



Desenvolvimento de um campo de força “coarse-grain” para carboidratos

Victor Holanda Rusu



Universidade Federal de Pernambuco
Recife, Pernambuco, Brasil
Fevereiro, 2014



Desenvolvimento de um campo de força “coarse-grain” para carboidratos

Victor Holanda Rusu

Tese de Doutorado submetida ao
Programa de Pós-Graduação em Química
da Universidade Federal de Pernambuco
como parte dos requisitos para a obtenção
do título de **Doutor em Química**.

Orientador: Prof. Dr. Roberto Dias Lins Neto

Universidade Federal de Pernambuco
Recife, Pernambuco, Brasil
Fevereiro, 2014

Catálogo na fonte
Bibliotecário Jefferson Luiz Alves Nazareno, CRB 4-1758

Rusu, Victor Holanda.

Desenvolvimento de um campo de força “coarse-grain” para carboidratos. / Victor Holanda Rusu. – Recife: O Autor, 2014.

119f.: fig.

Orientador: Roberto Dias Lins Neto.

Tese (Doutorado) - Universidade Federal de Pernambuco. CCEN. Química fundamental , 2014.

Inclui referências e apêndice.

1. Carboidratos. 2. Campo de força 3.Coarse-grain. .
I.Lins Neto, Roberto Dias. (orientador). II. Título.

547 (22. ed.)

FQ 2014-15

Tese submetida ao Corpo Docente do Programa de Pós-Graduação em Química do Departamento de Química Fundamental da Universidade Federal de Pernambuco como parte dos requisitos necessários para a obtenção do Grau de Doutor em Química.

Aprovada:

Prof. Roberto Dias Lins Neto (Orientador)
Departamento de Química Fundamental
Universidade Federal de Pernambuco

Profa. Thereza Amélia Soares
Departamento de Química Fundamental
Universidade Federal de Pernambuco

Prof. Ricardo Luiz Longo
Departamento de Química Fundamental
Universidade Federal de Pernambuco

Prof. Pedro Geraldo Pascutti
Instituto de Biofísica Carlos Chagas Filho
Universidade Federal do Rio de Janeiro

Prof. Maurício Domingues Coutinho Neto
Centro de Ciências Naturais e Humanas
Universidade Federal do ABC

**“Desenvolvimento de um campo de força
“coarse-grain” para carboidratos”**

Por

Victor Holanda Rusu

Departamento de Química Fundamental
Centro de Ciências Exatas e da Natureza
Universidade Federal de Pernambuco
Recife-PE- Brasil

10 de Fevereiro de 2014

*A minha fabulosa e amada família e a
minha querida e muito amada Débora.*

Resumo

Victor Holanda Rusu: Desenvolvimento de um campo de força “coarse-grain” para carboidratos. Doutorado em Química, orientador prof. Dr. Roberto Dias Lins Neto, Departamento de Química Fundamental, Universidade Federal de Pernambuco, Recife, Pernambuco, Brasil, 2014.

Na natureza, os carboidratos são geralmente encontrados na forma polimérica e/ou complexados com outras biomoléculas, tais como proteínas, lipídeos, etc. A descrição teórica do comportamento destes sistemas biomoleculares requer simulações relativamente longas e conseqüentemente, computacionalmente custosas. Uma maneira de diminuir os requerimentos computacionais é através do uso de campos de força “coarse-grain” (CG). Nesta abordagem, grupos de átomos são mapeados em esferas diminuindo o número total de partículas no sistema, mas com o custo da perda de detalhes químicos. Neste trabalho desenvolvemos uma nova maneira de expressar carboidratos de forma mais aproximada e combinando com o modelo de água SPC CG GROMOS. O novo campo de força, denominado PITOMBA, corretamente mapeia as conformações barco e cadeira, bem como, os anômeros alfa e beta. A validação dos parâmetros são mostrados para as amilose V e A e α -, β - and γ -ciclodextrinas (CD). Nosso trabalho abre a possibilidade de simular sistemas contendo carboidratos e a termodinâmica de formação de complexos com CDs com um ganho de tempo computacional de 1-2 ordens de grandeza em comparação aos campo de força atomísticos.

Palavras-chaves: carboidratos, campo de força, coarse-grain

Abstract

Victor Holanda Rusu: Development of a coarse-grained force field for carbohydrates. Ph.D. thesis supervised by prof. Dr. Roberto Dias Lins Neto, Departamento de Química Fundamental, Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil, 2014.

In nature, carbohydrates are typically found in polymeric forms and/or complexed with other of biomolecules (e.g., proteins, lipids, components of external cell wall membranes, etc). A theoretical description of the behavior of these large biomolecular systems requires simulations of relatively long time scales and consequently demanding computational resources. One way to decrease the computational requirements is to use coarse-grained (CG) force fields. In this approach, groups of atoms are mapped into beads decreasing the total number of particles in a system at the expense of some chemical detail. In this work we have developed a novel approach to depict carbohydrates at a coarser level in combination with a SPC CG model for water. The new force field, denoted as PTIOMBA, correctly predicts carbohydrate boat and chair conformations as well as beta and alpha anomers. Validation of the parameters are showcased for amylose V, A and α -, β - and γ -cyclodextrins. Our work opens the possibility of simulating carbohydrate molecular systems and host-guest binding thermodynamics with convenient computational time savings of about 1-2 orders of magnitude compared to the atomistic counterparts.

Keywords: carbohydrates, force field, coarse-grain

Agradecimentos

Muitos fatores contribuíram para a realização deste trabalho, de modo que gostaria de expressar meus sinceros agradecimentos:

A minha família, especialmente a minha mãe M^a Helena, esposa Débora, irmã Mariana, tia Júlia e tio Marco pela compreensão, apoio e crença de que algum dia eu terminaria o doutorado. Aos professores Roberto Dias Lins Neto, Riccardo Baron, Thereza A. Soares, Ricardo Luiz Longo, Mozart Neves Ramos e Gustavo Seabra pela amizade, investimento, dedicação, paciência, orientações, ensinamentos, oportunidades de trabalho e inspiração. Ao professor Riccardo Baron gostaria de agradecer também por ter me recebido no seu grupo na The University of Utah, durante o sanduíche. Aos professores do Departamento de Química Fundamental por contribuírem de maneira especial para a minha formação acadêmica.

Gostaria de agradecer aos colegas de grupo no BIOMAT, LQTC e BARON GROUP e especialmente a Gabriel da Hora, Frederico Pontes, Roberta Dias, Janilson da Silva Jr., Laércio Pol Fachin, Denys Ewerton, Keila Cunha, Agrinaldo Nascimento, Daniela Nadvorny, Carlos Henrique, Carol Pacheco, Renaldo Junior, Eduardo Castro, Diego de Paula, Júlio Cosme, Albano Carneiro Neto, Miguel Ângelo, Cristiano Bastos, Karina Barros, Paulo McMiller, Erico Teixeira, James Robertson, Nadeem Vellore, Joe Passman, Jibao Lu, Julie Kneller, Jordan Meadows, Yumna Subhani, Vitor Horta, Sven Miller e Kyle Trettin por tornarem a jornada mais fácil e divertida. E a Sereina Riniker por comentários importantes e por fornecer os arquivos de entrada das simulações do modelo de água SPC CG GROMOS.

Gostaria de agradecer também à CAPES pela bolsa de doutorado através do programa Nanobiotec-BR/CAPES e à oportunidade de realizar parte do doutorado na The University of Utah. Ao CNPq, FACEPE, INCT/INAMI e LNCC pelo apoio financeiro através de infraestrutura e/ou auxílio nas viagens para congresso e cursos. Agradeço pelos recursos computacionais ao Environmental Molecular Sciences

Laboratory (EMSL), ao Argonne National Laboratory, ao Extreme Science and Engineering Discovery Environment (XSEDE), ao San Diego Supercomputer Center (SDSC) e ao Center for High Performance Computing (CHPC) na The University of Utah.

Sumário

Resumo	3
Abstract.....	4
Agradecimentos	5
Lista de Tabelas	12
Lista de Figuras	15
Notação	20
Dimensões e Unidades.....	20
Siglas.....	20
Introdução.....	22
Introdução à Carboidratos.....	23
Dinâmica Molecular	28
Integração da Equação de Movimento	29
Ensembles	29
Cálculos de Energia Livre	30
Métodos da Perturbação e Integração Termodinâmica	31
Método da Razão de Aceitação de Bennett	33
Modelos Atomísticos de Carboidratos	34
Modelos “ <i>coarse-grain</i> ” para Carboidratos	37

Modelos com Três Esferas	37
Outros modelos “coarse-grain”	43
Hipótese da Tese.....	47
Construção do Campo de Força PITOMBA	48
Características Gerais do Modelo	48
Mapeamento Atomístico para “Coarse-grain”	48
Modelo de Ligação em “Pipa”	50
Estratégia de Parametrização	50
Condições de Simulações e Sistemas Simulados	51
Parametrização dos Termos Ligados	53
Contribuições de Ligação	54
Contribuições Angulares.....	56
Contribuições de Ângulos Impróprios	57
Contribuições de Ângulos Diédricos Próprios.....	58
Parametrização dos Termos Não-Ligados.....	62
Interações Eletrostáticas	62
Interações de van der Waals	64
Desempenho Computacional do Campo de Força PITOMBA	66
Validação do Campo de Força PITOMBA	68
Efeito do Passo de Integração no Campo de Força PITOMBA.....	68

Hélices Simples e Duplas com Ligações 1-4 α de Glicose	70
α -, β - e γ -Ciclodextrinas	72
Considerações Finais	76
Perspectivas	77
Referências Bibliográficas	78
Apêndice A	85
A.1. Programas Disponíveis para <i>Download</i>	85
A.1.1. <i>g_map</i>	85
A.1.2. <i>g_cg</i> dipolefitter	86
A.2. Programas Disponíveis por Interface <i>Web</i>	87
A.2.1. Rotational Profiler	87
A.2.2. MDWiZ.....	88
Apêndice B.....	89
Modelo de Água SPC CG GROMOS.....	89
Validade do Modelo	90
Propriedades do Modelo Implementado.....	90
Apêndice C.....	92
C. 1. Potencial Quártico de Ligação	92
C. 2. Força gerada por um Potencial Quártico	92

C. 3. Potencial Quártico e Dependência com λ	93
C. 4. Implementação em C no GROMACS 4.5.x e 4.6.x.....	94
C. 4. 1. Função bonds(...)	94
C. 4. 2. Função harmonic (...)	96
Apêndice D	98
D.1. Arquivos mdp das simulações.....	98
D.1.1. Simulações Atomísticas.....	98
D.1.2. Simulações “ <i>coarse-grain</i> ”	100
D.2. Arquivos de topologia do PITOMBA	104
D.2.1. α -D-Glicose	104
D.2.2. Maltose	105
D.2.3. Celobiose.....	107
D.2.4. α -ciclodextrina.....	109
D.2.5. Água SPC CG GROMOS	115
D.2. Arquivos de geometria do PITOMBA.....	116
D.2.1. α -D-Glicose	116
D.2.2. Maltose	116
D.2.3. Celobiose.....	116
D.2.4. α -ciclodextrina.....	116

Publicações Científicas.....	118
------------------------------	-----

Lista de Tabelas

- Tabela 1.** Populações relativas das três conformações do grupo hidroximetil, obtidas durante 20 ns de simulação da α - e β -D-glicose, descrito pelo campo de força GROMOS 45A3. Tabela adaptada de Lins, R. D.; Hünenberger, P. H. *J. Comput. Chem.* **2005**, *26*, 1400–1412, Copyright (2012) com permissão da American Chemical Society..... 36
- Tabela 2.** Energias interna e livre de Gibbs relativas da α -D-glicose em relação à conformação 4C_1 . Tabela adaptada de Pol-Fachin, L.; Rusu, V. H.; Verli, H.; Lins, R. D. *J. Chem. Theory Comput.* **2012**, *8*, 4681–4690, Copyright (2012) com permissão da American Chemical Society e adaptada de Lins, R. D.; Hünenberger, P. H. *J. Comput. Chem.* **2005**, *26*, 1400–1412, Copyright (2012) com permissão da American Chemical Society..... 36
- Tabela 3.** Parâmetros de célula do cristal de celulose I β nos modelo atomísticos (CHARMM22 e CSFF), CG celulose (Modelo de Bu *et al.*) e experimental^a. Tabela adaptada de Bu, L.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. *J. Phys. Chem. B* **2009**, *113*, 10994–11002, Copyright (2009) com permissão da American Chemical Society..... 39
- Tabela 4.** Massa, em unidades de massa atômicas, e átomos, na numeração IUPAC, que constituem cada tipo de superátomo no modelo PITOMBA..... 49
- Tabela 5.** Tempo de simulação, em nanosegundos, número de partículas na caixa de simulação e representação molecular de cada um dos sistemas simulados 52
- Tabela 6.** Tipo de ligação, constante de força, distância de equilíbrio e distância médias obtidas com o PITOMBA e mapeadas do modelo atomístico GROMOS 53A6GLYC para a glicose..... 55

