

Universidade Federal do Rio Grande do Sul
Instituto de Biociências
Bacharelado em Biotecnologia - Ênfase em Bioinformática

João Luiz de Meirelles

OTIMIZAÇÃO DO CAMPO DE FORÇA GROMOS PARA
SIMULAÇÃO DE HEXOPIRANOSSES E PENTOFURANOSSES
POR MEIO DE CORREÇÕES EMPÍRICAS DE
PARÂMETROS TORCIONAIS.

Porto Alegre
2021

João Luiz de Meirelles

OTIMIZAÇÃO DO CAMPO DE FORÇA GROMOS PARA
SIMULAÇÃO DE HEXOPIRANOSSES E
PENTOFURANOSSES POR MEIO DE CORREÇÕES
EMPÍRICAS DE PARÂMETROS TORCIONAIS

Trabalho de Conclusão de Curso submetido à Universidade Federal do Rio Grande do Sul, como requisito necessário para obtenção do grau de Bacharel em Biotecnologia com Ênfase em Bioinformática

Porto Alegre, maio de 2021

CIP - Catalogação na Publicação

de Meirelles, João Luiz
João Luiz de Meirelles OTIMIZAÇÃO DO CAMPO DE FORÇA
GROMOS PARASIMULAÇÃO DE HEXOPIRANÓSES EPENTOFURANÓSES
POR MEIO DE CORREÇÃO EMPÍRICAS DE PARÂMETROS
TORCIONAIS / João Luiz de Meirelles. -- 2021.
44 f.
Orientador: Hugo Verli.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Instituto
de Biociências, Curso de Biotecnologia:
Bioinformática, Porto Alegre, BR-RS, 2021.

1. Bioinformática Estrutural. 2. Carboidratos. 3.
Computação. I. Verli, Hugo, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

João Luiz de Meirelles

Esta Monografia foi julgada adequada para a obtenção do título de Bacharel em Biotecnologia com Ênfase em Bioinformática, sendo aprovada em sua forma final pela banca examinadora:

Orientador(a): Prof. Dr. Hugo Verli
Universidade Federal do Rio Grande do Sul -
UFRGS

Prof. Dr. Paulo Augusto Netz
Universidade Federal do Rio Grande do Sul -
UFRGS

Prof. Dr. Rodrigo Ligabue-Braun
Universidade Federal de Ciências da Saúde
de Porto Alegre - UFCSPA

Porto Alegre, maio de 2021

Agradecimentos

Um agradecimento especial ao Professor Dr. Hugo Verli, meu orientador. O meu tempo no laboratório foi vital para minha evolução profissional e como pessoa. Podendo trabalhar em diversos projetos incríveis sobre a tua orientação foi engrandecedor, permitindo aprender diversas ferramentas, habilidades e conceitos que me possibilitam produzir ciência e expandir os meus horizontes e aspirações. Muito obrigado!

Ao Professor Dr. Paulo Augusto Netz e ao Professor Dr. Rodrigo Ligabue-Braun pelo aceite do convite e pelas orientações para melhora do meu trabalho!

Agradeço aos meus pais Marilene e Luiz João e meus irmãos Pedro e Maurício. Vocês são as pessoas mais importantes na minha vida, amo muito vocês! Também cabe um agradecimento ao meu sobrinho Oliver, por me alegrar todo dia com sua fofura.

Aos meus colegas de curso por estarem comigo durante todos os semestres e cadeiras. Por todas as partidas de truco esperando o horário da próxima aula. Em especial, um agradecimento a Reykla, Igor, Daniel e Yuri.

Por último o meu agradecimento aos integrantes do laboratório GBE: Felipe, Crisciele, Brisa, Laís, Thiago, Marcelo e Shabir. Vocês foram e são muito importantes por ter me possibilitado aprender muito e obrigado por serem sempre solícitos as minhas várias perguntas.

If I have seen further than others, it is by standing upon the shoulders of giants.

Isaac Newton

Resumo

Carboidratos são biomoléculas extremamente importantes, participando da estrutura de ácidos nucleicos, glicoproteínas, reconhecimento célula-célula ou no desenvolvimento de doenças. Não existem métodos simples para sua síntese, identificação e obtenção. Neste contexto, uma das ferramentas para o estudo do papel funcional de carboidratos reside em ferramentas computacionais como a dinâmica molecular (DM). A acurácia destes modelos reside diretamente na qualidade dos parâmetros empregados, chamados de campos de força. O campo de força 53A6 GLYC, desenvolvido previamente pelo nosso grupo, é parametrizado para a simulação em alta fidelidade de hexoses. Porém, este campo de força apresenta limitações, como a restrita reprodução da cinética de interconversão entre as formas que um anel monossacarídeo pode adotar e falta de parâmetros para pentofuranoses. Assim, o objetivo deste trabalho é a re-parametrização do campo de força GROMOS para simulação de hexopiranoses e pentofuranoses por meio da correção empírica de parâmetros torsionais, buscando ajustá-los para melhor reprodução do dado experimental de RMN. Para tal, todos os sistemas foram simulados utilizando o pacote GROMACS. Para a execução de metadinâmicas, foi utilizado o software PLUMED. As análises de θ e ϕ do puckering e da constante de acoplamento (${}^3J_{H,H}$) durante as dinâmicas moleculares foram feitas com implementação autoral em Python, sendo o cálculo da constante de acoplamento baseado na equação de Hasnoot-Altona. Os dados de abundância de cadeiras e ${}^3J_{H,H}$ extraídos em diferentes valores de constante torcional foram comparados com dados experimentais de diferentes monossacarídeos em RMN, sendo escolhido o valor que se mostrou o mais próximo ao dado experimental. Com isso, a cinética de interconversão entre as formas de um anel monossacarídico do modelo se mostrou superior aos parâmetros anteriores. Os novos parâmetros possibilitam a representação mais fiel de sistemas envolvendo carboidratos, auxiliando futuros estudos de suas funções biológicas e terapêuticas.

Palavras-chave: Carboidratos, Dinâmica Molecular, Metadinâmica, Campo de Força.

Abstract

Carbohydrates are biomolecules important in nucleic acid structure, glycoproteins, cell-cell recognition, and disease development. There are no easy methods to their synthesis, recognition and obtention. In this context, computational tools, such as molecular dynamics, are essential to study carbohydrate functional role. The accuracy of these models is directly linked to the quality of the parameters, named the force field. The 53A6 GLYC force field, developed previously by our group, is parameterized for the high fidelity simulation of hexoses. However, this force field presents limitations, as the restricted reproduction of the interconversion kinetics that a monosaccharide ring can adopt and the lack of parameters to represent pentoses. The objective of this work is the reparameterization of the GROMOS force field to simulate hexoses and pentoses by the empiric correction of torsional parameters, adjusting them to reproduce experimental NMR data. To this end, all systems studied were simulated using GROMACS. To run metadynamics, PLUMED was utilized. The puckering θ and ϕ variables and the coupling constant (${}^3J_{H,H}$) during the molecular dynamics were calculated by a Python script. The coupling constant calculation used the Hasnoot-Altona equation. Chair conformation and ${}^3J_{H,H}$ extracted from different torsional constant values were validated by NMR experimental data, being chosen the torsional constant with most proximity to experimental data. The interconversion kinetics between the possible carbohydrates chair conformations was superior to the original parameters. The new validated parameters can represent the experimental data with more reliability, helping future studies to understand carbohydrate biological and therapeutical functions.

Keywords: Carbohydrates, Molecular Dynamics, Metadynamics, Force Field.

Lista de ilustrações

Figura 1 – Complexidade configuracional de monossacarídeos.	17
Figura 2 – Complexidade configuracional de dissacarídeos.	18
Figura 3 – <i>Puckering</i> de hexose com diferentes formas de anel.	19
Figura 4 – Diferentes termos presentes em campos de força para a representação de átomos e suas interações.	23
Figura 5 – Metadinâmica de α -Idose de 100 ns.	32
Figura 6 – Metadinâmica de Desoxirribose de 100 ns.	35
Figura 7 – Metadinâmica de α -Arabinose de 100 ns.	36

Lista de tabelas

Tabela 1	– Médias de constante de acoplamento em dinâmicas moleculares de α -Idose.	31
Tabela 2	– Médias de constante de acoplamento em dinâmicas moleculares de β -Idose.	33
Tabela 3	– Médias de constante de acoplamento de α -Altrose.	34
Tabela 4	– Posição e energia relativa dos mínimos e estado de transição em dados de QM e metadinâmica com parâmetros de GLYC 53A6.	35
Tabela 5	– Médias de constante de acoplamento de Desoxirribose.	35
Tabela 6	– Médias de constante de acoplamento de α -Arabinose.	36

Lista de abreviaturas e siglas

D	Dextrógiro
DFT	Teoria do Funcional da Densidade
DM	Dinâmica Molecular
HF	Hartree–Fock
L	Levógiro
MM	Mecânica molecular
NVT	<i>Ensemble</i> Canônico
NPT	<i>Ensemble</i> Isotérmico-Isobárico
QM	Mecânica quântica
RMN	Ressonância Magnética Nuclear

Lista de símbolos

${}^3J_{H,H}$ Constante de acoplamiento vicinal

α Alpha

β Beta

ϕ Phi

θ Theta

Sumário

1	INTRODUÇÃO	16
1.1	Carboidratos	16
1.1.1	Complexidade Configuracional	16
1.1.2	Complexidade Conformacional	17
1.2	Métodos experimentais	19
1.2.1	Cristalografia de Raios X	19
1.2.2	RMN	19
1.3	Dinâmica Molecular	20
1.3.1	Mecânica Quântica	21
1.3.2	Mecânica Molecular	21
1.4	Metadinâmica	24
2	JUSTIFICATIVA	25
3	OBJETIVOS	26
4	METODOLOGIA	27
4.1	Topologia de hexoses	27
4.2	Topologia de pentoses	27
4.3	Dinâmica Molecular	27
4.4	Metadinâmica	28
4.5	Cálculo do <i>Puckering</i>	28
4.6	Cálculo de ${}^3J_{H,H}$	28
5	RESULTADOS	30
5.1	Simulações de hexoses	30
5.1.1	Validação da hipótese	30
5.1.2	<i>Screening</i> de α -Idose	31
5.1.3	<i>Screening</i> de β -Idose	33
5.1.4	<i>Screening</i> de α -Altrose	33
5.2	Simulações de pentoses	34
5.2.1	Desoxirribose	34
5.2.2	α -Arabinose	35
6	CONCLUSÃO	37
7	PERSPECTIVAS	38

