Soc. B* 469-479.
 58. Igarashi, K., Ishida, T., Hori, C., Samejima, M. (2008) Characterization of an endoglucanase belonging to a new subfamily of glycoside hydrolase family 45 of the basidiomycete *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* 74:5628–5634
 59. Nakamura, A. Ishida, T., Samejima, M., Igarashi, K. (2015) ‘Newton's cradle’ proton relay with amide imidic acid tautomerization in inverting cellulase visualized by Neutron crystallography. *Science Adv.* 1:e1500263
 60. Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S., Onufriev, A. (2005) H⁺⁺: A server for estimating pK_as and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* 33:W368–W371.
 61. Myers, J., Grothaus, G., Narayanan, S., Onufriev, A. (2006) A simple clustering algorithm can be accurate enough for use in calculations of pK_as in macromolecules. *Proteins* 63:928–938.
 62. MacKerell, A.D.Jr., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102:3586–3616.
 63. MacKerell, A.D. Jr., Feig, M., Brooks, C.L., III. (2004) Extending the treatment of

- backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* 25:1400–1415.
64. Cukier, R. I. (2004) Theory and simulation of proton-coupled electron transfer, hydrogen-atom transfer, and proton translocation in proteins. *Biochim. Biophys. Acta* 1655:37-44.
 65. Driemeier, C., Francisco, L. H. (2015) X-ray diffraction from faulted cellulose I constructed with mixed $I\alpha$ vs. $I\beta$ stacking. *Cellulose* 21:3161-3169.

ANEXO: Nomenclatura e estrutura dos aminoácidos

Nome, Abreviações	Fórmula estrutural	Nome, Abreviações	Fórmula estrutural
<i>Aminoácidos de cadeia lateral polar carregada</i>			
Lisina Lys K		Ácido aspártico Asp D	
Arginina Arg R		Ácido Glutâmico Glu E	
Histidina His H			
<i>Aminoácidos de cadeia lateral apolar</i>			
Glicina Gly G		Metionina Met M	
Alanina Ala A		Prolina Pro P	
Valina Val V		Fenilalanina Phe F	
Leucina Leu L		Triptofano Trp W	
Isoleucina Ile I			
<i>Aminoácidos de cadeia lateral polar sem carga</i>			
Serina Ser S		Glutamina Gln Q	
Treonina Thr T		Tirosina Tyr Y	
Asparagina Asn N		Cisteína Cys C	