Tabela 7. Tipo de ângulo, constante de força, ângulo de equilíbrio, ângulos médios e posições relativa dos ângulos do PITOMBA e mapeadas do modelo atomístico GROMOS 53A6GLYC para o resíduo de maltose e celobiose. Os erros representam o desvio padrão.....	56
Tabela 8. Tipo de ângulo impróprio, constante de força, ângulo de equilíbrio, ângulos médios e posições relativa dos ângulos do PITOMBA e mapeadas do modelo atomístico GROMOS 53A6GLYC para o resíduo de maltose e celobiose.	58
Tabela 9. Tipo de ângulo próprio, constante de força, deslocamento de fase, multiplicidade para ligação glicosídica em dissacarídeos presentes no PITOMBA.....	60
Tabela 10. Valores das cargas elétricas atribuídas a cada esfera CG no PITOMBA.	62
Tabela 11. Energia de hidratação em água SPC CG calculada e escalonada de cada esfera CG do PITOMBA. Os valores experimentais estão associado as moléculas que compõe cada esfera CG, com valência completa por hidrogênio.	65
Tabela 12. Parâmetros de Lennard-Jones (C6 e C12) para cada esfera CG no PITOMBA.....	65
Tabela 13. Energia média, de 1000 passo de simulação, de cada termo do PITOMBA em função do passo de integração para diferentes simulações da maltose em solução.	69
Tabela 14. Número médio de moléculas de água dentro das cavidades das α -, β - e γ -ciclodextrinas durante 100 ns de simulação com os campos de força GROMOS 53A6 _{GLYC} e PITOMBA.	73
Tabela 15. Parâmetros requeridos pelo modelo de água SPC CG GROMOS.	90

Tabela 16. Densidade (ρ), energia potencial por esfera CG (V_{pot}) e coeficiente de difusão (D), em 298 K, para modelo SPC CG GROMOS implementado nesta tese, em comparação com diferentes modelos de água CG e com dados experimentais. Tabela adaptada de Riniker, S.; van Gunsteren, W. F. *J. Chem. Phys.* **2011**, *134*, 084110, Copyright (2011) com permissão da AIP Publishing LLC..... 91

Lista de Figuras

Figura 1. Representação esquemática da D-glicose.....	23
Figura 2. Representação esquemática da α -D-glicose e β -D-glicose.....	24
Figura 3. Representação esquemática da maltose.	25
Figura 4. Representação esquemática da celobiose.....	25
Figura 5. Representação esquemática dos ângulos ϕ e ψ da ligação glicosídica.....	26
Figura 6. Representação esquemática dos tipos de ciclodextrinas em função do número de monômeros do anel.....	27
Figura 7. Representação dos valores de ϕ e ψ para a maltose (a) e celobiose (b) com os valores médios de cada diedro no centro de cada elipse. A magnitude dos eixos das elipses representam os desvios padrão de cada modelo. Figura adaptada de Carbohydrate Research, 344/16, Carlos A. Stortz, Glenn P. Johnson, Alfred D. French, Gábor I. Csonka, 12, Copyright (2009), com permissão da Elsevier. Adaptada de The Lancet, 16, Carlos A. Stortz, Glenn P. Johnson, Alfred D. French, Gábor I. Csonka, Comparison of different force fields for the study of disaccharides, 12, Copyright (2009), com permissão da Elsevier.....	35
Figura 8. Representações atomísticas e M3B sobrepostas da molécula de maltose. A posição das esferas CG M3B (B1, B4 e B6) coincidem com os átomos de carbono (C1, C4 e C6) do modelo atomístico. As três esferas do modelo M3B estão conectadas por ligações (linhas tracejadas) e uma única ligação entre resíduos representa a ligação glicosídica. Figura adaptada de Molinero, V.; Goddard, W. A. <i>J. Phys. Chem. B</i> 2004 , 108, 1414–1427, Copyright (2004) com permissão da American Chemical Society.....	37

- Figura 9.** Determinação da temperatura de transição vítrea para a α -D-glicose com 12,2 % de água na representação M3B. Figura adaptada de Molinero, V.; Goddard, W. A. *J. Phys. Chem. B* **2004**, *108*, 1414–1427, Copyright (2004) com permissão da American Chemical Society. 38
- Figura 10.** Açúcares parametrizados no modelo MARTINI por López *et al.*: α -glicose(G); β -frutose (F); sucrose (SUC); trealose (T); maltose (M); celobiose (C); nigeriose (N); laminarabiose (L); kojibiose (K) e soforose (S). Figura adaptada de López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210, Copyright (2009) com permissão da American Chemical Society. 40
- Figura 11.** Mapeamento atômico para CG de mono (A); di- (B) e oligossacarídeos (C) no modelo MARTINI. Figura adaptada de López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210, Copyright (2009) com permissão da American Chemical Society. 41
- Figura 12.** Estruturas da amilose V com 26 monômeros em nonano (A) e em água (B). Na parte superior representação atômica e na inferior CG MARTINI. (A) Ambas as representações em nonano convergiram para um estrutura em V estável em 25 ns; (B) A transferência para água promove um aumento de flexibilidade e o desenovelamento das estruturas em hélice. Figura adaptada de López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210, Copyright (2009) com permissão da American Chemical Society. 42
- Figura 13.** Estrutura atômica com as definições das esfera CG para o modelo de Hynninen *et al.* para a (a) glicose e água; (b) celobiose e (c) celotrealose. A representação CG da celotetraose está indicada pela seta. Figura adaptada de Hynninen, A.-P.; Matthews, J. F.; Beckham, G. T.; Crowley, M. F.; Nimlos,

M. R. *J. Chem. Theory Comput.* **2011**, 7, 2137–2150, Copyright (2011) com permissão da American Chemical Society. 42

Figura 14. Estruturas finais da simulação da transição entre as fases cristalina ($\lambda = 1$) e amorfa ($\lambda = 0$) da celulose. Figura adaptada de Srinivas, G.; Cheng, X.; Smith, J. C. *J. Chem. Theory Comput.* **2011**, 7, 2539–2548, Copyright (2011) com permissão da American Chemical Society. 43

Figura 15. Representação esquemática da cadeia de celulose na representação atomística (topo) e na representação CG de Bellesia (meio). Na parte inferior a comparação geométrica dos modelos CG de Bellesia e atomístico para a molécula de glicose. Figura adaptada de Bellesia, G.; Chundawat, S. P. S.; Langan, P.; Redondo, A.; Dale, B. E.; Gnanakaran, S. *J. Phys. Chem. B* **2012**, 116, 8031–8037, Copyright (2012) com permissão da American Chemical Society. 44

Figura 16. Estruturas finais da simulação da transição entre as fases I β ($\lambda = 0$) e III $_T$ ($\lambda = 1$) da celulose. Figura adaptada de Bellesia, G.; Chundawat, S. P. S.; Langan, P.; Redondo, A.; Dale, B. E.; Gnanakaran, S. *J. Phys. Chem. B* **2012**, 116, 8031–8037, Copyright (2012) com permissão da American Chemical Society. 45

Figura 17. Mapeamento CG para o resíduo de glicose nos quatro diferentes modelos CG desenvolvidos por Markutsya. Figura adaptada de Markutsya, S.; Devarajan, A.; Baluyut, J. Y.; Windus, T. L.; Gordon, M. S.; Lamm, M. H. *J. Chem. Phys.* **2013**, 138, 214108–214117, Copyright (2013) com permissão da AIP Publishing LLC. 46

Figura 18. Representação do resíduo de glicose no campo de força PITOMBA e os átomos que constituem cada esfera CG na representação de átomos unidos do GROMOS96. 49

- Figura 19.** Representação esquemática da maltose, no modelo PITOMBA, enfatizando as ligações do modelo em forma de pipa..... 50
- Figura 20.** Histogramas normalizados por área das distâncias CV2–CV6 e CVM–CV6 mapeadas de 100 ns uma simulação atomística da glicose com o campo de força 53A6_{GLYC}. 55
- Figura 21.** Perfis rotacionais dos ângulos ϕ (GROMOS 53A6_{GLYC}) e χ (CG PITOMBA) para a molécula de maltose (a) e celobiose (b). O perfil atomístico foi calculado com o nível de teoria HF/6-31G(d). Os ângulos estão definidos na representação CG..... 60
- Figura 22.** Valores do ângulo diédrico da ligação glicosídica da a) maltose e b) celobiose, no modelo PITOMBA, durante 100 ns de simulação CG PITOMBA e atomísticas GROMOS 53A6_{GLYC}. As moléculas atomísticas foram mapeadas para CG..... 61
- Figura 23.** a) Momento de dipolo do resíduo de glicose nos modelos PITOMBA e GROMOS 53A6_{GLYC} em função do tempo ao longo de 100 ns de simulação; b) Histograma do ângulo que entre o momento de dipolo do resíduo de glicose e o seu valor médio nos modelos PITOMBA e GROMOS 53A6_{GLYC} ao longo de 100 ns de simulação..... 63
- Figura 24.** Momento de dipolo do resíduo de glicose nos modelos GROMOS 53A6_{GLYC} (vermelho) e PITOMBA (azul). A posição do vetor atomístico foi deslocado para melhor visualização do vetor CG, mas ambos são sobrepostos. 64
- Figura 25.** Ganho computacional do campo de força PITOMBA comparado ao campo de força GROMOS 53A6_{GLYC}, para as moléculas de glicose, γ -ciclodextrina e amilose A, respectivamente. Os cálculos foram realizado num computador com arquitetura Intel i7, com quatro núcleos de processamento 2,9 GHz e 8 GB de memória RAM e sistema operacional OS X 10.9.1..... 66

- Figura 26.** Valores do ângulo diédrico da ligação glicosídica da maltose, no modelo PITOMBA, durante 500 passos de simulação no CG PITOMBA. 70
- Figura 27.** Representação da amilose V em um passo de um trajetória de 100 ns em solução no PITOMBA. A distância entre dois resíduos numa volta da amilose V é $\sim 20 \text{ \AA}$ 71
- Figura 28.** Raio de giro da amilose V ao longo de 100 ns de simulação no campo de força PITOMBA em solução. Valor médio de $11,9 \text{ \AA}$ 71
- Figura 29.** Distância entre resíduos numa volta de amilose A ao longo de 100 ns de simulação no campo de força PITOMBA em solução. Valor médio de $24,1 \text{ \AA}$.72
- Figura 30.** Representação esquemática da γ -ciclodextrina no campo de força GROMOS 53A6_{GLYC} e visões lateral (a) e de cima (b) do cilindro (azul) utilizado na contagem das moléculas de água..... 74
- Figura 31.** Histograma do número de moléculas de água dentro das cavidades atomísticas no modelo GROMOS 53A6_{GLYC} da (a) α -, (c) β - e (e) γ -CDs, e do número de moléculas de água atomísticas dentro das cavidades CG PITOMBA do (b) α -, (d) β - e (f) γ -CDs. O número de moléculas de água atomísticas no modelo CG foi obtido ao multiplicar por cinco o número de esferas dentro das cavidades. 75
- Figura 32.** Modelo de água SPC CG GROMOS. Figura adaptada de Riniker, S.; van Gunsteren, W. F. *J. Chem. Phys.* **2011**, *134*, 084110, Copyright (2011) com permissão da AIP Publishing LLC. 89

Notação

Dimensões e Unidades

Utilizou-se de diversas unidades para reportar os resultados. Tentou-se manter as unidades o mais próximas possíveis das usuais.