7.1	Programa para cálculo de ${}^3J_{H,H}$	38
7.2	Parâmetros de α-Altrose	38
7.3	Exploração de outros monossacarídeos	38
7.4	Desoxirribose	38
	REFERÊNCIAS	39

1 Introdução

1.1 Carboidratos

Carboidratos constituem uma importante classe de biomacromoléculas assim como proteínas, lipídeos e ácidos nucleicos. Seu nome, com referência a carbonos hidratados, é derivado da sua fórmula geral $C_x(H_2O)_y$. São as biomacromoléculas mais abundantes na natureza, sendo presentes em diversos sistemas e processos biológicos como função estrutural, reconhecimento célula-célula [1], sinalização celular e no desenvolvimento de doenças [2]. São compostos majoritariamente de carbono, oxigênio e hidrogênio, presentes em um equilíbrio entre formas acíclicas e cíclicas em solução, com predominância da forma cíclica em anel com 99.9% de ocorrência [3]

São moléculas que podem apresentar tanto simplicidade como monômeros livres, quanto alta complexidade como polissacarídeos ramificados. Mais de centenas de unidades monoméricas são encontradas na natureza [4], uma diversidade maior que 22 aminoácidos naturais e 5 bases nitrogenadas.

Mesmo sendo tão abundantes em sistemas vivos, o entendimento de sua estrutura, dinâmica e função ainda é muito limitado. Isso se dá por vários desafios técnicos e experimentais no seu estudo devido à sua inerente complexidade química [5] [6]. Porém, avanços em tecnologias como RMN, cristalografia de raios X e modelagem molecular tornam possível o estudo de sua estrutura e dinâmica em resolução atômica, ajudando a preencher essa lacuna de conhecimento.

1.1.1 Complexidade Configuracional

Tudo começa em sua estrutura mais simples, o monossacarídeo. Sua forma cíclica pode ser composta de seis átomos, chamada de piranose, ou de cinco átomos, chamada de furanose. O seu monômero varia pelo número de átomos de carbono da cadeia, como trioses, tetroses, pentoses (por exemplo, ribose) e hexoses (por exemplo, glicose).

Por conta de múltiplos centros quirais em sacarídeos, sua complexidade configuracional se dá pelo arranjo de substituintes nesses centros. Para diferenciação entre glicose, manose e idose, por exemplo, a posição dos substituintes pode ser equatorial ou axial (Fig. 1A). Possuem dois enantiômeros D e L, os dois possíveis estereoisômeros de cada monossacarídeo (Fig. 1B), sendo na natureza o enantiômero D o mais abundante, e dois estados anoméricos α e β , caracterizados pela orientação da hidroxila no carbono 1 (Fig. 1C). Uma maior diversidade química pode ser adicionada por meio de fosforilação [7] ou sulfatação [8] dos grupos hidroxila.

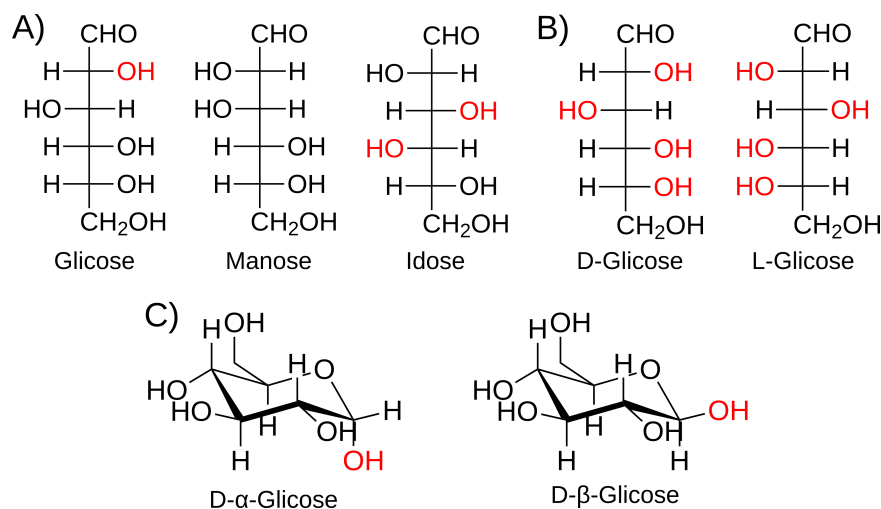


Figura 1 – Complexidade configuracional de monossacarídeos. A) Variação de monossacarídeos de hexose. Em vermelho as diferentes orientações da hidroxila que diferenciam os monômeros. B) Enantiômeros de glicose, em vermelho as orientações de hidroxilas espelhadas e C) estados anoméricos α e β , caracterizados pela orientação da hidroxila no carbono 1 em vermelho.

Uma estrutura mais complexa é o dissacarídeo, dois monômeros ligados por meio de uma ligação glicosídica. Exemplos comuns são a lactose e a maltose. A ligação glicosídica é formada entre um hemiacetal de um monossacarídeo e uma hidroxila de outro monossacarídeo. Diferente de aminoácidos e nucleotídeos, onde a ligação sempre ocorre entre os mesmos dois átomos e pode ser recuperada em uma ordem linear, a ligação glicosídica pode ocorrer entre o carbono um e o carbono quatro, como também entre o carbono um e carbono três, por exemplo.

Na figura 2, podemos ver um exemplo de dissacarídeo, onde a ligação glicosídica se dá entre o carbono um do anel com o carbono três do outro anel (Fig. 2A). Outra complexidade é a possibilidade de ramificação, ou seja, cada monômero pode estar ligado a dois ou mais monossacarídeos (Fig. 2B).

1.1.2 Complexidade Conformacional

A palavra conformação foi utilizada pela primeira vez como definição de forma molecular por Hawort [9] para demonstrar as diferentes possíveis geometrias de um anel de glicose. A conformação de um carboidrato pode ser classificada por meio da nomenclatura descrita por Schwartz [10] e adotada pela IUPAC [11] por meio do *puckering*.

As 38 conformações de *puckering* de hexoses canônicas representam estados discretos possíveis de serem adotados pelo anel de monossacarídeos. O anel pode estar em forma de cadeira (C), envelope (E), meia-cadeira (H), bote torcido (S) e bote (B). Um número sobrescrito (4C) representa o carbono do anel acima do plano, enquanto um número subscrito (C_1) representa o carbono do anel abaixo do plano. Exemplos de cadeiras e botes

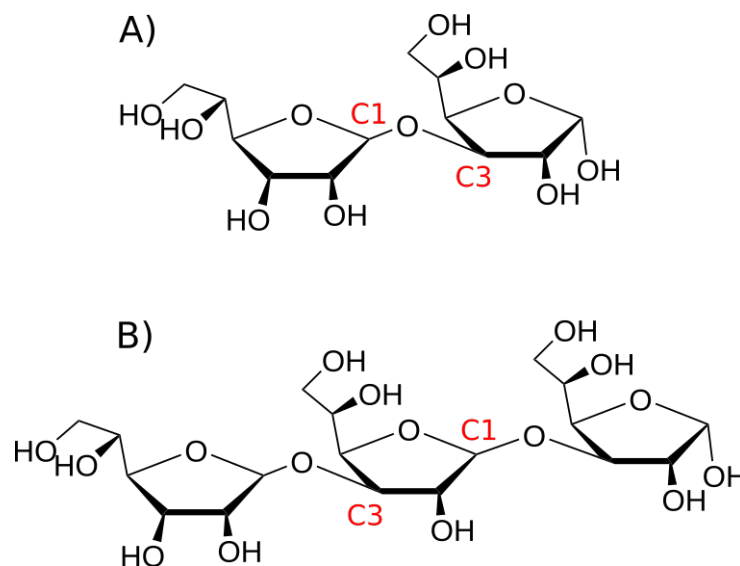


Figura 2 – Complexidade configuracional de dissacarídeos. A) Dissacarídeo com ligação entre carbono um e carbono três de diferentes monômeros, anotado em vermelho. B) Monômero central da cadeia fazendo uma ligação glicosídica no seu carbono um e outra no seu carbono três, anotados em vermelho.

são representados na figura 3.

Partindo dessas definições, Cremer e Pople [12] estabeleceram um sistema de coordenadas que possibilita o mapeamento das geometrias de *puckering* de hexose em uma esfera, sendo suas coordenadas ϕ , θ e Q . Para pentose, somente os parâmetros ϕ e θ são necessários, sendo comumente substituídas pelas coordenadas cartesianas Z_x e Z_y .

O uso dos valores de *puckering* possibilita avaliar a conformação do anel de carboidrato partindo de sua estrutura atômica, o que torna possível estudar a sua inerente flexibilidade e os diferentes fenômenos biológicos relacionados a conformação do anel [13] [14].

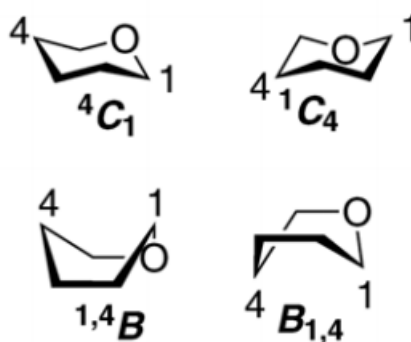


Figura 3 – *Puckering* de hexose, com exemplos de cadeira 4C_1 , onde o carbono quatro se encontra acima do plano e o carbono um abaixo do plano, e cadeira 1C_4 , onde carbono um se encontra acima do plano e o quatro abaixo. No bote ${}^{1,4}B$, o carbono um e o carbono quatro estão acima do plano, enquanto no bote $B_{1,4}$ eles estão abaixo do plano.