Siglas

CG – do inglês *coarse-grain*

CD – Ciclodextrina

PITOMBA – Parameter Interface To Oligosaccharide Molecules Based on Atoms

BAR – Método da Razão de Aceitação

1→1 α – Ligação glicosídica da α -D-glicopiranosil-(1→1)- α -D-glicopiranosose

1→1 β – Ligação glicosídica de β -D-glicopiranosil-(1→1)- β -D-glicopiranosose

1→2 α – Ligação glicosídica da α -D-glicopiranosil-(1→2)- α -D-glicopiranosose

1→2 β – Ligação glicosídica de β -D-glicopiranosil-(1→2)- β -D-glicopiranosose

1→3 α – Ligação glicosídica da α -D-glicopiranosil-(1→3)- α -D-glicopiranosose

1→3 β – Ligação glicosídica de β -D-glicopiranosil-(1→3)- β -D-glicopiranosose

1→4 α – Ligação glicosídica da α -D-glicopiranosil-(1→4)- α -D-glicopiranosose

1→4 β – Ligação glicosídica de β -D-glicopiranosil-(1→4)- β -D-glicopiranosose

1→6 α – Ligação glicosídica da α -D-glicopiranosil-(1→6)- α -D-glicopiranosose

1→6 β – Ligação glicosídica de β -D-glicopiranosil-(1→6)- β -D-glicopiranosose

Introdução

Carboidratos têm grande importância na natureza, principalmente na forma de longas estruturas poliméricas tais como amilose, glicogênio (reserva energética de animais), celulose, quitina (componentes estruturais de células e exoesqueletos) e combinações com outras biomoléculas tais como proteínas (na forma de glicoproteínas) e lipídios (lipopolissacarídeos).¹⁻⁵ O tamanho e flexibilidade destes sistemas, em geral, limitam a simulação computacional via modelos atomísticos. Uma alternativa é o uso de representações mais simplificadas, tais como as “*coarse-grain*” (CG). Nesta abordagem, grupos de átomos são mapeados em uma única esfera, sendo estas responsáveis por mimetizar o volume ocupado dos grupos de átomos que as compõe e as energias de interação com as demais partículas do sistema. Há duas grandes razões, do ponto de vista computacional, para o uso desta representação, i. há uma redução no número de partículas no sistema e ii. os movimentos mais lentos devido às massas reduzidas maiores, assim o uso de passos maiores na integração das equações de movimento. No entanto, o uso de modelos CG representa um compromisso entre tempo computacional e perda de informação química, tais como, detalhes estruturais do sistema, interações específicas entre sítios de ligação e superfícies termodinâmicas simplificadas.

Neste contexto, este capítulo tem como objetivos dar uma breve introdução aos tipos de açúcares tratados nesta tese e à simulação computacional via dinâmica molecular de carboidratos tanto nas representações atomísticas, bem como nas representações CG, à teoria do cálculo de energias de livres de Gibbs em simulação computacionais, utilizada para ajustar os potenciais de interação entre as esferas CG e, por fim, apresentar a hipótese geral desta tese.

Introdução à Carboidratos

Carboidratos é a classe de biomoléculas mais abundantes na natureza e podem ser classificadas como: monossacarídeos (unidades de polihidroxiáldeídos ou polihidroxicetonas), dissacarídeos (dímeros de monossacarídeos), oligossacarídeos (cadeias curtas contendo no máximo 20 monossacarídeos) e polissacarídeos (cadeias curtas contendo mais de 20 monossacarídeos). O monossacarídeo mais abundante na natureza é a dextrose (**Figura 1**) que é formada por um anel de seis membros e contém cinco grupos hidroxila.⁶

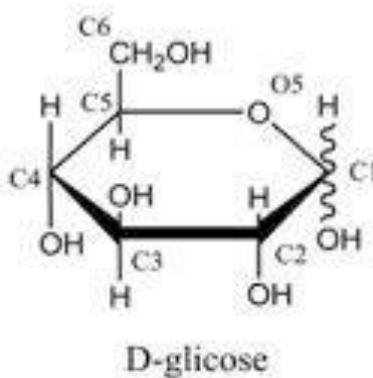


Figura 1. Representação esquemática da D-glicose.

Os carboidratos com mais de cinco átomos de carbono geralmente aparecem na forma de anéis cíclicos, parecidos com o furano (cinco membros) ou pirano (seis membros). Desta forma os carboidratos com seis membros são conhecidos como piranoses. Carboidratos com grupo aldeídos são conhecidos como aldoses e os com grupo cetona são conhecidos como cetoses. As aldoses com seis átomos de carbono tendem a formar anéis de seis membros (a exemplo da glicose), em que um dos membros do anel é um oxigênio (**Figura 1**).⁶

Cada átomo de carbono da cadeia de glicose apresenta uma numeração específica, como pode ser vista na **Figura 1**. Exceto pelo oxigênio que pertence ao anel (oxigênio cinco), os demais oxigênios se ligam a apenas um átomo de carbono e têm a numeração igual ao do carbono associado a ele. Por exemplo, o oxigênio ligado ao carbono C1 tem numeração O1. Esta mesma numeração é empregada para os átomos de

hidrogênio ligados aos carbonos. Quando há mais de dois hidrogênios ligados ao mesmo carbono um segundo índice é adicionado para indicar cada hidrogênio separadamente. Os átomos de hidrogênio ligados aos átomos de oxigênio são denominados HOs, assim o hidrogênio ligado ao O1 tem nomenclatura HO1.⁶

Na representação esquemática da dextrose na **Figura 1** observa-se a presença de um carbono anomérico (C1) que pode apresentar duas configurações do grupo OH, a saber, α ou β . As duas configurações distintas da D-glicose estão representadas na **Figura 2**. A glicose recebe o nome de α -glicose quando o grupo OH do carbono anomérico está em posição oposta ao grupo CH₂OH e recebe o nome de β -glicose quando o grupo OH do carbono anomérico está em posição similar ao grupo CH₂OH.⁶

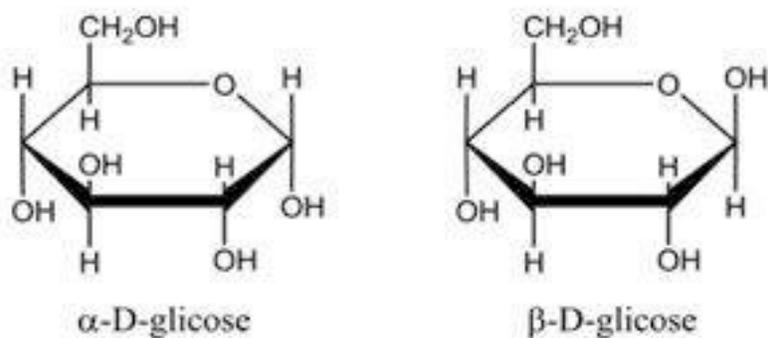


Figura 2. Representação esquemática da α -D-glicose e β -D-glicose

A ligação química entre dois monossacarídeos é denominada ligação glicosídica e ocorre entre qualquer posição que contenha o grupo OH no açúcar. Quando a ligação envolve o carbono 1 de um monossacarídeo e o carbono 4 de outro, e ainda, os monossacarídeos são α -glicose, forma-se um composto conhecido como α -D-glicopiranosil-(1 \rightarrow 4)- α -D-glicopiranosose ou maltose (**Figura 3**).⁶

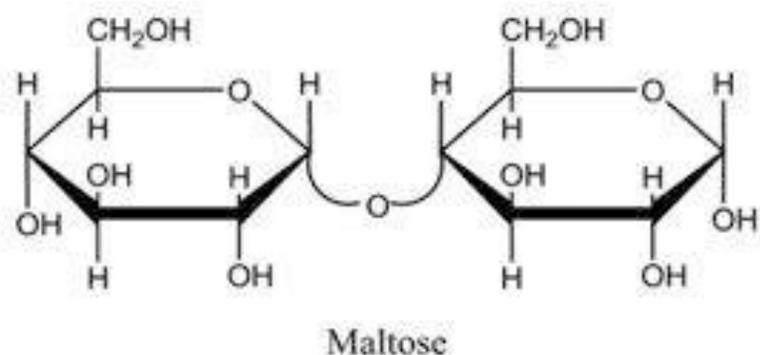


Figura 3. Representação esquemática da maltose.

A escrita da nomenclatura IUPAC para carboidratos é muito extensa mesmo para dissacarídeos.⁶ Portanto, nesta tese adotaremos uma nomenclatura simplificada. Os compostos do tipo z-D-glicopiranosil-(X→Y)-z-D-glicopiranosose serão abreviados como X→Y z, em que z pode ser α ou β .

A maltose é menor unidade que pode conter ligações 1→4 α , o que significa dizer que ela é um excelente ponto de partida para entender o comportamento de cadeias mais longas contendo ligações 1→4 α . Quando a ligação glicosídica acontece nas posições 1 e 4 entre β -glicoses, o dissacarídeo é denominado celobiose (**Figura 4**). A maltose é obtida a partir da quebra enzimática do amido e a celobiose da degradação enzimática da celulose. Estes compostos apresentam solubilidade em água bastante distintas sendo 1,08 g mL⁻¹ e 0,12 g mL⁻¹ para a maltose e celobiose, respectivamente.

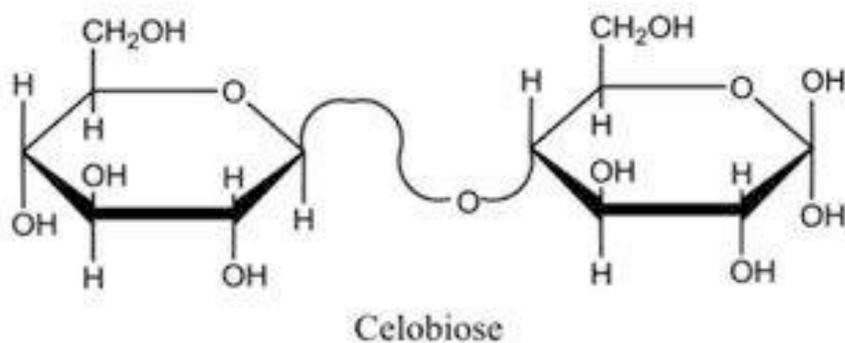


Figura 4. Representação esquemática da celobiose

A ligação glicosídica nas posições 1 e 4 está associada a dois novos diedros (ϕ e ψ) que definem a disposição espacial entre os monômeros no dissacarídeo. De acordo com a IUPAC⁶ os ângulos diedros ϕ e ψ (**Figura 5**) são definidos como:

$$\begin{aligned}\phi &= H1 - C1 - O4' - C4' \\ \psi &= C1 - O4' - C4' - H4'\end{aligned}\quad (1)$$

Cada tipo de dissacarídeo tem um perfil rotacional diferente e conseqüentemente tem valores médios para cada ângulo individualmente. A maltose, por exemplo, apresenta valores de ϕ que variam entre -90° e 60° e valores de ψ entre -100° e 60° . Mas os valores angulares da geometria de menor energia são em torno de -20° para ϕ e -20° para ψ . Na celobiose, porém os valores mudam para uma faixa de -50° a 90° para ϕ e de -80° até 40° para ψ . Com valores de mínimo de energia em torno de 50° para ϕ e de 0° para ψ .

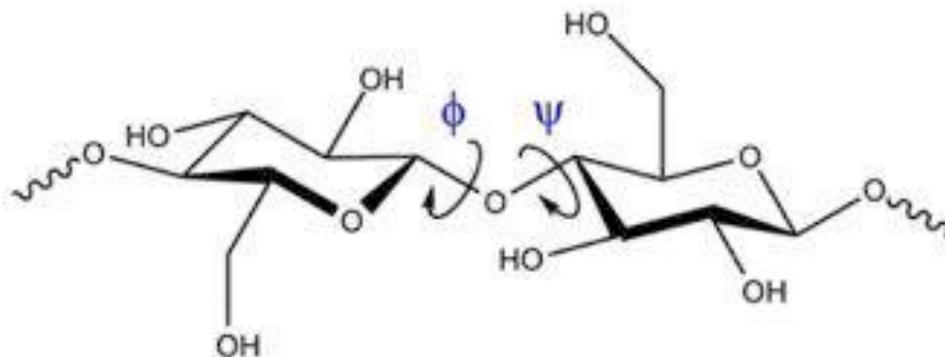


Figura 5. Representação esquemática dos ângulos ϕ e ψ da ligação glicosídica

A partir da maltose, pode-se construir cadeias maiores de carboidratos contendo ligações $1 \rightarrow 4 \alpha$. Exemplos típicos são amilose V e amiloses A e B. As amiloses A e B formam hélices duplas, cada uma contém seis glicoses por volta, que diferem apenas no arranjo das hélices e no número de águas de coordenação.⁷ Devido à flexibilidade intrínseca dos polissacarídeos e as atuais limitações nas técnicas experimentais de detecção de estrutura, ainda existe controvérsia a respeito da direção de rotação das hélices, podendo ser rotação de mão direita ou de mão esquerda.⁸⁻¹³ Similarmente às amiloses A e B, a amilose V contém seis glicoses por volta, no entanto, esta última é

formada por uma única fita de polissacarídeos. Outra característica bem diferente entre as amiloses citadas é o grau de empacotamento das fitas. A distância entre dois açúcares numa volta completa de hélice, nas amiloses A e B está em torno de 20 Å, enquanto que para a amilose V a distância é em torno de 8 Å.¹⁴⁻¹⁶

Do ponto de vista da ocorrência destes compostos a amilose A é, em geral, encontrada em cereais, enquanto que a amilose B é encontrada em túberas, tais como batatas e inhame. A amilose V, em geral, é obtida a partir das amiloses A e B quando diminui-se a polaridade do solvente, ou quando adiciona-se sais de cristalização. Com a celobiose é possível construir polímeros de açúcares 1→4 β com grande importância energética como por exemplo, a celulose. A celulose pode ser encontrada em diversas formas, sendo as mais comuns a celulose I e II. A diferença entre elas está no arranjo das ligações de hidrogênio entre as fitas.