1.2 Métodos experimentais

1.2.1 Cristalografia de Raios X

A cristalografia de raios X é um dos métodos mais utilizados para a determinação da estrutura atômica de biomacromoléculas [15]. Cristais de moléculas são gerados a partir de uma solução concentrada. Raios X são emitidos nesses cristais, sendo o padrão de difração resultante utilizado para a elucidação do arranjo atômico. Raios X são escolhidos por seu comprimento de onda ser do tamanho de distâncias interatômicas (1 Å ou 10^{-8} cm), possibilitando uma resolução que defina estruturas atômicas, providenciando dados de ângulo de ligação, comprimento de ligação, ângulos torsionais e distâncias intramoleculares.

Contanto, a cristalografia não é adequada para o estudo da estrutura de carboidratos por sua dependência da estabilidade estrutural no ambiente cristalino. É necessário que o cristal possua uma repetição ordenada de estruturas para que sua difração possa ser recuperada em uma estrutura consenso. Na maioria dos casos, partes ou todo o glicano se tornam ausentes no mapa de alta resolução eletrônica [16] por conta da alta flexibilidade conformacional de carboidratos, causando uma ausência de padrão de difração recorrente. Casos onde é possível visualizar todo o glicano podem ocorrer por interações com a superfície da proteína que estabiliza o sacarídeo, criando um padrão definido.

1.2.2 RMN

A ressonância magnética nuclear (RMN) é uma das técnicas espectroscópicas mais utilizadas para a elucidação da estrutura química e conformação de compostos orgânicos em solução [17]. Ela é baseada na observação de campos magnéticos ao redor de um núcleo

atômico após excitação por um campo magnético, onde núcleos com ressonância (^1H ou ^{13}C por exemplo) absorvem a radiação eletromagnética na frequência específica de seu isótopo, gerando uma excitação. Essa excitação é captada pelo equipamento, dando acesso à estrutura eletrônica e grupos funcionais da molécula utilizada como amostra.

A RMN consegue capturar a flexibilidade de carboidratos, acessando o equilíbrio configuracional e conformacional dessas moléculas. Ela é muito utilizada para a determinação das conformações do anel em mono, oligo e polissacarídeos [18] [19] [20]. O primeiro exemplo na literatura é de Lemieux [21], observando diferentes padrões de constante de acoplamento vicinal entre prótons.

Um dos fenômenos na RMN que possibilita a análise de atributos estruturais é o acoplamento vicinal. Ele ocorre pois o *spin* de um núcleo polariza o *spin* de núcleos próximos, causando uma interferência que resulta em um sinal de RMN separado em mais de um pico. O acoplamento próton-próton vicinal ($^3J_{H,H}$) representa a separação por três ligações química, sendo essas observações utilizadas por Karplus [22] para relacionar o valor da constante com o ângulo diedral entre fragmentos H-C-C-H por meio da equação:

$$^3J_{H,H} = 7.8 - 1.0 \cos \phi + 5.6 \cos 2\phi$$

Haasnoot e Altona [23] demonstraram que os substituintes eletronegativos do fragmento H-C-C-H afetam o valor do acoplamento vicinal e propuseram uma nova equação que leva em conta a eletronegatividade desses substituintes:

$$j_o^{-1} \cdot ^3J_{H,H} = 13.86 \cos^2 \phi - 0.81 \cos \phi + \sum_{i=1}^I \Delta\chi_i [0.56 - 2.32 \cos^2 (\xi_i \phi_o^{-1} \phi + 17.9 |\Delta\chi_i|)] ,$$

onde $j_o^{-1} = 1 \text{ Hz}$, $\phi_o = 1^\circ$, ϕ é o ângulo diedral entre os hidrogênios acoplados, I o número de substituintes diferentes de hidrogênios do fragmento H-C-C-H, $\Delta\chi_i = \chi_i - \chi_H$, sendo χ_i a constante de eletronegatividade de Huggins [24] do átomo i e ξ_i é o sinal contrário do sinal do ângulo diedral do fragmento H'-C-C-S $_i$.

Por meio dessa equação, Rao [25] calculou o valor de $^3J_{H,H}$ para vários monossacarídeos em 4C_1 , mostrando concordâncias com valores experimentais. Portanto, a espectroscopia por RMN permite a visualização de estados conformacionais em solução para moléculas flexíveis como carboidratos, sem a interferência do ambiente cristalino.

1.3 Dinâmica Molecular

A dinâmica molecular (DM) inclui os métodos teóricos e computacionais utilizados para a abstração de átomos e suas interações, sendo uma metodologia poderosa para

o estudo da estrutura tridimensional de compostos e sua dinâmica [26]. Esses modelos são vitais tanto para o estudo de biomacromoléculas, como proteínas, ácidos nucleicos e carboidratos, como também para pequenas moléculas, como fármacos. Esses métodos podem acessar diferentes níveis de informação atômica, desde o grânulo mais fino como a nuvem eletrônica [27], quanto a grânulos mais grossos onde lipídeos inteiros são simplificados a poucas variáveis [28].

Sabendo dos limites experimentais para o estudo da estrutura de carboidratos devido a sua complexidade química, a DM fornece dados com resolução atômica e temporal inacessíveis a experimentos como RMN e cristalografia de raio X [29], possibilitando a exploração de diversos fenômenos biológicos [30] [31].

1.3.1 Mecânica Quântica

A representação de átomos, moléculas e suas interações de granularidade mais fina conhecida se dá por meio da mecânica quântica (QM), onde a nuvem eletrônica é representada explicitamente por meio da solução aproximada da equação de Schrödinger. Isso torna possível o estudo das propriedades relacionadas a elétrons de sistemas, como a quebra de ligações e criação de ligações, importantes para fenômenos biológicos como a catálise enzimática. Métodos computacionais como o Hartree–Fock (HF) [32] e a Teoria do Funcional da Densidade (DFT) [33] servem como aplicações práticas para a execução de mecânica quântica, observando fenômenos biológicos que necessitam de resolução eletrônica [34] [35].

Softwares como Psi4 [36] e ORCA [37] possibilitam a execução de QM com diferentes bases e métodos para diversos usos [38] [39]. Porém, a QM tem um custo computacional exorbitante, tornando a simulação de biomacromoléculas inacessíveis em escalas de tempo significativas. Se tornam necessários métodos que reduzam a complexidade computacional do problema, tornando acessível a dinâmica desses compostos em escalas de tempo que permitam o estudo de fenômenos com relevância biológica.

1.3.2 Mecânica Molecular

Como uma forma de reduzir a dimensionalidade, a mecânica molecular (MM) é utilizada para a simplificação da representação física de átomos e suas interações. Ela faz uso da mecânica clássica para descrever o movimento de moléculas pela resolução das equações de Newton [40]. Com a aproximação de Born–Oppenheimer, é possível aproximar átomos a esferas rígidas, tornando a nuvem eletrônica um elemento implícito [41]. Isso diminui a complexidade computacional em ordens de grandeza, permitindo o acesso de escalas de tempo muito superiores aos da QM para sistemas muito maiores. *Softwares* como GROMACS [42], CHARMM [43] e AMBER [44] possibilitam a simulação de diversos

sistemas biológicos por meio da mecânica molecular.

No entanto, para que a MM se aproxime da realidade, o modelo utilizado para estimar as forças entre os átomos precisa representar fielmente o dado experimental. Esse modelo é chamado de campo de força, onde parâmetros são utilizados para representação de interações entre os átomos. Esses parâmetros representam forças entre átomos não ligados e ligados, sendo representadas na figura 4.

Os primeiros dois parâmetros representados na figura 4 são os não ligados, sendo a esquerda as forças de Van der Waals e a direita a força eletrostática. Esses são os termos mais custosos, já que um átomo comumente é ligado a poucos átomos, mas interage de forma não ligada com vários.

Os termos entre átomos ligados são os quatro abaixo, sendo eles, da esquerda para a direita: ligação química entre dois átomos, ângulo de ligação entre três átomos, diedro entre quatro átomos e diedro impróprio. O parâmetro diedral, também chamado de torcional, é o parâmetro que apresenta maior relevância para definição da flexibilidade conformacional de moléculas. Dependendo de sua constante, uma molécula pode se tornar mais rígida ou mais flexível em sua dinâmica.

Para a simulação de carboidratos, nosso grupo desenvolveu o campo de força GROMOS 53A6 GLYC [45], possibilitando o estudo da dinâmica dessas biomoléculas em diversos ambientes, como ligados a lipídeos [46] e em glicoproteínas [47]. Porém, este modelo apresenta limitações, como a restrita reprodução da cinética de interconversão entre as formas que um anel monossacarídeo pode adotar [48], aumentando artificialmente a barreira energética necessária para a mudança de uma conformação de cadeira para a outra, como por exemplo de 4C_1 para 1C_4 . Outro limite do campo de força reside na falta de topologias de pentoses, carboidratos com cinco átomos no anel.

Em busca de resolver o limite de representação da cinética de interconversão do anel, esse trabalho visa a otimização dos parâmetros torcionais em busca de uma representação com maior fidelidade ao dado experimental da flexibilidade de anéis monossacarídeos. Por conta da falta de parâmetros para pentose, o trabalho também visa a geração de topologias para pentoses em busca de expandir o universo de carboidratos possíveis de serem simulados.

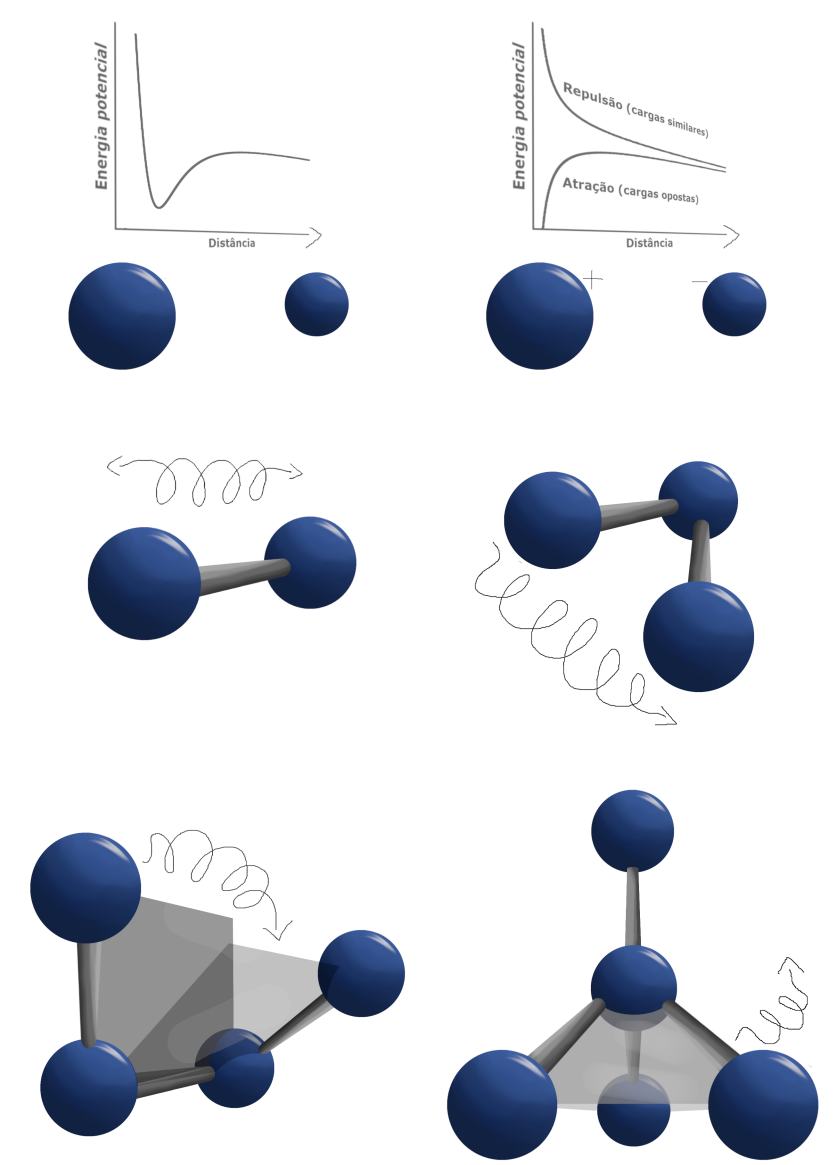


Figura 4 – Diferentes termos presentes em campos de força para a representação de átomos e suas interações. Os primeiros dois são os parâmetros não ligados, sendo o primeiro o parâmetro de Van der Waals e o segundo o parâmetro eletrostático. Os outros quatro são os parâmetros ligados, sendo estes, da esquerda para a direita: o parâmetro de ligação entre dois átomos, o parâmetro de ângulo de ligação entre três átomos, parâmetro diedral entre quatro átomos e o diedro impróprio. Laís Barth Arend, 2020 [49]

1.4 Metadinâmica

A metadinâmica é um método de simulação onde, em vez de utilizar todos os graus de liberdade como na dinâmica molecular, algumas variáveis são utilizadas para a exploração do espaço de energia livre por meio da adição de potenciais de *bias*, permitindo escapar de mínimos locais e quantificar o espaço de energia livre ao mesmo tempo [50].

Por meio desse método é possível a avaliação e estudo de eventos raros que necessitam de um longo tempo de dinâmica molecular para ocorrer, obtendo uma resolução acurada da energia relativa de ocorrência do evento em tempos hábeis [51].

O evento de interconversão entre conformações do anel de monossacarídeo é um evento raro, tendo barreiras energéticas altas que dificultam seu estudo por dinâmicas moleculares em tempos de simulação alcançados por computadores comuns. A metadinâmica se torna uma metodologia adequada para esse estudo, sendo possível identificar as energias relativas de conformações do anel utilizando tanto mecânica clássica [52] quanto mecânica quântica [53], fornecendo dados essenciais para o entendimento da dinâmica do anel e suas energias relativas.

2 Justificativa

Carboidratos são compostos de extrema importância para uma grande variedade de fenômenos biológicos, como na estrutura de ácidos nucleicos, glicoproteínas, reconhecimento célula-célula ou no desenvolvimento de doenças. Porém, problemas técnicos e experimentais relacionados ao estudo de sua estrutura, dinâmica e interação com outras moléculas dificultam o entendimento de suas funções.

Ferramentas de modelagem computacional e métodos de simulação como dinâmica molecular e metadinâmica nos permitem acessar essas informações, possibilitando o estudo das propriedades estruturais dessas biomoléculas. No entanto, para que os resultados desses modelos sejam confiáveis, seus parâmetros necessitam de validação pelo dado experimental, visando a aproximação da realidade. O campo de força GROMOS 53A6 GLYC parametrizado pelo nosso grupo apresentou dificuldades em reproduzir a cinética de monossacarídeos de hexopirranose, tornando o anel mais rígido do que o esperado com base em dados experimentais.

Portanto, com a conclusão deste trabalho, o campo de força GROMOS 53A6 GLYC para a modelagem computacional de carboidratos será otimizado, tanto na geração de novas topologias que representam corretamente a cinética de interconversão de cadeiras quanto na criação de topologias de pentoses, possibilitando o maior entendimento dessa importante classe de moléculas.

3 Objetivos

Este trabalho visa a re-parametrização do campo de força GROMOS 53A6 GLYC [45] para a simulação de hexopiranosos e pentofuranosos por meio da correção empírica de parâmetros torcionais, contando com os seguintes objetivos específicos:

- Avaliação do efeito de diferentes potenciais torcionais suavizados no equilíbrio pseudo rotacional de hexopiranosos;
- Otimização de parâmetros diedrais para diferentes monossacarídeos por correção empírica, gerando topologias mais fidedignas ao dado experimental;
- Implementação de programa para cálculo de constante de acoplamento vicinal derivada de dinâmicas moleculares;
- Geração de topologias para a dinâmica molecular de pentoses no campo de força GROMOS 53A6 GLYC.

4 Metodologia

4.1 Topologia de hexoses

O campo de força 53A6 GLYC [45] contém topologias para diversas hexoses, carboidratos com seis átomos no anel. Essas topologias foram utilizadas para a dinâmica e metadinâmica de α -Idose, β -Idose e α -Altrose, sendo somente modificados os parâmetros diedrais que representam o ângulo diedral entre 3 carbonos do anel e a hidroxila de um dos carbonos (C-C-C-O). Esses diedros são os responsáveis por conferir ou remover flexibilidade do anel monossacarídeo. Divisões por diferentes valores foram aplicadas nas constantes torcionais, estudando o seu efeito sobre a energia livre do anel.

4.2 Topologia de pentoses

O campo de força 53A6 GLYC [45] não contém topologias para pentoses, carboidratos com cinco átomos no anel. Para sua criação, topologias de hexoses foram utilizadas como base, removendo átomos e parâmetros não presentes em pentoses. Um exemplo é a desoxirribose, presente em nucleotídeos de DNA. Em comparação com uma glicose, esse composto possui um carbono a menos e um oxigênio a menos.

4.3 Dinâmica Molecular

Para a execução de dinâmicas moleculares, o protocolo de simulação foi baseado em protocolos utilizados anteriormente para estudo da dinâmica de carboidratos [54]. O software GROMACS 5.1.4 [42] foi utilizado, tendo utilizado topologias para os sistemas de hexose, explicitado em 4.1 e para topologias de pentose explicitado em 4.2. A estrutura inicial para cada sistema de monossacarídeo foi extraída do Protein Data Bank (PDB) [55].

O sistema foi solvatado com moléculas de água representadas pelo modelo SPC [56] em uma caixa dodecaédrica com 1 nm de distância às suas arestas.

O sistema solvatado foi minimizado por meio do algoritmo de *Steepest Descent* e posteriormente equilibrado por meio de NVT e NPT, cada um por 1 ns com forças de restrição de 1000 kJ/mol.

Com os sistemas equilibrados, cinco dinâmicas moleculares foram executadas, sem forças de restrição, 1 μ s cada, totalizando 5 μ s para cada divisão de parâmetros torcionais como screening inicial. Para manter cada sistema em temperatura constante (NPT), se faz necessário a aplicação de um termostato, sendo escolhido o método *V-rescale* em 298 K.

Para manter pressão constante, se faz necessário o uso de um barostato, sendo utilizado o algoritmo de Parrinello-Ramadhan. O corte de interações eletrostáticas foi feito por meio do método de *Particle-Mesh Ewald* (PME). Partindo do resultado dessas dinâmicas, foi decidido uma divisão que se aproximasse de valores experimentais de ${}^3J_{H,H}$.

Todas as simulações foram executadas utilizando o supercomputador Santos Dumont, presente no Laboratório Nacional de Computação Científica brasileiro, por meio da submissão de *jobs* utilizando o software de gerenciamento de filas *Slurm*.

4.4 Metadinâmica

Cálculos de metadinâmica foram executados utilizando GROMACS 5.1.4 junto do pacote PLUMED 2.5 [57]. As mesmas etapas de preparação do sistema utilizada para dinâmica molecular, sendo estes a geração de topologia, solvatação, minimização por *Steepest Descent* e equilibração por NVT e NPT.

Após o sistema estar solvatado, minimizado e equilibrado, a metadinâmica foi executada seguindo os mesmos algoritmos de barostato, termostato e corte de interações eletrostáticas utilizados nas dinâmicas moleculares. A altura de gaussiana utilizada foi de 0.5 kJ/mol, com largura de gaussiana de 0.1 radianos, fator de bias 10, temperatura 298 K e frequência de adição de gaussianas a cada 500 passos.

As variáveis coletivas escolhidas para a metadinâmica de hexoses foram os parâmetros θ e ϕ do *puckering*, enquanto para pentoses os parâmetros Z_x e Z_y do *puckering* foram escolhidos. O tempo de simulação total para cada metadinâmica foi de 100 ns. Essa escala de tempo se demonstrou suficiente para explorar o espaço energético relacionado ao *puckering* de um monossacarídeo, possibilitando a análise de suas conformações e energia livre.

4.5 Cálculo do *Puckering*

O cálculo das variáveis de *puckering* θ e ϕ de hexoses foram calculadas por meio de um script autoral em Python 3.7 com ajuda da ferramenta *driver* do PLUMED 2.5, possibilitando o cálculo desses ângulos durante cada passo da dinâmica molecular para monitorar as diferentes conformações do anel monossacarídico.