As cadeias de α-D-glicose com ligações glicosídicas do tipo 1→4 α podem também apresentar-se na forma de estruturas cíclicas chamadas de ciclodextrinas (CD). As compostas por seis, sete e oito monômeros são denominadas α-, β- e γ-ciclodextrinas, respectivamente (**Figura 6**).¹⁷

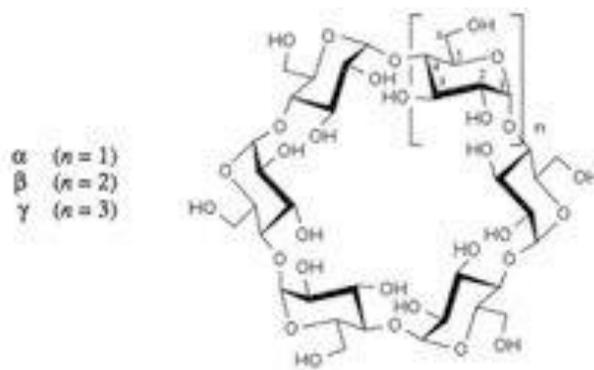


Figura 6. Representação esquemática dos tipos de ciclodextrinas em função do número de monômeros do anel.

As CDs são bastante úteis para solubilizar compostos hidrofóbicos e têm grandes aplicações na indústria farmacêutica na entrega de drogas e terapia gênica.¹⁸⁻²⁰

Dinâmica Molecular

Dinâmica molecular é uma técnica de amostragem computacional em que a evolução temporal de um grupo de partículas é simulado através da integração das equações de movimento destas partículas. Quando as equações de movimento são governadas pelas leis de Newton e as interações interpartículas são calculadas por um potencial clássico, denomina-se Dinâmica Molecular Clássica, ou simplesmente Dinâmica Molecular.²¹

O potencial clássico é uma soma de várias contribuições, podendo ser classificadas em dois grupos, i. contribuições associadas a termos ligados e ii. contribuições associadas a termos não ligados. A quantidade de termos e os parâmetros de cada termo depende do campo de força e do sistema em estudado. O campo de força GROMOS^{22,23}, por exemplo, tem suas contribuições divididas como mostra a **Equação 2**.

$$V = \sum_l V_l^{ligação} + \sum_a V_a^{\hat{angulo}} + \sum_i V_i^{impr\u00f3prios} + \sum_d V_d^{di\u00e9dros} + \sum_j V_j^{lennard-jones} + \sum_c V_c^{coulomb}, \quad (2)$$

em que, l representa o n\u00famero de termos de liga\u00e7\u00e3o $V_l^{ligação}$, a representa o n\u00famero de termos angulares $V_a^{\hat{angulo}}$, i representa o n\u00famero de termos de diedros impr\u00f3prios $V_i^{impr\u00f3prios}$, d representa o n\u00famero de termos de \u00e2ngulos di\u00e9dricos $V_d^{di\u00e9dros}$, j representa o n\u00famero de termos de intera\u00e7\u00e3o de Lennard-Jones $V_j^{lennard-jones}$, c representa o n\u00famero de termos de intera\u00e7\u00e3o el\u00e9tricas de Coulomb $V_c^{coulomb}$.

As for\u00e7as s\u00e3o obtidas atrav\u00e9s do negativo do gradiente do potencial, como mostra a **Equa\u00e7\u00e3o 3**

$$\vec{F} = -\vec{\nabla}V. \quad (3)$$

A partir das forças o movimento dos átomos é simulado numericamente ao resolver-se a equação de movimento de Newton.

$$\frac{d^2\vec{r}_i}{dt^2} = \frac{\vec{F}_i}{m_i}, \quad (4)$$

em que, \vec{r} representa a posição, m a massa, \vec{F} a força sobre a i -ésima partícula do sistema e t representa o tempo.

Integração da Equação de Movimento

Existem diversos algoritmos para resolver numericamente a **Equação 4**, tais como, *verlet*²⁴ e *velocity-verlet*.²⁵ Dentre os mais comumente implementados encontra-se o algoritmo de salto do sapo, do inglês “*leapfrog*”. O algoritmo recebe este nome pois as posições (\vec{r}) são conhecidas no tempo \vec{t} , mas as velocidades (\vec{v}) são conhecidas apenas em $t - 1/2 dt$.²¹

As posições e velocidades são obtidas através das relações

$$\begin{aligned} \vec{v}\left(t + \frac{1}{2}dt\right) &= \vec{v}\left(t - \frac{1}{2}dt\right) + \vec{F}(t) \frac{dt}{m} \\ \vec{r}(t + dt) &= \vec{r}(t) + \vec{v}\left(t + \frac{1}{2}dt\right)dt \end{aligned} \quad (5)$$

Ensembles

Para resolver a **Equação 4** é necessário algumas condições iniciais, tais como, posições (\vec{r}) e velocidades (\vec{v}) iniciais. As posições são definidas pelo usuário e as velocidades podem ser obtidas, por exemplo, a partir de uma distribuição de Maxwell-Boltzmann de velocidades, na temperatura da simulação. O número de partículas e o volume da caixa de simulação são definidos pelas condições iniciais e a energia total é a soma da energia potencial inicial e da energia cinética gerada pela distribuição de velocidades. Caso, estas variáveis sejam mantidas constantes ao longo da trajetória diz-se

que o conjunto (*ensemble*) de configurações pertencem a um *ensemble* micro-canônico ou N,V,E. E quando controla-se a temperatura do sistema ao longo da simulação, denomina-se *ensemble* canônico ou N,V,T. Em química, em geral, as reações ocorrem a pressão (p) e temperatura constantes, neste caso, quando há o controle da temperatura e pressão, denomina-se *ensemble* isotérmico-isobárico ou N,p,T.

A temperatura é calculada como

$$T = \frac{1}{n_f n k_b} \sum_i^n \frac{\overline{\vec{p}_i \cdot \vec{p}_i}}{m_i}, \quad (6)$$

em que k_B é a constante de Boltzmann, n_f é a quantidade de graus de liberdade de cada partícula, n é o número de partículas, \vec{p}_i é o momento e m_i é a massa da i -ésima partícula no sistema. O controle da temperatura do sistema pode, por exemplo, ser feito alterando de forma apropriada o valor do momento das partículas durante a trajetória.

De acordo com o teorema virial a pressão sobre um sistema microscópico tem forma²¹

$$pV = \frac{1}{3} \sum_i^n \frac{\vec{p}_i^2}{m_i} + \frac{1}{3} \sum_i^n \vec{r}_i \cdot \vec{F}, \quad (7)$$

em que \vec{F} é a força gerada pela interação entre as partículas. Desta forma, para controlar a pressão sobre o sistema basta calcular o lado direito da **Equação 7** e alterar o volume da caixa de simulação para que alterar a pressão.

Cálculos de Energia Livre

Diferentes propriedades moleculares podem ser expressas a partir de médias de *ensembles*, tais como energia cinética e potencial. No entanto, algumas propriedades termodinâmicas dependem não apenas de um único *ensemble*. Exemplos destas propriedades são entropia e energia livre. O espaço de fase acessível ao um sistema

molecular com muitos graus de liberdade não pode ser explorado completamente numa simulação. No entanto, as derivadas da energia livre em relação ao volume, temperatura ou ao potencial que descreve as interações das partículas podem ser expressos como médias de *ensembles*, e consequentemente medidos a partir de simulações computacionais.²¹

Para um *ensemble* isotérmico-isobárico (N,p,T) a energia livre de Gibbs pode ser expressa na forma

$$G = -k_B T \ln Z, \quad (8)$$

em que k_B é a constante de Boltzmann, T é a temperatura e Z é a função de partição, determinada pela hamiltoniana ($\mathcal{H}(\vec{p}, \vec{q})$) que descreve a energia total do sistema; \vec{q} e \vec{p} representam as coordenadas e os momentos generalizados das partículas no sistema. Para um sistema descrito por N partículas Z pode ser escrita como

$$Z = \frac{1}{h^{3N} N!} \frac{1}{V_0} \iint e^{\frac{-\mathcal{H}(\vec{p}, \vec{q}) + pV}{k_B T}} d\vec{p} d\vec{q} dV, \quad (9)$$

em que h é a constante de Planck. A entropia pode ser obtida a partir da fórmula

$$G = H - TS, \quad (10)$$

em que H é a entalpia do sistema.

Métodos da Perturbação e Integração Termodinâmica

Os métodos de integração e perturbação termodinâmicas calculam diferenças de energia livre entre dois estados A e B ao criar um dependência da hamiltoniana com um parâmetro λ que liga os dois estados.²¹ De maneira geral a hamiltoniana que liga os dois estados pode ser escrito na forma

$$\mathcal{H}(\vec{p}, \vec{q}, \lambda) = (1 - \lambda)\mathcal{H}_A(\vec{p}_A, \vec{q}_A, \lambda) + \lambda\mathcal{H}_B(\vec{p}_B, \vec{q}_B, \lambda). \quad (11)$$

Desta forma, para $\lambda = 0$ o sistema é descrito pelo hamiltoniana para \mathcal{H}_A e para $\lambda = 1$, o sistema é descrito pela hamiltoniana para \mathcal{H}_B . Assim a energia livre de Gibbs passa a ser uma função de λ , na forma

$$G(\lambda) = -k_B T \ln Z(\lambda). \quad (12)$$

As diferenças de energia entre os estados A e B passa a ser escrita na forma

$$\Delta G_{BA} = G_B(\lambda_B) - G_A(\lambda_A) = -k_B T \ln \frac{Z(\lambda_B)}{Z(\lambda_A)}. \quad (13)$$

Expressando-se a **Equação 13** com uma média sobre um *ensemble*, tem-se que

$$\Delta G_{BA} = -k_B T \ln \left\langle e^{\frac{\mathcal{H}_B(\vec{p}, \vec{q}, \lambda_B) - \mathcal{H}_A(\vec{p}, \vec{q}, \lambda_A)}{k_B T}} \right\rangle_{\lambda_A}, \quad (14)$$

em que $\langle \dots \rangle_\lambda$ representa a média sobre o *ensemble* no espaço de fase \vec{q} e \vec{p} para o valor de λ . A **Equação 14** é conhecida como Equação de Perturbação Termodinâmica,²¹ pois o valor de ΔG_{BA} só pode ser calculado quando os estados A e B são muito próximos. Caso os estados sejam distantes, pode-se criar estados intermediários mais próximos que conectem os estados A e B .

A derivada da **Equação 12** em relação a λ gera

$$\left. \frac{\partial G(\lambda)}{\partial \lambda} \right|_T = -\frac{k_B T}{Z(\lambda)} \frac{\partial Z(\lambda)}{\partial \lambda} = \left\langle \frac{\partial \mathcal{H}(\vec{p}, \vec{q}, \lambda)}{\partial \lambda} \right\rangle_\lambda. \quad (15)$$

Integrando-se a **Equação 15** no intervalo de λ_A até λ_B tem-se

$$\Delta G_{BA} = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial \mathcal{H}(\vec{p}, \vec{q}, \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda. \quad (16)$$

A **Equação 16** é conhecida como Equação de Integração Termodinâmica.²¹ O valor de ΔG_{BA} pode ser calculado durante uma única simulação entre estados A e B , contanto que a mudança em λ seja reversível.