4.6 Cálculo de ${}^3J_{H,H}$

O cálculo de constantes de acoplamento (${}^3J_{H,H}$) durante as dinâmicas moleculares de monossacarídeos foram executados com um script autoral em Python 3.7. Cada passo da simulação é convertido em um arquivo PDB, sendo os hidrogênios apolares implícitos

adicionados por meio da ferramenta *pdb2gmx* do programa GROMACS. A equação utilizada para o cálculo de ${}^3J_{H,H}$ foi a equação de Haasnoot-Altona [23]:

$$j_o^{-1} \cdot {}^3J_{H,H} = 13.86 \cos^2 \phi - 0.81 \cos \phi + \sum_{i=1}^I \Delta\chi_i [0.56 - 2.32 \cos^2 (\xi_i \phi_o^{-1} \phi + 17.9 |\Delta\chi_i|)] ,$$

onde $j_o^{-1} = 1$ Hz, $\phi_o = 1^\circ$, ϕ é o ângulo diedral entre os hidrogênios acoplados, I o número de substituintes diferentes de hidrogênios do fragmento H-C-C-H, $\Delta\chi_i = \chi_i - \chi_H$, sendo χ_i a constante de eletronegatividade de Huggins [24] do átomo i e ξ_i é o sinal contrário do sinal do ângulo diedral do fragmento H'-C-C-S $_i$.

Com o valor de ${}^3J_{H,H}$ para cada passo (2 fs) da dinâmica molecular, a média desses valores foi utilizada como valor final. Essa média em solução se torna comparável com o valor experimental, calculado por RMN como uma média em solução da constante.

O código responsável pelo cálculo da constante de acoplamento reside nesse [repositório](#) do GitHub, sendo possível de ser usado por meio de instalação local ou por meio de uma imagem Docker.

5 Resultados

Com a informação de que o campo de força 53A6 GLYC torna os anéis de carboidratos rígidos em comparação ao dado experimental, os parâmetros diedrais relacionados a esses anéis foram identificados. Partindo dos valores originais da constante torcional, topologias novas foram criadas para hexoses, executando divisões em busca da diminuição desses valores para remediar a rigidez artificial do anel imposta pelo modelo.

Metadinâmicas e dinâmicas foram executadas, partindo desses diferentes valores em busca de entender o efeito de diminuir a constante torcional. Na metadinâmica, podemos avaliar esse efeito na superfície de energia, buscando entender se a sua alteração causa a perda da definição dos mínimos. No caso da dinâmica, a reprodução do *ensemble* em solução aquosa possibilita visualizar o efeito na dinâmica dos carboidratos.

Os resultados das diferentes simulações foram comparados a dados experimentais para definição de uma nova constante que se aproxime aos dados reais, em busca do aprimoramento do campo de força 53A6 GLYC.

5.1 Simulações de hexoses

Partindo dos parâmetros originais do campo de força 53A6 GLYC, foram criadas novas topologias para diferentes monossacarídeos, diminuindo a constante torcional. O primeiro passo foi a validação da hipótese, pois a diminuição desse valor poderia acarretar em um modelo menos acurado, perdendo a definição dos mínimos. Após ser validada, para cada monossacarídeo, foi feita uma triagem de diferentes constantes, otimizando para a proximidade ao dado experimental de RMN.

5.1.1 Validação da hipótese

Para validar a hipótese de que a diminuição da constante mantém a definição dos mínimos, a hexose α -Idose foi escolhida por possuir dados de RMN na literatura [58] e por ter a presença de três grupos hidroxilas axiais que provocam uma grande flexibilidade do anel em comparação a monômeros de Glicose e Manose [59]. Metadinâmicas de 100 ns foram executadas com o valor de constante original, dividido por dois e dividido por cinco, representados na figura 5. O resultado nos mostra que a posição dos mínimos e suas hierarquias se mantêm após a divisão, com o mínimo global na conformação 4C_1 e o próximo mínimo local em 1C_4 . A única mudança foi a diminuição da barreira energética entre os mínimos, sendo a interconversão entre cadeiras (de 4C_1 para 1C_4) aproximadamente 70

kJ/mol no parâmetro original (Fig. 4A), mas 30 kJ/mol no parâmetro dividido por 5 (Fig. 4C).

Esse resultado valida a hipótese de que a diminuição das constantes torsionais mantém a definição e a hierarquia dos mínimos em hexoses. Ao mesmo tempo, valida que a barreira de transição entre os mínimos é diminuída, reduzindo a energia necessária para a interconversão entre as cadeiras. Sabendo que a rigidez dos parâmetros originais pode ser resolvida com uma diminuição dos valores torcionais, seguimos para um *screening* de dinâmicas moleculares com diferentes parâmetros, comparando-os a resultados experimentais de RMN em busca de aprimorar os parâmetros de 53A6 GLYC para maior proximidade aos dados experimentais.

5.1.2 Screening de α -Idose

Sabendo que a diminuição da constante torcional tanto mantém a definição dos mínimos quanto diminui a barreira energética entre eles, 50 dinâmicas moleculares de 1 μ s cada foram feitas para o monossacarídeo α -Idose em diferentes valores de parâmetro, totalizando 50 μ s por parâmetro. As divisões utilizadas foram escolhidas por meio dos resultados das metadinâmicas no passo de validação da hipótese. Para cada um dos valores, as constantes ${}^3J_{H,H}$ foram calculadas para cada passo da dinâmica, sendo sua média e comparação com dados experimentais representados na tabela 1.

Os dados indicam que o escalonamento por um fator de 0,25 contém o menor erro agregado, ou seja, o valor de constante mais próximo ao dado experimental obtido por RMN. Esse resultado demonstra, junto com a análise dos resultados da metadinâmica, que o afrouxamento dos parâmetros torsionais do anel diminui as energias de transição de α -Idose, se tornando mais fidedigno ao dado experimental. Portanto, a nova topologia de α -Idose recebe um valor original escalonado por 0,25 no campo de força 53A6 GLYC.

Tabela 1 – Médias de constante de acoplamento em dinâmicas moleculares de α -Idose. exp significa valores experimentais de ${}^3J_{H,H}$ [58], enquanto calc significa valores de ${}^3J_{H,H}$ calculados durante as 50 simulações de 1 μ s. Em negrito o parâmetro com menor erro em comparação com o dado experimental.

K	${}^3J_{1,2}$		${}^3J_{2,3}$		${}^3J_{3,4}$		${}^3J_{4,5}$		$J^{exp} - J^{calc}$
	exp	calc	exp	calc	exp	calc	exp	calc	
0.22	6.0	6.4	8.1	7.6	7.9	7.5	5.0	4.3	2.0
0.25	6.0	6.8	8.1	8.0	7.9	7.9	5.0	4.6	1.3
0.29	6.0	7.0	8.1	8.3	7.9	8.2	5.0	4.8	1.7

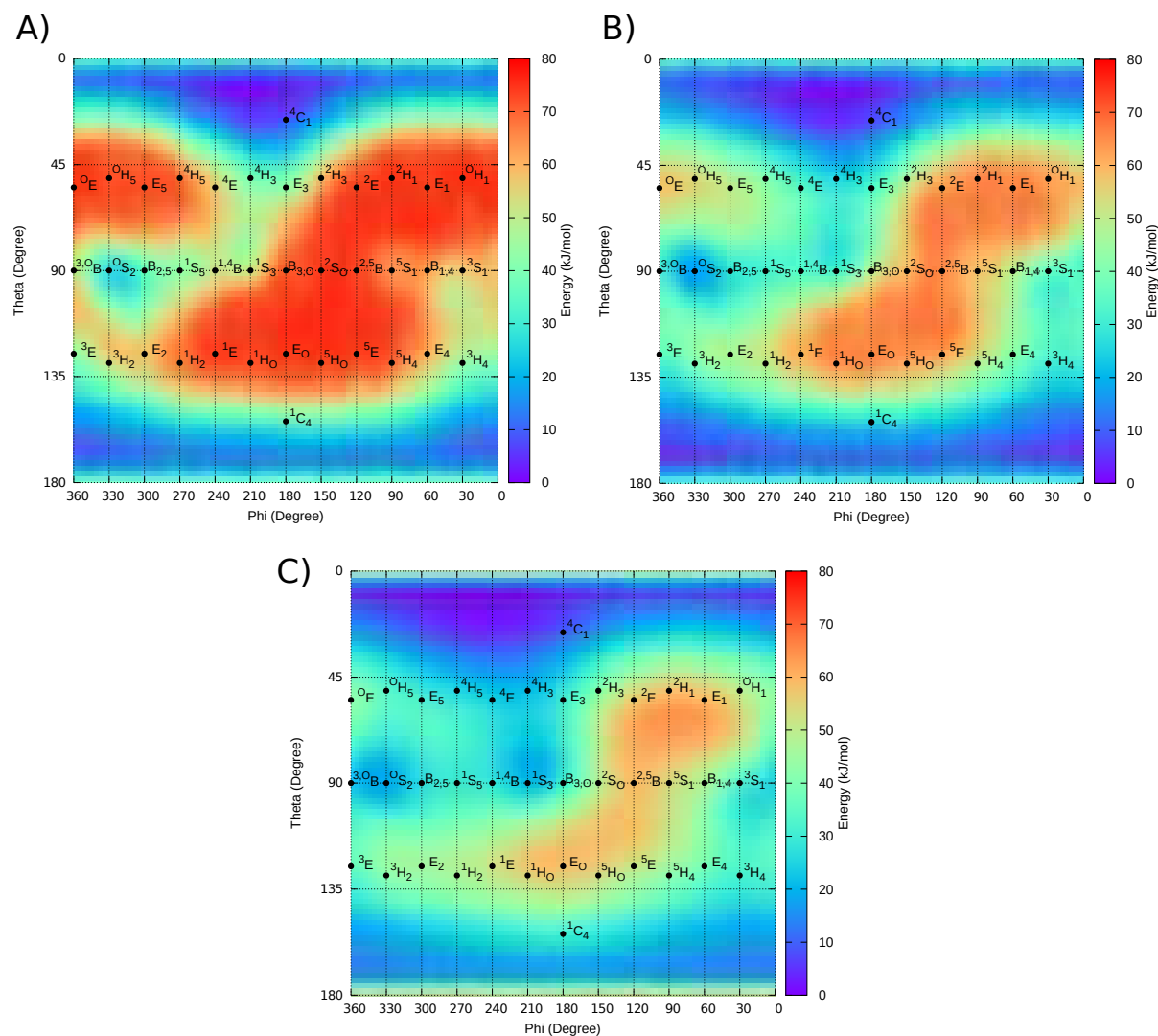


Figura 5 – Metadinâmica de α -Idose de 100 ns com diferentes valores de constante torcional, com eixo X representando o valor de θ em graus, eixo Y o valor de ϕ em graus e as cores da superfície a energia em kJ/mol. Pontos em preto representam as conformações de hexoses relacionadas ao par de θ e ϕ . A) Parâmetro original, B) Parâmetro dividido por 2 e C) Parâmetro dividido por 5. Podemos ver que a localização dos mínimos e suas hierarquias se mantêm, enquanto as barreiras energéticas diminuem.