Método da Razão de Aceitação de Bennett

Outro método que utiliza-se da **Equação 13** é método da razão de aceitação de Bennett, conhecido em inglês como “Bennett Acceptance Ratio” ou BAR.²⁶ Este se baseia no fato que a **Equação 13** pode ser expressa na forma

$$\Delta G_{BA} = -k_B T \ln \frac{Z(\lambda_B)}{Z(\lambda_A)} = -k_B T \ln \frac{\langle W(q^N) e^{-\frac{\mathcal{H}(\vec{p}, \vec{q}, \lambda_B)}{k_B T}} \rangle_{\lambda_A}}{\langle W(q^N) e^{-\frac{\mathcal{H}(\vec{p}, \vec{q}, \lambda_A)}{k_B T}} \rangle_{\lambda_B}}, \quad (17)$$

em que $W(q^N)$ é uma função de ponderação que ainda será definida. No caso em que as amostragens dos estados λ_A e λ_B sejam independentes e produzam n_{λ_A} e n_{λ_B} configurações, $W(q^N)$ satisfaz a condição

$$W(q^N) = C \left\{ \frac{Z(\lambda_B)}{n_{\lambda_B}} e^{-\frac{\mathcal{H}(\vec{p}, \vec{q}, \lambda_A)}{k_B T}} + \frac{Z(\lambda_A)}{n_{\lambda_A}} e^{-\frac{\mathcal{H}(\vec{p}, \vec{q}, \lambda_B)}{k_B T}} \right\}^{-1}, \quad (18)$$

em que a constante C ainda será determinada. Das **Equações 17 e 18** tem-se

$$\frac{Z(\lambda_B)}{Z(\lambda_A)} = e^c \frac{\langle f(\mathcal{H}(\vec{p}, \vec{q}, \lambda_B) - \mathcal{H}(\vec{p}, \vec{q}, \lambda_A) + C) \rangle_{\lambda_A}}{\langle f(\mathcal{H}(\vec{p}, \vec{q}, \lambda_A) - \mathcal{H}(\vec{p}, \vec{q}, \lambda_B) + C) \rangle_{\lambda_B}}, \quad (19)$$

em que a constante f é a função de Fermi $f(x) = \{1 + e^{(x)}\}^{-1}$, e a constante C é definida como

$$e^C = \frac{Z(\lambda_A)n_{\lambda_B}}{Z(\lambda_B)n_{\lambda_A}}, \quad (20)$$

Escrevendo-se os valores esperados da **Equação 19** na forma de média sobre configurações e substituindo-se na **Equação 17** tem-se

$$\Delta G_{BA} = k_B T \left\{ \ln \frac{\sum_{n_{\lambda_B}} f(\mathcal{H}(\vec{p}, \vec{q}, \lambda_B) - \mathcal{H}(\vec{p}, \vec{q}, \lambda_A) + C)}{\sum_{n_{\lambda_A}} f(\mathcal{H}(\vec{p}, \vec{q}, \lambda_A) - \mathcal{H}(\vec{p}, \vec{q}, \lambda_B) + C)} - \ln \frac{n_{\lambda_B}}{n_{\lambda_A}} + C \right\}, \quad (21)$$

Para simulações suficientemente grandes que satisfazem a condição de balanço detalhado, o primeiro termo dentro das chaves da **Equação 21** torna-se zero e, portanto, a energia livre é dada por

$$\Delta G_{BA} = k_B T \left\{ C - \ln \frac{n_{\lambda_B}}{n_{\lambda_A}} \right\}, \quad (22)$$

Na prática a energia livre é calculada resolvendo-se as **Equações 21 e 22** iterativamente para um valor inicial de C , até que a convergência seja atingida. Este protocolo tem sido adotado em programas de simulação de dinâmica molecular, tais como GROMACS.²⁷ O sucesso do BAR depende da sobreposição dos estados A e B . E novamente, caso os estados sejam distantes, pode-se criar estados intermediários mais próximos que conectem os estados A e B .

Modelos Atomísticos de Carboidratos

Existem mais de vinte campos de força atomísticos para carboidratos. Cada de campo de força tem suas vantagens e desvantagens. Stortz *et al.*²⁸ demonstrou que os campos de força GROMOS 45A4,²⁹ GLYCAM06³⁰ e MM3³¹ são as melhores escolhas

para descrever as conformações experimentais de 54 diferentes conformêmeros com diferentes orientações do grupos hidroximetil, hidroxila e das ligações glicosídicas (**Figura 7**). Embora as orientações do grupo hidroximetil fornecidas pelo GROMOS 45A4 estejam de acordo com os dados experimentais (**Tabela 1**) a proporção entre os arranjos do anel não apresentam concordância com os valores experimentais (**Tabela 2**).³²⁻³⁵

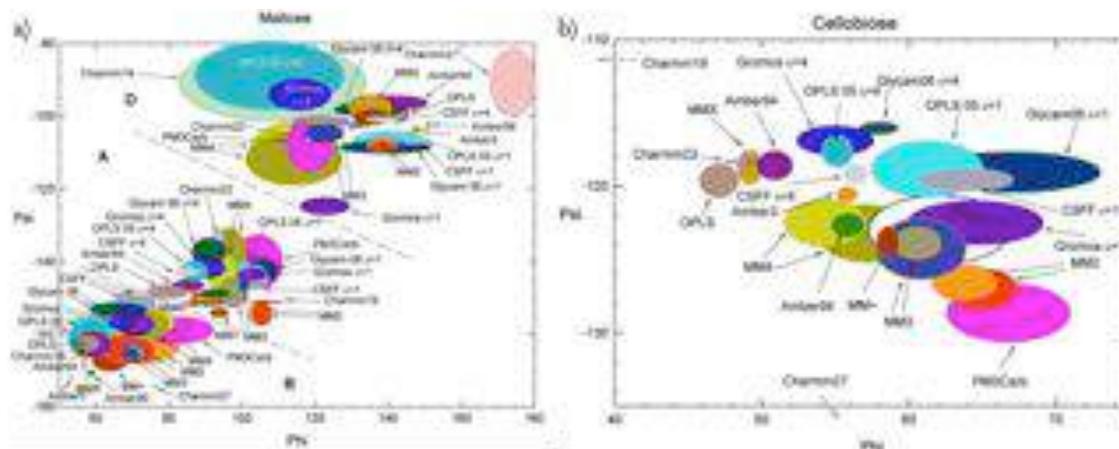


Figura 7. Representação dos valores de ϕ e ψ para a maltose (a) e celobiose (b) com os valores médios de cada diedro no centro de cada elipse. A magnitude dos eixos das elipses representam os desvios padrão de cada modelo. Figura adaptada de Carbohydrate Research, 344/16, Carlos A. Stortz, Glenn P. Johnson, Alfred D. French, Gábor I. Csonka, 12, Copyright (2009), com permissão da Elsevier. Adaptada de The Lancet, 16, Carlos A. Stortz, Glenn P. Johnson, Alfred D. French, Gábor I. Csonka, Comparison of different force fields for the study of disaccharides, 12, Copyright (2009), com permissão da Elsevier.

Para corrigir a população entre dos diferentes arranjos de anel obtidas pelo campo de força GROMOS 45A4, os potenciais torcionais dos átomos que compõe os anéis de açúcares do campo de força GROMOS 45A4 foram melhor ajustados aos seus respectivos perfis calculados com o método HF/6-31G(d).^{36,37} Este novo conjunto de parâmetros foi introduzido dentro do campo de força GROMOS 53A6, sendo denominado GROMOS 53A6_{GLYC}.³⁸ Lembrando que o campo de força GROMOS 53A6 para carboidratos é o mesmo do GROMOS 45A4.

Tabela 1. Populações relativas das três conformações do grupo hidroximetil, obtidas durante 20 ns de simulação da α - e β -D-glicose, descrito pelo campo de força GROMOS 45A3. Tabela adaptada de Lins, R. D.; Hünenberger, P. H. *J. Comput. Chem.* **2005**, *26*, 1400–1412, Copyright (2012) com permissão da American Chemical Society.

Conformação	α -D-glicose	β -D- glicose
60 (gt)	46 (44 ^a , 54 ^b , 53 ^c)	45 (45 ^a , 62 ^b , 61 ^c)
180 (tg)	0 (0 ^a , -1 ^b , 7 ^c)	0 (2 ^a , -7 ^b , 8 ^c)
300 (gg)	54 (56 ^a , 47 ^b , 40 ^c)	55 (53 ^a , 45 ^b , 31 ^c)

As conformações gt, tg e gg foram definidas nos intervalos 0 – 120, 120 – 240 e 240 – 360, respectivamente do ângulo diédrico do grupo hidroximetil (O5-C5-C6-O6). Os valores entre parênteses correspondem às populações relativas obtidas por espectros de RMN por ^aNishida et al., ³⁹⁻⁴¹ ^bBrochier-Salon e Molin,⁴² e ^cThibaudeau et al.⁴³

Tabela 2. Energias interna e livre de Gibbs relativas da α -D-glicose em relação à conformação ⁴C₁. Tabela adaptada de Pol-Fachin, L.; Rusu, V. H.; Verli, H.; Lins, R. D. *J. Chem. Theory Comput.* **2012**, *8*, 4681–4690, Copyright (2012) com permissão da American Chemical Society e adaptada de Lins, R. D.; Hünenberger, P. H. *J. Comput. Chem.* **2005**, *26*, 1400–1412, Copyright (2012) com permissão da American Chemical Society.

Conformação	GROMOS 45A4		GROMOS 53A6 _{GLYC}	RMN / %
	ΔE / kJ mol ⁻¹	ΔG / kJ mol ⁻¹	ΔG / kJ mol ⁻¹	
⁴ C ₁	0.0	0.0	0.0	100
¹ C ₄	+34.8	+18.0	+49.0	0

O novo campo de força GROMOS 53A6_{GLYC} aumenta a diferença de energia entre os arranjos de anel (**Tabela 2**), apresentando uma melhor concordância com dados experimentais de RMN para as populações das conformações de anel de 16 diferentes açúcares (Tabela 4 da Referência ³⁸). As correções nos termos torcionais do campo de força GROMOS 53A6_{GLYC} não interferem significativamente nas orientações do grupo hidroximetil (Tabela 5 da Referência ³⁸), mas apresenta a proporção correta (100%) de conformação ⁴C₁ em 1 μ s de simulação de α -D-glicose em solução, o que não é observado com o campo de força GLYCAM06.⁴⁴

Modelos “*coarse-grain*” para Carboidratos

Diversas estratégias de parametrização de campos de força CG foram desenvolvidas ao longo dos últimos anos,⁴⁵⁻⁴⁹ inclusive na quantidade de esferas CG contidas nos modelos, especialmente em carboidratos.⁵⁰⁻⁶⁰ Para facilitar a comparação dos modelos já desenvolvidos para carboidratos, classificou-se os campos de força pelo número de esferas utilizadas para representar a molécula de glicose.

Modelos com Três Esferas

O M3B foi o primeiro campo de força CG para carboidratos, sendo desenvolvido por Molinero e Goddard III, em 2004.⁵⁰ O modelo mapeia a α -D-glicose com três esferas cujas posições coincidem com os átomos C1, C4 e C6 (**Figura 8**). A molécula de água é representada por uma esfera CG.

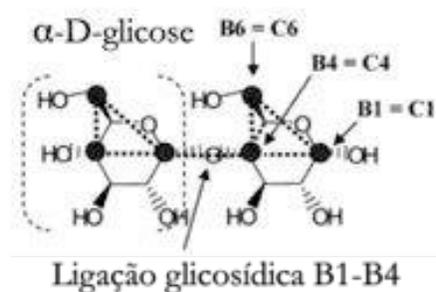


Figura 8. Representações atômicas e M3B sobrepostas da molécula de maltose. A posição das esferas CG M3B (B1, B4 e B6) coincidem com os átomos de carbono (C1, C4 e C6) do modelo atômico. As três esferas do modelo M3B estão conectadas por ligações (linhas tracejadas) e uma única ligação entre resíduos representa a ligação glicosídica. Figura adaptada de Molinero, V.; Goddard, W. A. *J. Phys. Chem. B* **2004**, *108*, 1414–1427, Copyright (2004) com permissão da American Chemical Society.

Os parâmetros de interação para os termos não ligados foram ajustados empiricamente para mimetizar o comportamento da α -D-glicose descrita pelo campo de força UFF DREIDING.⁶¹ O modelo utiliza potenciais harmônicos de ligação e ângulo, associados a potenciais de diédricos com funções cosseno para expressar as interações ligadas e o potencial de Morse para calcular as interações não ligadas, não incluindo cargas elétricas nas esferas CG.

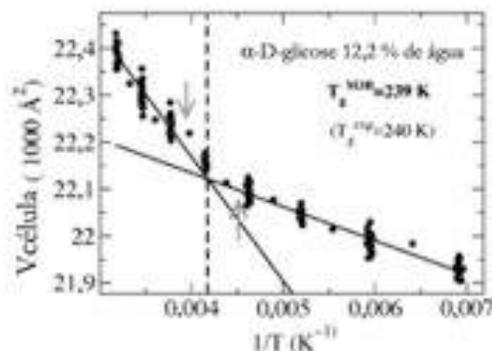


Figura 9. Determinação da temperatura de transição vítrea para a α -D-glicose com 12,2 % de água na representação M3B. Figura adaptada de Molinero, V.; Goddard, W. A. *J. Phys. Chem. B* **2004**, *108*, 1414–1427, Copyright (2004) com permissão da American Chemical Society.

O M3B consegue reproduzir a temperatura de transição vítrea experimental (**Figura 9**), as estruturas cristalinas da V-amilose, a energia de coesão entre a α -D-glicose e a água (comparados com o modelo UFF DREIDING). No entanto, a descrição das interações não ligadas por um potencial de Morse não é a maneira mais computacionalmente eficiente quando não se pretende descrever quebra e formação de ligações químicas. De fato, Liu *et al.*⁵¹, em 2007, abordam este problema utilizando o métodos de ajuste de força por mínimos quadrados, do inglês “*Least-Square Force Matching*”,^{45,62} para criar uma tabela de interações intermoleculares a partir do campo de força atômico OPLS/AA⁶³ usando o modelo de água SPC/E.⁶⁴ A representação CG da α -D-glicose é a mesma do modelo M3B (**Figura 8**), o único açúcar estudado é a α -D-glicose. Os autores mostram que o modelo CG pode ser usado em diferentes temperaturas e pressões daquelas usadas na parametrização do modelo.

O modelo proposto por Liu *et al.* é ~ 3 ordens de grandeza mais rápido do que atômico, no entanto, os autores não comentam sobre o passo de integração das equações de movimento, visto que o modelo M3B é ~ 4 ordens de grandeza mais rápido do que o modelo atômico ao utilizar um passo de integração uma ordem de grandeza maior (10 fs) do que o comumente usado em simulações atômicas sem restrições nas ligações (1-2 fs).

Usando a mesma representação CG para a glicose (**Figura 8**), Bu *et al.*, em 2009, construíram um modelo para descrever a celulose I β .⁵² No entanto, os autores utilizam o modelo generalizado de Born com uma função “switch” para representar o solvente (solvente implícito). Os termos ligados foram obtidos através do método de inversão de Boltzmann a partir de simulações atomísticas com os modelos CHARMM22⁶⁵ e CSFF^{66,67} com água TIP3P.⁶⁸ As interações não ligadas foram definidas a partir de parâmetros de Lennard-Jones escalonados, sendo que os parâmetros ϵ foram obtidos do modelo M3B e os parâmetros σ a partir de distância entre os átomos correspondentes no modelo atomístico. Este modelo consegue descrever os parâmetros de célula experimentais do cristal de celulose I β (**Tabela 3**).

Tabela 3. Parâmetros de célula do cristal de celulose I β nos modelo atomísticos (CHARMM22 e CSFF), CG celulose (Modelo de Bu *et al.*) e experimental^a. Tabela adaptada de Bu, L.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. *J. Phys. Chem. B* **2009**, *113*, 10994–11002, Copyright (2009) com permissão da American Chemical Society.

Parâmetro	Cristal ^b	Modelo atomístico ^c	Modelo CG
$a / \text{Å}$	7,8	8,5	10,8
$b / \text{Å}$	8,2	8,1	8,1
$c / \text{Å}$	10,4	10,5	10,5
$\gamma / ^\circ$	96,5	~90	90,0
Torsão / $^\circ$	0	1,4 – 1,7	2,5

^a Os parâmetros de célula b e c definem o plano (100)

^b Referências ^{69,70}

^c Referências ⁷¹

Um modelo CG mais completo para carboidratos foi desenvolvido em 2009 por López *et al.* como uma extensão ao campo de força MARTINI.^{53,72} O modelo inclui os monômero de α -D-glicose e β -D-frutose, bem como diversos dissacarídeos. Posteriormente este modelo foi combinado com lipídeos para formar o primeiro conjunto de parâmetros CG para glicolipídeos.⁷³

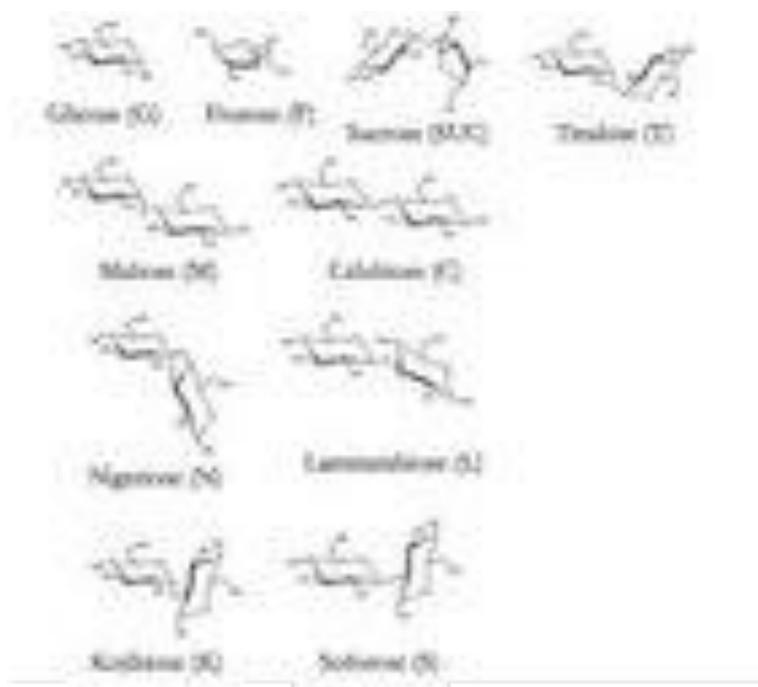


Figura 10. Açúcares parametrizados no modelo MARTINI por López *et al.*: α -glicose(G); β -frutose (F); sucrose (SUC); trealose (T); maltose (M); celobiose (C); nigeriose (N); laminarabiose (L); kojibiose (K) e sofrorse (S). Figura adaptada de López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. *J. Chem. Theory Comput.* **2009**, 5, 3195–3210, Copyright (2009) com permissão da American Chemical Society.

O mapeamento atômico para CG é feito através do centro de massa dos átomos que compõe cada esfera CG. Uma característica importante do modelo MARTINI é que o mapeamento de monossacarídeos é diferente de dissacarídeos e oligossacarídeos (**Figura 11**).

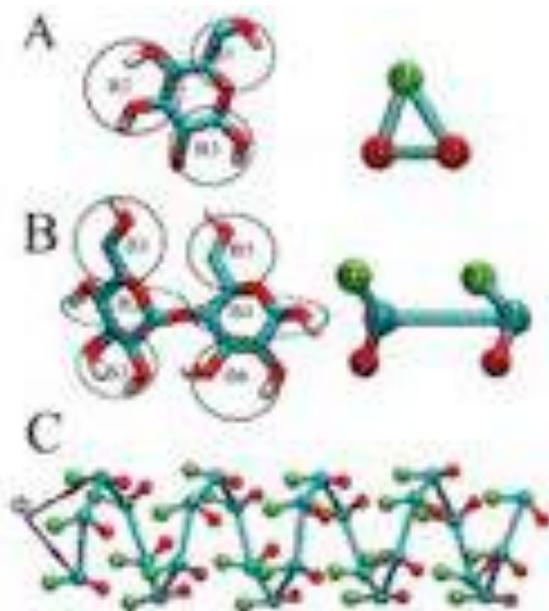


Figura 11. Mapeamento atomístico para CG de mono (A); di- (B) e oligossacarídeos (C) no modelo MARTINI. Figura adaptada de López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210, Copyright (2009) com permissão da American Chemical Society.

Por ser uma extensão do MARTINI, o campo de força é baseado na função potencial do campo de força GROMOS. Os tipos de esferas CG foram escolhidas da lista de esferas CG MARTINI utilizando-se cálculos de energia livre e as constantes de força dos termos ligados foram ajustadas manualmente para representar as distribuições atomísticas mapeadas para CG.

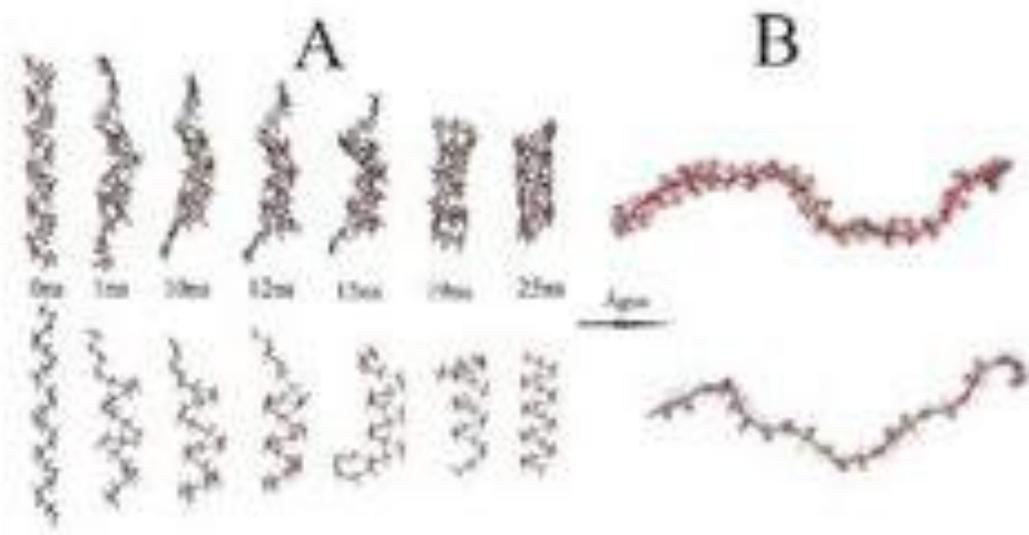


Figura 12. Estruturas da amilose V com 26 monômeros em nonano (A) e em água (B). Na parte superior representação atomística e na inferior CG MARTINI. (A) Ambas as representações em nonano convergiram para um estrutura em V estável em 25 ns; (B) A transferência para água promove um aumento de flexibilidade e o desenovelamento das estruturas em hélice. Figura adaptada de López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210, Copyright (2009) com permissão da American Chemical Society.

O modelo MARTINI é capaz de prever a estrutura da amilose V em solventes de baixa polaridade (**Figura 12A**), bem como descrever a alta flexibilidade da amilose V em água (**Figura 12B**).

Em 2011, uma nova extensão ao MARTINI foi desenvolvida que adiciona parâmetros para diversos oligômeros da β -D-glicose para simular celulose I β .⁵⁴ O modelo foi testado em combinação com o modelo MARTINI para proteínas, ao simular o domínio de ligação de açúcar da celulase fúngica (Cel7A) da *Trichoderma reesei*.

Seguindo a mesma estratégia utilizada por Liu *et al.* (interações ligadas mapeadas através do método de inversão de Boltzmann e interações não ligadas através do método de “force matching”), Hynninen, em 2011, desenvolveu um outro modelo CG para β -D-glicose.⁵⁵ No entanto, as posições das esferas foram mapeadas através do centro de massa do átomos que as compõe (**Figura 13**). O modelo é capaz de descrever a função de distribuição radial das fibras de celulose semelhantemente ao campo de força CHARMM (conjunto de parâmetros não reportados).⁵⁵

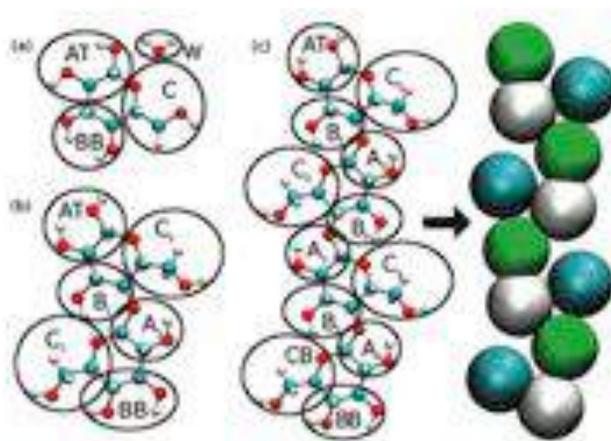


Figura 13. Estrutura atomística com as definições das esfera CG para o modelo de Hynnine *et al.* para a (a) glicose e água; (b) celobiose e (c) celotrealose. A representação

CG da celotetraose está indicada pela seta. Figura adaptada de Hynninen, A.-P.; Matthews, J. F.; Beckham, G. T.; Crowley, M. F.; Nimlos, M. R. *J. Chem. Theory Comput.* **2011**, 7, 2137–2150, Copyright (2011) com permissão da American Chemical Society.