5.1.3 Screening de β -Idose

Com o êxito dos resultados em α -Idose, a β -Idose foi estudada, buscando a otimização de seus parâmetros diedrais. Essa hexose foi escolhida pelos mesmos motivos da α -Idose, tendo dados de RMN na literatura e menor rigidez de anel [59].

Partindo de uma progressão aritmética, diferentes valores de constante foram gerados, sendo estes a divisão do parâmetro original por dois, quatro, seis, oito e dez. A tabela 2 representa as médias de constante de acoplamento para cada um desses parâmetros, provenientes de 5 dinâmicas de 1 μ s cada, totalizando 5 μ s para cada torcional.

Os dados indicam que tanto a divisão por oito como a divisão por dez tiveram o menor erro, se aproximando ao máximo dos dados experimentais. Isto sugere que valores ainda menores deveriam ser explorados, buscando exaurir o espaço amostral e alcançar a convergência com os dados experimentais.

Tabela 2 – Médias de constante de acoplamento em dinâmicas moleculares de β -Idose. exp significa valores experimentais de ${}^3J_{H,H}$ [58], enquanto calc significa valores de ${}^3J_{H,H}$ calculados durante as 5 simulações de 1 μ s. Em negrito os parâmetros com menor erro em comparação com o dado experimental.

K	${}^3J_{1,2}$		${}^3J_{2,3}$		${}^3J_{3,4}$		${}^3J_{4,5}$		$J^{exp} - J^{calc}$
	exp	calc	exp	calc	exp	calc	exp	calc	
0.1	1.6	2.5	3.8	4.0	3.7	3.6	1.8	1.4	1.6
0.13	1.6	2.5	3.8	4.1	3.7	3.7	1.8	1.4	1.6
0.17	1.6	2.5	3.8	4.2	3.7	3.8	1.8	1.5	1.7
0.25	1.6	2.5	3.8	4.3	3.7	3.9	1.8	1.5	1.9
0.5	1.6	2.6	3.8	4.7	3.7	4.4	1.8	1.6	2.9

5.1.4 Screening de α -Altrose

Na busca de otimizar os parâmetros de α -Altrose, foi executada a mesma progressão aritmética de divisão de constante diedral de dois, quatro, seis, oito e dez utilizada em β -Idose, sendo seus resultados representados na tabela 3.

Porém, em contraste com os resultados de β -Idose e α -Idose, o resultado mais próximo ao dado experimental foi o parâmetro dividido por dois. Isso indica que a diminuição em excesso da constante torcional ocasionou em um modelo que se distancia do dado experimental, acarretando em um erro agregado maior na comparação com dados de RMN. Um dos motivos para esse resultado pode partir de um aumento excessivo da flexibilidade do anel na diminuição da constante torcional, gerando um modelo menos fidedigno.

Tabela 3 – Médias de constante de acoplamento de α -Altrose. exp significa valores experimentais de ${}^3J_{H,H}$ [60], enquanto calc significa valores de ${}^3J_{H,H}$ calculados durante as 5 simulações de 1 μ s. Em negrito o parâmetro com menor erro em comparação com o dado experimental.

K	${}^3J_{1,2}$		${}^3J_{2,3}$		${}^3J_{3,4}$		${}^3J_{4,5}$		$J^{exp} - J^{calc}$
	exp	calc	exp	calc	exp	calc	exp	calc	
0.1	3.4	2.7	5.6	3.6	3.7	3.0	7.6	9.6	5.4
0.13	3.4	2.8	5.6	3.7	3.7	3.0	7.6	9.6	5.2
0.17	3.4	2.9	5.6	3.8	3.7	3.0	7.6	9.6	5.0
0.25	3.4	3.1	5.6	4.0	3.7	2.9	7.6	9.6	4.7
0.5	3.4	3.5	5.6	4.4	3.7	2.8	7.6	9.6	4.2

5.2 Simulações de pentoses

O campo de força 53A6 contém topologias de pentoses, porém suas simulações apresentam diversos limites de representação do dado experimental. Topologias completas foram criadas utilizando os parâmetros do 53A6 GLYC para hexoses, somente removendo os átomos não existentes em pentoses. Um exemplo é a Desoxirribose, presente em nucleotídeos de DNA. Em comparação com uma Glicose, esse composto possui um carbono a menos e um oxigênio a menos. Partindo dessas topologias, foram executadas metadinâmicas de 100 ns e dinâmicas moleculares de 1 μ s, visando comparação com dados experimentais para validação do modelo.

5.2.1 Desoxirribose

As energias livres recuperadas de sua metadinâmica de 100ns foram comparadas a dados de QM da literatura [61], representados na figura 6. Podemos ver as energias relacionadas a cada par de valores Z_x e Z_y do *puckering*, sendo a energia relativa dos mínimos e estado de transição representados na tabela 4.

Em comparação aos dados de mecânica quântica, podemos notar que os parâmetros diretamente derivados do GLYC 53A6 conseguiram ser próximos em relação ao estado de transição (TS), mas o mínimo global (MIN1) foi invertido com o mínimo local (MIN2).

As constantes de acoplamento calculadas na dinâmica molecular simulada por 1 μ s foram comparados com dados de RMN da literatura [62], em busca de comparar a proximidade ao dado experimental. Seus dados residem na tabela 5. Podemos notar proximidade em valores de ${}^3J_{1,2}$ e ${}^3J_{2,3}$, por exemplo, enquanto ${}^3J_{3,4}$ se mostrou muito longe do dado experimental, gerando um alto erro acumulado.

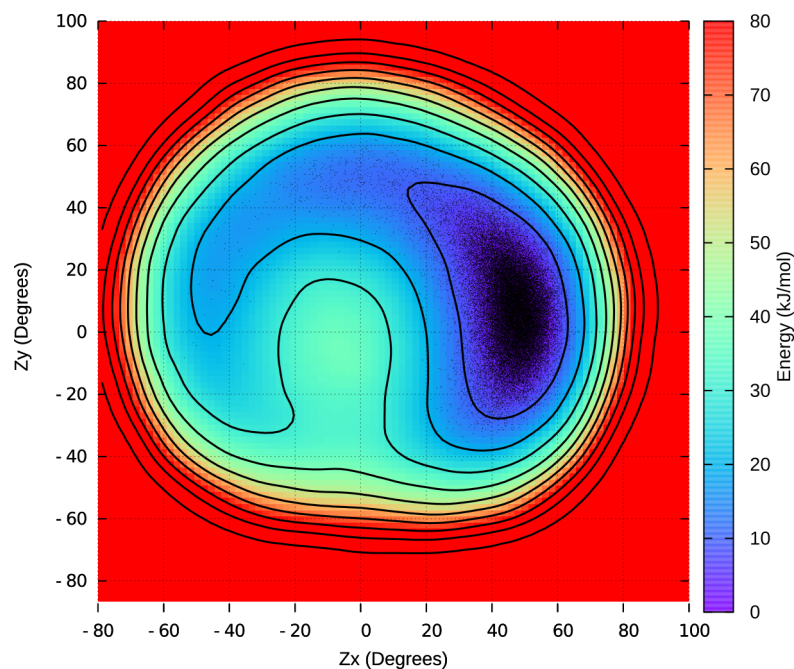


Figura 6 – Metadinâmica de Desoxirribose de 100 ns, com eixo X os valores de Z_x em graus, e no eixo Y valores em Z_y em graus. Os pontos em preto sobrepostos representam os valores Z_x e Z_y calculados em cada passo da dinâmica molecular de $1\mu s$. Cada linha na superfície demarca uma diferença de 10 kJ entre as regiões.

Tabela 4 – Posição e energia relativa dos mínimos e estado de transição em dados de QM [61] e metadinâmica com parâmetros de GLYC 53A6, sendo Z_x e Z_y em graus e ΔE em kJ/mol. Notamos uma inversão dos mínimos dos dados simulados em comparação com dados de QM.

Metodo	MIN1			MIN2			TS		
	Z_x	Z_y	ΔE	Z_x	Z_y	ΔE	Z_x	Z_y	ΔE
QM (MP2c)	-37.1	16.3	0	41.7	4.6	10.0	9.0	41.6	11.7
Metadinâmica	-41.2	18.9	13.4	51.3	0.7	0	-5.2	43.4	15.4

Tabela 5 – Médias de constante de acoplamento de Desoxirribose. exp significa valores experimentais de ${}^3J_{H,H}$ [62], enquanto calc significa valores de ${}^3J_{H,H}$ calculados durante a simulação de $1\mu s$.

Res	${}^3J_{1,2}$		${}^3J_{2,3}$		${}^3J_{3,4}$		${}^3J_{4,5}$		$J^{exp} - J^{calc}$
	exp	calc	exp	calc	exp	calc	exp	calc	
Desoxi	2.6	1.4	6.7	7.4	4.2	8.5	4.6	3.5	7.3

5.2.2 α -Arabinose

Resultados da metadinâmica de α -Arabinose não puderam ser comparados com a literatura por falta de trabalhos com esse monômero. Porém, os mínimos e locais e globais foram semelhantes aos encontrados na metadinâmica de desoxirribose, sendo representados na figura 7, tendo uma leve diferença por ter um segundo estado de transição de menor

barreira energética em $Z_x = 0$ e $Z_y = -40$.