Outros modelos “coarse-grain”

Modelos com diferentes números de esferas CG para representar monossacarídeos também foram desenvolvidos. Srinivas *et al.*, em 2011, introduziu um modelo com uma esfera e sem solvente para simular as transições de fase entre as fases cristalina e amorfa das fibras de celulose.⁵⁶

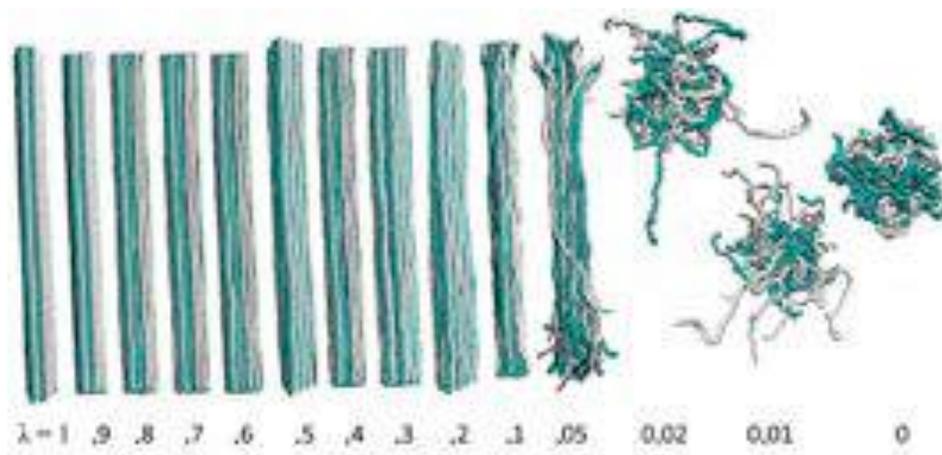


Figura 14. Estruturas finais da simulação da transição entre as fases cristalina ($\lambda = 1$) e amorfa ($\lambda = 0$) da celulose. Figura adaptada de Srinivas, G.; Cheng, X.; Smith, J. C. *J. Chem. Theory Comput.* **2011**, 7, 2539–2548, Copyright (2011) com permissão da American Chemical Society.

O modelo é baseado no campo de força CHARMM (conjunto de parâmetros não reportados) e a transição de fase é obtida através da modificação do parâmetro λ , que conecta as hamiltonianas dos estados amorfo ($\lambda = 0$) e cristalino ($\lambda = 1$) (**Figura 14**).⁵⁶

Tentando reproduzir no modelo CG o movimento coletivo dos átomos de simulações atomísticas, Glass *et al.* desenvolveu um modelo de esfera única e sem solvente, em 2012.⁵⁷ A parametrização foi feita através do método REACH (Realistic Extension Algorithm), no qual as constantes de força entre as esferas CG são obtidas

diretamente da matriz de variância-covariância calculada a partir das simulações atomísticas.⁴⁷⁻⁴⁹ O modelo REACH não usa qualquer dado experimental, apenas dados de simulações atomísticas e foi aplicado para descrever propriedades mecânicas da celulose I β , tais como, velocidade do som transmitida ao longo das fibras e módulo de Young.

O primeiro modelo contendo cinco esferas foi proposto em 2012 por Bellesia *et al.* para estudar as transições de fase entre celulose I β e celulose III_I.⁵⁸ O modelo mapeia o centro de massa dos átomos C1, C2, C3, C4, C5 e O5 na esfera X, a esfera J o átomo O1, as esferas Y e Z representam os átomos C6 e O6, respectivamente. A esfera W representa o centro de massa dos oxigênio O2 e O3 (**Figura 15**).



Figura 15. Representação esquemática da cadeia de celulose na representação atomística (topo) e na representação CG de Bellesia (meio). Na parte inferior a comparação geométrica dos modelos CG de Bellesia e atomístico para a molécula de glicose. Figura adaptada de Bellesia, G.; Chundawat, S. P. S.; Langan, P.; Redondo, A.; Dale, B. E.; Gnanakaran, S. *J. Phys. Chem. B* **2012**, *116*, 8031–8037, Copyright (2012) com permissão da American Chemical Society.

Similarmente ao modelo desenvolvido por Srinivas, o modelo de Bellesia não utiliza solvente e foi parametrizado usando-se dados cristalinos das diferentes fases da celulose. Os termos não ligados foram obtidos a partir de potenciais de Lennard-Jones escalados. Novamente a transição entre os estados é obtida através da variação do

parâmetro λ que conecta os estados celulose I β ($\lambda = 0$) e celulose III $_I$ ($\lambda = 1$) (**Figura 16**).

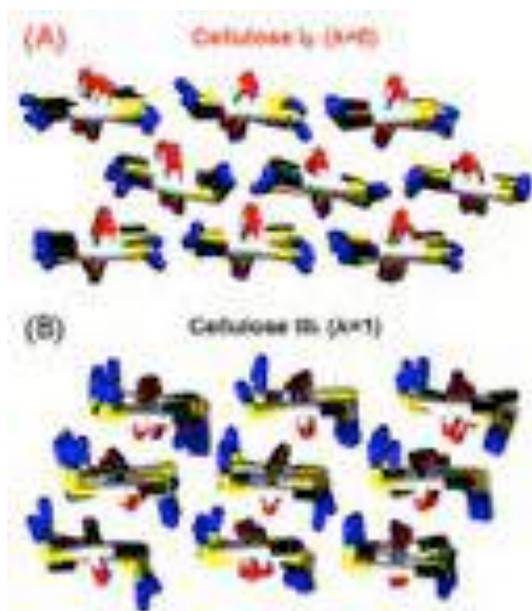


Figura 16. Estruturas finais da simulação da transição entre as fases I β ($\lambda = 0$) e III $_I$ ($\lambda = 1$) da celulose. Figura adaptada de Bellesia, G.; Chundawat, S. P. S.; Langan, P.; Redondo, A.; Dale, B. E.; Gnanakaran, S. *J. Phys. Chem. B* **2012**, *116*, 8031–8037, Copyright (2012) com permissão da American Chemical Society.

O modelo prevê uma maior instabilidade da celulose III $_I$ comparada à celulose I β devido ao termo de entalpia, e propõe que a menor estabilidade térmica do cristal da celulose III $_I$ é devido a sua maior entropia.

De maneira sistemática, em 2013, Markutsya *et al.* aplicaram o método de “*force matching*” para diferentes esquemas de mapeamento da glicose, de uma a quatro esferas (**Figura 17**).⁶⁰ Os autores sugerem que os modelos com uma e duas esferas são os melhores para simular propriedades associadas a cadeias, tais como transições entre estruturas amorfas e cristalinas. Eles também sugerem que o aumento no número de esferas permite uma melhor descrição dos ângulos diédricos da ligação glicosídica e que os modelos com três e quatro têm graus de liberdade suficiente para descrever mudanças conformacionais associadas às unidades de açúcar em polissacarídeos. No entanto, verificou-se que o modelo de quatro esferas assumiu orientações não físicas em simulações de celulose I α .

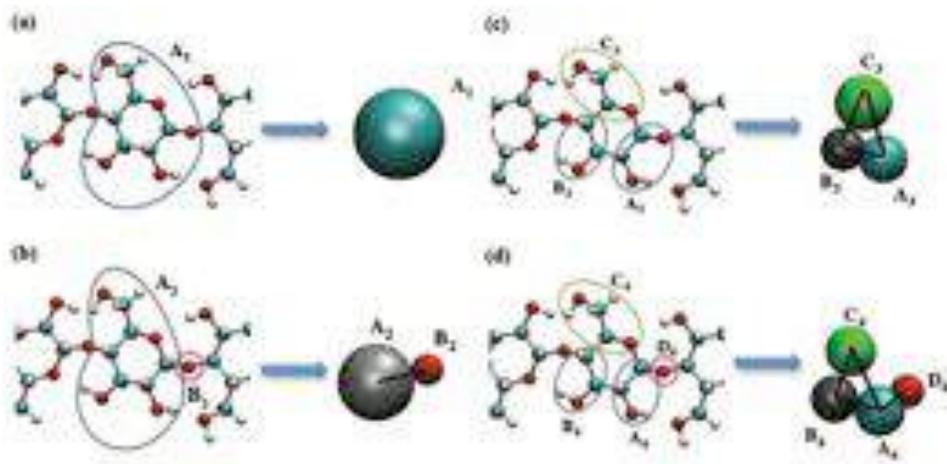


Figura 17. Mapeamento CG para o resíduo de glicose nos quatro diferentes modelos CG desenvolvidos por Markutsya. Figura adaptada de Markutsya, S.; Devarajan, A.; Baluyut, J. Y.; Windus, T. L.; Gordon, M. S.; Lamm, M. H. *J. Chem. Phys.* **2013**, *138*, 214108–214117, Copyright (2013) com permissão da AIP Publishing LLC.

O mesmo grupo de pesquisa também mapeou, em 2012, a molécula de β -D-glicose contendo seis esferas e uma esfera para molécula de água.⁵⁹ O modelo foi parametrizado por “*force matching*” a partir de simulações com o campo de força CHARMM com água TIP3P.⁶⁸

O modelo de seis esferas permitiu um mapeamento mais fácil dos ângulos diédricos das ligações glicosídicas, bem como a possibilidade de calcular as populações dos arranjos do anel da glicose. No entanto, os autores reportam populações de arranjo de anel não esperada, a saber, 20% de barco nas simulações atomísticas e 40% nas simulações CG, lembrando que dados de RMN mostram que a β -D-glicose assume apenas conformações 4C_1 em água. Estes resultados demonstram o cuidado que deve ser tomado na escolha dos modelos atomísticos utilizados na parametrização de modelos CG.

Hipótese da Tese

No contexto dos campos de força e modelos já desenvolvidos, propõe que é possível desenvolver um modelo CG universal para carboidratos, denominado PITOMBA, baseado em quatro esferas capaz de:

- Permitir a diferenciação visual as ligações glicosídicas de açúcares α e β ;
- Contenha um único conjunto de esferas CG para mapear monossacarídeos e oligossacarídeos;
- Seja baseado em construção de blocos, i. e., oligossacarídeos sejam construídos a partir da combinação de monossacarídeos, e estes sejam construídos a partir da combinação de esferas CG. Sendo permitido trocar e/ou adicionar as esferas de acordo com o tipo de açúcar;
- Incluir interações eletrostáticas explicitamente;
- Descrever o perfil dos ângulos diédricos das ligação glicosídica, sem que o modelo acesse regiões não físicas;
- Descrever as geometrias das cadeias de açúcares em solução;
- Manter a compatibilidade com o modelo de parametrização do GROMOS;
- Ser facilmente implementado em programas de alto desempenho.

Construção do Campo de Força PITOMBA

O desenvolvimento do campo de força PITOMBA[†] inclui diversas etapas, i. definição da filosofia do modelo; ii. definição das estratégias de parametrização; iii. parametrização e iv. validação do modelo. Neste contexto, este capítulo tem como objetivos apresentar a filosofia do modelo PITOMBA, as suas características gerais, a estratégia de parametrização e os parâmetros do modelo. Além de descrever as condições de simulação tanto na parametrização bem como na validação do modelo. A lista de programas desenvolvidos para auxiliar o mapeamento, parametrização e análise das trajetórias CG PITOMBA pode ser vista no Apêndice A.

Características Gerais do Modelo

O PITOMBA foi construído com o objetivo de reproduzir propriedades das cadeias e motivos estruturais de açúcares em solução. Para isso, ele foi concebido como um modelo de quarto esferas, em que as interações eletrostáticas são consideradas ao atribuir cargas pontuais a cada esfera CG e o modelo de água SPC CG GROMOS⁷⁴ como solvente, mas podendo ser usado com o modelo de água SPC.⁷⁵ O PITOMBA baseia-se no modelo atomístico GROMOS 53A6_{GLYC},³⁸ em dados experimentais e na forma funcional do campo de força GROMOS.^{22,23} A filosofia de construção do modelo permite incluir, sem nenhuma extensão, cadeias de açúcares com diferentes ligações glicosídicas. Além de múltiplas cadeias na forma de ligações glicosídicas do tipo 1 → 1 α , 1 → 4 α , 1 → 6 α , 1 → 1 β , 1 → 4 β e 1 → 6 β .