Médias de constante de acoplamento calculadas a partir da dinâmica molecular foram comparadas com dados experimentais encontrados na literatura [63]. Com base nos dados, essa topologia se mostra mais acurada em comparação aos dados de RMN do que a topologia de desoxirribose. Uma hipótese seria a maior semelhança química entre monômeros de arabinose com monômeros de glicose e hexose. Podemos ver também que novamente os valores de ${}^3J_{3,4}$ são os valores mais afastados dos experimentais, indicando que os parâmetros relacionados aos átomos constituintes desse diedro podem ser otimizados em busca de um resultado mais aproximado ao experimental.

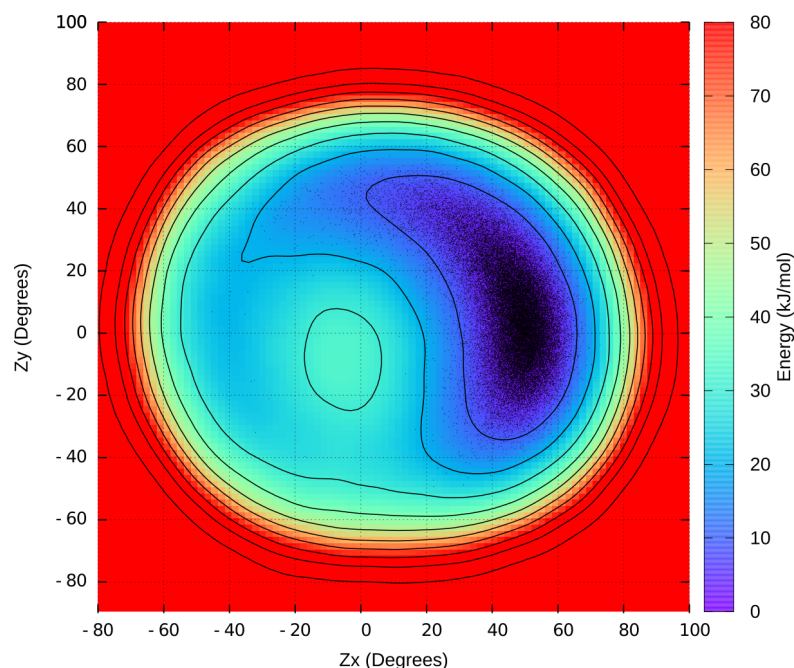


Figura 7 – Metadinâmica de α -Arabinose de 100 ns, com eixo X os valores de Z_x em graus, e no eixo Y valores em Z_y em graus. Os pontos em preto sobrepostos representam os valores Z_x e Z_y calculados em cada passo da dinâmica molecular de $1\mu s$. Cada linha na superfície demarca uma diferença de 10 kJ entre as regiões.

Tabela 6 – Médias de constante de acoplamento de α -Arabinose. exp significa valores experimentais de ${}^3J_{H,H}$ [63], enquanto calc significa valores de ${}^3J_{H,H}$ calculados durante a simulação de $1\mu s$.

Res	${}^3J_{1,2}$		${}^3J_{2,3}$		${}^3J_{3,4}$		${}^3J_{4,5}$		$J^{exp} - J^{calc}$
	exp	calc	exp	calc	exp	calc	exp	calc	
α -Ara	1.7	1.5	3.3	4.1	5.9	8.6	3.1	3.8	4.4

6 Conclusão

- A suavização dos parâmetros torcionais de hexoses mantém a definição e hierarquia de mínimos e estados de transição.
- O programa para cálculo de constante de acoplamento vicinal (${}^3J_{H,H}$) foi implementado, sendo depositado em um [repositório](#) no GitHub.
- A divisão por quatro das constantes diedrais na topologia de α -Idose resultou em um modelo mais acurado em comparação a dados experimentais de RMN.
- A divisão por oito ou por dez resultou em uma topologia mais acurada para monômeros de β -Idose em comparação a dados de RMN, sendo necessário maior tempo de simulação para decisão entre os dois valores.
- A divisão do parâmetro torcional em α -Altrose para valores de dois a dez resultou em uma topologia menos acurada em comparação a dados de RMN, tendo como perspectiva a exploração de valores entre o original e dois.
- Topologias iniciais de Desoxirribose derivadas de parâmetros de hexose foram criadas, tendo como limites a inversão dos mínimos e a baixa acurácia em comparação a dados de RMN.
- Topologias iniciais de α -Arabinose derivadas de parâmetros de hexose foram criadas, se mostrando relativamente próximas ao dado experimental de RMN, mas com espaço para melhora em átomos relacionados ao diedro de ${}^3J_{3,4}$.

7 Perspectivas

7.1 Programa para cálculo de ${}^3J_{H,H}$

Um manuscrito será preparado para a publicação do programa criado para o cálculo de constante de acoplamento derivado de dados de dinâmica molecular e dados de estruturas estáticas para a revista *Bioinformatics*.

7.2 Parâmetros de α -Altrose

Um *screening* de valores intermediários entre a divisão por dois e o valor original podem ser executados e comparados com dados de RMN, em busca de explorar a hipótese do aumento excessivo da flexibilidade do diedro.

7.3 Exploração de outros monossacarídeos

Baseado nos resultados obtidos com monômeros de Idose e α -Altrose, o efeito da divisão da constante torcional será avaliada em outras hexopiranoses como Glicose e Manose, buscando validar a universalização da hipótese e aproximação ao dado experimental de RMN.

7.4 Desoxirribose

Partindo da topologia de desoxirribose derivada de hexoses, a otimização dos seus parâmetros buscaria corrigir tanto a inversão entre mínimo global e mínimo local quanto aproximar o valor de ${}^3J_{2,3}$ ao valor experimental em RMN.

Referências

- [1] N. Sharon and H. Lis, “Lectins as cell recognition molecules 10.1126/science.2552581,” *Science*, vol. 246, no. 4927, pp. 227–234, 1989. 16
- [2] B. Blomme, C. Van Steenkiste, N. Callewaert, and H. Van Vlierberghe, “Alteration of protein glycosylation in liver diseases,” *Journal of Hepatology*, vol. 50, no. 3, pp. 592–603, 2009. 16
- [3] Y. Zhu, J. Zajicek, and A. S. Serianni, “Acyclic forms of [1-13c] aldohexoses in aqueous solution: quantitation by 13c nmr and deuterium isotope effects on tautomeric equilibria,” *The Journal of organic chemistry*, vol. 66, no. 19, pp. 6244–6251, 2001. 16
- [4] S. Herget, P. V. Toukach, R. Ranzinger, W. E. Hull, Y. A. Knirel, and C.-W. Von der Lieth, “Statistical analysis of the bacterial carbohydrate structure data base (bcsdb): characteristics and diversity of bacterial carbohydrates in comparison with mammalian glycans,” *BMC structural biology*, vol. 8, no. 1, pp. 1–20, 2008. 16
- [5] R. P. Joosten and T. Lütteke, “Carbohydrate 3d structure validation,” *Current opinion in structural biology*, vol. 44, pp. 9–17, 2017. 16
- [6] G. Jeffrey and R. Rosenstein, “Crystal-structure analysis in carbohydrate chemistry,” *Advances in carbohydrate chemistry*, vol. 19, pp. 7–22, 1964. 16
- [7] F. J. Culley, K. B. Bodman-Smith, M. A. Ferguson, A. V. Nikolaev, N. Shantilal, and J. G. Raynes, “C-reactive protein binds to phosphorylated carbohydrates,” *Glycobiology*, vol. 10, no. 1, pp. 59–65, 2000. 16
- [8] J.-Y. Le Questel, S. Cros, W. Mackie, and S. Pérez, “Computer modelling of sulfated carbohydrates: applications to carrageenans,” *International journal of biological macromolecules*, vol. 17, no. 3-4, pp. 161–175, 1995. 16
- [9] W. N. Haworth *et al.*, “Constitution of sugars,” 1929. 17
- [10] J. Schwarz, “Rules for conformation nomenclature for five- and six-membered rings in monosaccharides and their derivatives,” *Journal of the Chemical Society, Chemical Communications*, no. 14, pp. 505–508, 1973. 17
- [11] IUPAC-IUB, “Joint commission on biochemical nomenclature,” *Eur J Biochem*, vol. 257, pp. 293–298, 1998. 17
- [12] D. Cremer and J. A. Pople, “A General Definition of Ring Puckering Coordinates,” *Journal of the American Chemical Society*, vol. 97, no. 6, pp. 1354–1358, 1975. 18

- [13] C. B. Barnett, K. A. Wilkinson, and K. J. Naidoo, "Pyranose ring transition state is derived from cellobiohydrolase i induced conformational stability and glycosidic bond polarization," *Journal of the American Chemical Society*, vol. 132, no. 37, pp. 12800–12803, 2010. 18
- [14] X. Biarnes, A. Ardevol, J. Iglesias-Fernández, A. Planas, and C. Rovira, "Catalytic itinerary in 1, 3-1, 4- β -glucanase unraveled by qm/mm metadynamics. charge is not yet fully developed at the oxocarbenium ion-like transition state," *Journal of the American Chemical Society*, vol. 133, no. 50, pp. 20301–20309, 2011. 18
- [15] B. Rupp, *Biomolecular crystallography: principles, practice, and application to structural biology*. Garland Science, 2009. 19
- [16] M. R. Wormald, A. J. Petrescu, Y. L. Pao, A. Glithero, T. Elliott, and R. A. Dwek, "Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography, and molecular modelling," *Chemical Reviews*, vol. 102, no. 2, pp. 371–386, 2002. 19
- [17] H. Friebolin and J. K. Becconsall, *Basic one-and two-dimensional NMR spectroscopy*, vol. 7. Wiley-vch Weinheim:, 2005. 19
- [18] S. J. Angyal and V. Pickles, "Equilibria between pyranoses and furanoses. ii. aldoses," *Australian Journal of chemistry*, vol. 25, no. 8, pp. 1695–1710, 1972. 20
- [19] E. F. Hounsell, "1h nmr in the structural and conformational analysis of oligosaccharides and glycoconjugates," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 27, no. 5-6, pp. 445–474, 1995. 20
- [20] J. Lu and H. van Halbeek, "Complete 1h and 13c resonance assignments of a 21-amino acid glycopeptide prepared from human serum transferrin," *Carbohydrate research*, vol. 296, no. 1-4, pp. 1–21, 1996. 20
- [21] R. Lemieux, R. Kullnig, H. Bernstein, and W. Schneider, "Configurational effects on the proton magnetic resonance spectra of six-membered ring compounds1," *Journal of the American Chemical Society*, vol. 80, no. 22, pp. 6098–6105, 1958. 20
- [22] M. Karplus, "Vicinal proton coupling in nuclear magnetic resonance," *Journal of the American Chemical Society*, vol. 85, no. 18, pp. 2870–2871, 1963. 20
- [23] C. A. Haasnoot, F. A. de Leeuw, and C. Altona, "The relationship between proton-proton NMR coupling constants and substituent electronegativities-I. An empirical generalization of the karplus equation," *Tetrahedron*, vol. 36, no. 19, pp. 2783–2792, 1980. 20, 29

- [24] M. L. Huggins, “Bond energies and polarities,” *Journal of the American Chemical Society*, vol. 75, no. 17, pp. 4123–4126, 1953. 20, 29
- [25] K. Vijayalakshmi and V. Rao, “Theoretical studies on the conformation of aldopyranoses,” *Carbohydrate Research*, vol. 22, no. 2, pp. 413–424, 1972. 20
- [26] M. Karplus and G. A. Petsko, “Molecular dynamics simulations in biology,” *Nature*, vol. 347, no. 6294, pp. 631–639, 1990. 21
- [27] J. Foresman and E. Frish, “Exploring chemistry,” *Gaussian Inc., Pittsburg, USA*, 1996. 21
- [28] S. J. Marrink, A. H. De Vries, and A. E. Mark, “Coarse grained model for semi-quantitative lipid simulations,” *The Journal of Physical Chemistry B*, vol. 108, no. 2, pp. 750–760, 2004. 21
- [29] R. J. Woods, “Computational carbohydrate chemistry: what theoretical methods can tell us,” *Glycoconjugate journal*, vol. 15, no. 3, pp. 209–216, 1998. 21
- [30] H. B. Mayes, L. J. Broadbelt, and G. T. Beckham, “How sugars pucker: electronic structure calculations map the kinetic landscape of five biologically paramount monosaccharides and their implications for enzymatic catalysis,” *Journal of the American Chemical Society*, vol. 136, no. 3, pp. 1008–1022, 2014. 21
- [31] H. S. Lee, Y. Qi, and W. Im, “Effects of n-glycosylation on protein conformation and dynamics: Protein data bank analysis and molecular dynamics simulation study,” *Scientific reports*, vol. 5, no. 1, pp. 1–7, 2015. 21
- [32] C. F. Fischer, “Hartree–fock method for atoms. a numerical approach,” 1977. 21
- [33] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, “A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu,” *The Journal of chemical physics*, vol. 132, no. 15, p. 154104, 2010. 21
- [34] J. Vondrášek, L. Bendová, V. Klusák, and P. Hobza, “Unexpectedly strong energy stabilization inside the hydrophobic core of small protein rubredoxin mediated by aromatic residues: correlated ab initio quantum chemical calculations,” *Journal of the American Chemical Society*, vol. 127, no. 8, pp. 2615–2619, 2005. 21
- [35] C. B. Barnett and K. J. Naidoo, “Ring puckering: a metric for evaluating the accuracy of am1, pm3, pm3carb-1, and sec-dftb carbohydrate qm/mm simulations,” *The Journal of Physical Chemistry B*, vol. 114, no. 51, pp. 17142–17154, 2010. 21
- [36] J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, *et al.*, “Psi4: an

- open-source ab initio electronic structure program,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 2, no. 4, pp. 556–565, 2012. 21
- [37] F. Neese, F. Wennmohs, U. Becker, and C. Riplinger, “The orca quantum chemistry program package,” *The Journal of Chemical Physics*, vol. 152, no. 22, p. 224108, 2020. 21
- [38] S. E. Wheeler and J. W. Bloom, “Toward a more complete understanding of noncovalent interactions involving aromatic rings,” *The journal of physical chemistry A*, vol. 118, no. 32, pp. 6133–6147, 2014. 21
- [39] S. M. Ryno, C. Risko, and J.-L. Bredas, “Impact of molecular packing on electronic polarization in organic crystals: the case of pentacene vs tips-pentacene,” *Journal of the American Chemical Society*, vol. 136, no. 17, pp. 6421–6427, 2014. 21
- [40] W. Van Gunsteren and H. J. Berendsen, “Algorithms for macromolecular dynamics and constraint dynamics,” *Molecular Physics*, vol. 34, no. 5, pp. 1311–1327, 1977. 21
- [41] R. N. Barnett and U. Landman, “Born-oppenheimer molecular-dynamics simulations of finite systems: Structure and dynamics of (h 2 o) 2,” *Physical review B*, vol. 48, no. 4, p. 2081, 1993. 21
- [42] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, “Gromacs: fast, flexible, and free,” *Journal of computational chemistry*, vol. 26, no. 16, pp. 1701–1718, 2005. 21, 27
- [43] B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, *et al.*, “Charmm: the biomolecular simulation program,” *Journal of computational chemistry*, vol. 30, no. 10, pp. 1545–1614, 2009. 21
- [44] R. Salomon-Ferrer, D. A. Case, and R. C. Walker, “An overview of the amber biomolecular simulation package,” *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 3, no. 2, pp. 198–210, 2013. 21
- [45] L. Pol-Fachin and H. Verli, “Assessment of Glycoproteins Dynamics from Computer Simulations,” *Mini-Reviews in Organic Chemistry*, vol. 8, no. 3, pp. 229–238, 2011. 22, 26, 27
- [46] P. R. Arantes, C. Pedebos, M. D. Polêto, L. Pol-Fachin, and H. Verli, “The Lazy Life of Lipid-Linked Oligosaccharides in All Life Domains,” *Journal of Chemical Information and Modeling*, vol. 60, no. 2, pp. 631–643, 2020. 22

- [47] L. Pol-Fachin, “Insights into the effects of glycosylation and the monosaccharide-binding activity of the plant lectin CrataBL,” *Glycoconjugate Journal*, vol. 34, no. 4, pp. 515–522, 2017. 22
- [48] W. Plazinski and A. Plazinska, “Molecular dynamics simulations of hexopyranose ring distortion in different force fields,” *Pure and Applied Chemistry*, vol. 89, no. 9, pp. 1283–1294, 2017. 22
- [49] L. B. Arend 2020. 23
- [50] A. Laio and M. Parrinello, “Escaping free-energy minima,” *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 12562–12566, 2002. 24
- [51] O. Valsson, P. Tiwary, and M. Parrinello, “Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint,” *Annual review of physical chemistry*, vol. 67, pp. 159–184, 2016. 24
- [52] E. Autieri, M. Sega, F. Pederiva, and G. Guella, “Puckering free energy of pyranoses: A nmr and metadynamics-umbrella sampling investigation,” *The Journal of chemical physics*, vol. 133, no. 9, p. 09B604, 2010. 24
- [53] L. Petersen, A. Ardèvol, C. Rovira, and P. J. Reilly, “Mechanism of cellulose hydrolysis by inverting gh8 endoglucanases: a qm/mm metadynamics study,” *The Journal of Physical Chemistry B*, vol. 113, no. 20, pp. 7331–7339, 2009. 24
- [54] J. L. de Meirelles, F. C. Nepomuceno, J. Peña-García, R. R. Schmidt, H. Pérez-Sánchez, and H. Verli, “Current status of carbohydrates information in the protein data bank,” *Journal of chemical information and modeling*, vol. 60, no. 2, pp. 684–699, 2020. 27
- [55] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, “The protein data bank,” *Nucleic acids research*, vol. 28, no. 1, pp. 235–242, 2000. 27
- [56] J. Zielkiewicz, “Structural properties of water: Comparison of the spc, spce, tip4p, and tip5p models of water,” *The Journal of chemical physics*, vol. 123, no. 10, p. 104501, 2005. 27
- [57] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, “Plumed 2: New feathers for an old bird,” *Computer Physics Communications*, vol. 185, no. 2, pp. 604–613, 2014. 28
- [58] J. R. Snyder and A. S. Serianni, “D-Idose: A One- and Two-Dimensional,” *Journal of Organic Chemistry*, vol. 51, no. 12, pp. 2694–2702, 1986. 30, 31, 33

- [59] B. M. Sattelle, B. Bose-Basu, M. Tessier, R. J. Woods, A. S. Serianni, and A. Almond, "Dependence of pyranose ring puckering on anomeric configuration: methyl idopyranosides," *The Journal of Physical Chemistry B*, vol. 116, no. 22, pp. 6380–6386, 2012. 30, 33
- [60] E. Autieri, M. Sega, F. Pederiva, and G. Guella, "Puckering free energy of pyranoses: A NMR and metadynamics-umbrella sampling investigation," *Journal of Chemical Physics*, vol. 133, no. 9, pp. 1–15, 2010. 34
- [61] M. Huang, T. J. Giese, T. S. Lee, and D. M. York, "Improvement of DNA and RNA sugar pucker profiles from semiempirical quantum methods," *Journal of Chemical Theory and Computation*, vol. 10, no. 4, pp. 1538–1545, 2014. 34, 35
- [62] T. J. Church, I. Carmichael, A. S. Serianni, V. Uni, N. Dame, R. V. January, V. Re, M. Recci, and V. June, "C - 1 H and ^{13}C - ^{13}C Spin-Coupling Constants in Methyl - D -Ribofuranoside and Methyl 2-Deoxy- - D - erythro - pentofuranoside : Correlations with Molecular Structure and Conformation," vol. 7863, no. 18, pp. 8946–8964, 1997. 34, 35
- [63] A. S. Serianni and R. Barker, "[^{13}C]-Enriched Tetroses and Tetofuranosides: An Evaluation of the Relationship between NMR Parameters and Furanosyl Ring Conformation," *Journal of Organic Chemistry*, vol. 49, no. 18, pp. 3292–3300, 1984. 36