Mapeamento Atomístico para “Coarse-grain”

O campo de força PITOMBA emprega quatro esferas CG (CV1, CV2, CVM e CV6) para representar cada resíduo de glicose, posicionadas no centro de massa de grupo de átomos que constituem cada esfera, como mostra a **Figura 18** e **Tabela 4**.

[†] A PITOMBA é uma fruta típica de Pernambuco, arredondada, com casca marrom e interior composto por uma semente envolvida por uma polpa, que dentre outras coisas, contém açúcar. E ainda é uma sigla em inglês para **Parameter Interface To Oligosaccharides Molecules Based on Atoms**.

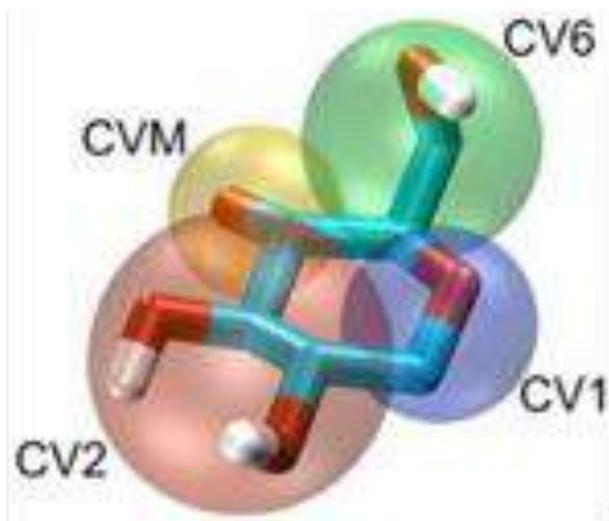


Figura 18. Representação do resíduo de glicose no campo de força PITOMBA e os átomos que constituem cada esfera CG na representação de átomos unidos do GROMOS96

Escolheu-se mapear o resíduo ao invés da molécula de glicose porque as cadeias de polissacarídeos não incluem os átomos O1, HO1 e HO4, visto que na reação de condensação há a perda de uma molécula de água.

Tabela 4. Massa, em unidades de massa atômicas, e átomos, na numeração IUPAC, que constituem cada tipo de superátomo no modelo PITOMBA.

Tipo da Esfera CG	Massa / u. m. a.	Átomos constituintes
CV1	29	O5 – C1 – H1
CV2	60	C2 – H2 – O2 – HO2 – C3 – H3 – O3 – HO3
CVM	29	C4 – H4 – O4
CV6	44	C5 – H5 – C6 – H61 – H62 – O6 – HO6

O uso de quatro esferas na representação do resíduo de glicose, permite o desenvolvimento de um modelo com algumas propriedades, tais como, i. definição explícita de um ângulo diédrico ao longo da ligação glicosídica; ii. cargas elétricas derivadas matematicamente das equações de dipolo e de conservação de cargas; iii.

correta representação das orientações equatorial e axial das ligações glicosídicas 1→1, 1→4 e 1→6 e iv. mapeamento direto entre as representações atomísticas e CG.

Modelo de Ligação em “Pipa”

Com o objetivo de diminuir o número de termos do modelo os quatro ângulos internos dos anéis de açúcar foram substituídos por duas ligações cruzadas (CV2 – CV6 e CV1 – CVM). Desta forma, as ligações intra e inter resíduos de glicose se assemelha a uma pipa (**Figura 19**).

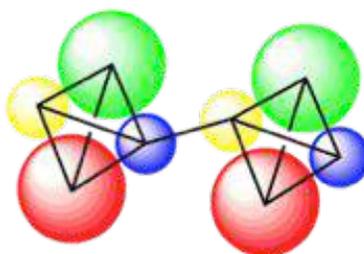


Figura 19. Representação esquemática da maltose, no modelo PITOMBA, enfatizando as ligações do modelo em forma de pipa.

Com o uso do modelo de ligações em pipa, é mais fácil parametrizar o PITOMBA seguindo a estratégia de parametrização adotada, discutida na próxima seção.

Estratégia de Parametrização

A parametrização do PITOMBA segue quatro passos:

1. Os estiramentos de ligação, deformações angulares e deformações de ângulo impróprio do PITOMBA foram definidos para ajustar os valores médios de simulações de 100 ns com o campo de força GROMOS 53A6_{GLYC} para as moléculas de glicose, maltose e celobiose em água SPC;
2. As cargas pontuais das esferas CG foram definidos para ajustar o momento de dipolo do resíduo de glicose, em água SPC, para cada 1000 passos da trajetória durante os 100 ns de simulação;
3. Os parâmetros de van der Waals para cada esfera CG foram definidos para reproduzir as energias de solvatação experimental de suas moléculas de

referência (**Tabela 4**). A molécula de referência é definida como sendo a molécula, cuja valência foi preenchida com hidrogênio, associada a cada esfera CG. As energias de solvatação foram obtidas no modelo de água SPC CG GROMOS.

4. Os termos diédricos da ligação glicosídica foram ajustados para reproduzir o perfil HF/-31G(d)^{36,37} da maltose e celobiose, da mesma forma que foram parametrizados no campo de força GROMOS 53A6_{GLYC}.

Condições de Simulações e Sistemas Simulados

Todas as simulações atomísticas foram realizadas usando-se o programa GROMACS 4.6.1²⁷ por 100 ns usando condições periódicas de contorno nas direções x , y e z no *ensemble* N,p,T e modelo de água SPC. A quantidade de partículas em casa simulação pode ser vista na **Tabela 5**. A temperatura foi mantida constante em 298 K com o auxílio do termostato “*velocity rescale*” com tempo de relaxação de 0,2 ps.⁷⁶ A pressão foi mantida em 1 atm usando-se o barostato de Berendsen com tempo de acoplamento de 0,5 ps e compressibilidade isotérmica (κ_T) de $4,51 \cdot 10^{-5}$ (kJ mol⁻¹ nm⁻³)⁻¹.⁷⁷ O método LINCS foi utilizado para restringir as ligações O-H, permitindo assim o uso de um passo de integração de 2 fs.⁷⁸ Aplicou-se um raio de corte nas interações de van der Waals de 14 Å. As interações eletrostáticas de longo alcance foram tratadas com o modelo de campo de reação com constante dielétrica relativa (ϵ_r) de 61 e raio de corte de 14 Å.⁷⁹

As simulações do PITOMBA foram realizadas com uma versão modificada do GROMACS 4.6.1, que inclui o termo quártico de interação do modelo de água SPC CG GROMOS. Os açúcares foram embebidos em água SPC CG GROMOS por 100 ns no *ensemble* N,p,T. Exceto pelo κ_T de $7,51 \cdot 10^{-4}$ (kJ mol⁻¹ nm⁻³)⁻¹, a temperatura e pressão foram mantidas constante usando-se as condições e algoritmos das simulações atomísticas. Nenhuma restrição foi adicionada às simulações. As interações entre os termos não ligados foram mantidas nas condições requeridas pelo modelo SPC CG GROMOS, i. e., raio de corte de 20 Å e correções de longo alcance com o método de campo de reação com ϵ_r de 78,5 e ϵ_{CS} de 2,5.⁷⁴

As energias de hidratação das esferas CG em água SPC CG GROMOS foram estimadas usando o método de Razão de Aceitação de Bennett²⁶ com o auxílio da ferramenta `g_bar` do GROMACS. Primeiramente as eliminou-se as interações eletrostáticas e depois as interações de van der Waals. Cada contribuição de energia foi calculada usando 41 valores de λ igualmente espaçados e potenciais “*soft-core*”. Para cada valor de λ foi simulado por 1 ns sendo os 500 ps iniciais tomados como equilíbrio. Os parâmetros dos potenciais “*soft-core*” foram $\lambda^p = 1,0$, $\sigma = 0,3$ nm e $\alpha = 0,5$ nm.⁸⁰

Os cálculos de perfil rotacional da maltose e celobiose foram realizados com o programa GAUSSIAN09⁸¹ utilizando-se o método HF/6-31G(d).^{36,37} A superfície de perfil rotacional foi obtida com 169 pontos espaçados por 30° em cada direção dos ângulos diédricos ϕ e ψ , da qual extraiu-se os valores para o cálculo do perfil rotacional. A lista dos sistemas simulados com o GROMACS, tanto para parametrização do PITOMBA, bem como para sua validação pode ser visto na **Tabela 5** e os arquivos de entrada de alguns sistemas estão descritos no Apêndice D.

Tabela 5. Tempo de simulação, em nanosegundos, número de partículas na caixa de simulação e representação molecular de cada um dos sistemas simulados

Sistema	Tempo simulado / ns	Número de partículas na simulação*	Representação molecular
---------	---------------------	------------------------------------	-------------------------

Glicose	100	3197 (1478)	GROMOS 53A6 _{GLYC} e PITOMBA
Maltose	100	6520 (920)	GROMOS 53A6 _{GLYC} e PITOMBA
Maltose	2	(920)	PITOMBA
Maltose	3,5	(920)	PITOMBA
Celobiose	100	6517 (1476)	GROMOS 53A6 _{GLYC} e PITOMBA
Água GROMOS SPC CG	100	1482	PITOMBA
Amilose V com 10 monômeros	100	2	PITOMBA
Amilose V com 18 monômeros	100	99531 (17270)	PITOMBA
Amilose V com 36 monômeros	100	810174 (184130)	PITOMBA
Amilose A com 18 monômeros em cada fita	100	99450 (17294)	GROMOS 53A6 _{GLYC} e PITOMBA
α -ciclodextrina	100	21429 (2892)	GROMOS 53A6 _{GLYC} e PITOMBA
β -ciclodextrina	100	21425 (2896)	GROMOS 53A6 _{GLYC} e PITOMBA
γ -ciclodextrina	100	21418 (2896)	GROMOS 53A6 _{GLYC} e PITOMBA

* Valores em parêntesis para o modelo PITOMBA

Parametrização dos Termos Ligados

Os termos ligados do PITOMBA foram obtidos a partir dos passos 1 e 4 da estratégia de parametrização. Os termos associados à 1ª etapa (estiramentos de ligação, deformações angulares e deformações de ângulo impróprio) serão apresentados primeiro

e por fim o termo de ângulo diédrico, muito embora este termo só foi definido após os passos 1, 2 e 3.

Contribuições de Ligação

Seguindo a forma funcional do termo de ligação do GROMOS (**Equação 23**) mapeou-se a distâncias de ligação das esferas do PITOMBA para os 100 ns de simulação.

$$V_l = \frac{1}{4}K_l(b_l^2 - b_{0,l}^2)^2, \quad (23)$$

em que K_l é a constante de força de ligação, b_l e $b_{0,l}$ representam a distância e distância de equilíbrio do l -ésimo termo de ligação, respectivamente.

Como esperado, as ligações CG mapeadas, a longo dos 100 ns de simulação, se comportaram como distribuições normais, exceto pelas ligações CV2–CV6 e CVM–CV6 (**Figura 20**). Estas ligações estão associadas à posição do ângulo diédrico do hidroximetil (O5-C5-C6-O6), que pode apresentar as conformações gg (-60°), tg (180°) e gt ($+60^\circ$), sendo a primeira um pouco mais favorável nos campos de força GROMOS 45A4 e 53A6_{GLYC} do que a última e a segunda praticamente não observada.^{29,38} Desta forma, ao invés de tirar uma média das distâncias nas conformações gg e gt, o que representaria a conformação tg, decidiu-se definir as distâncias de ligação entre as esferas CV2–CV6 e CVM–CV6 usando apenas as conformações gg do grupo hidroximetil.

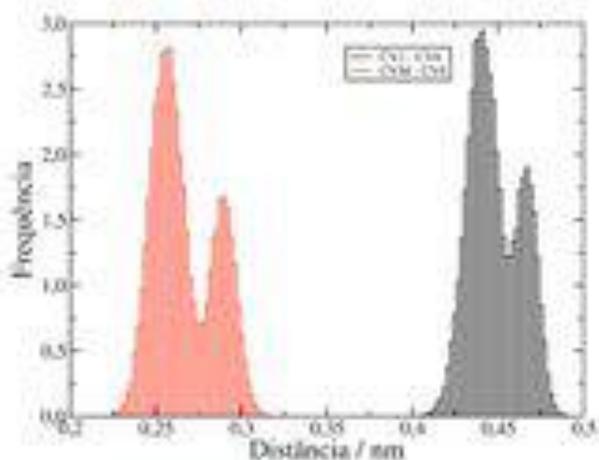


Figura 20. Histogramas normalizados por área das distâncias CV2–CV6 e CVM–CV6 mapeadas de 100 ns uma simulação atomística da glicose com o campo de força 53A6_{GLYC}.

É importante ressaltar que a distância de ligação CV1–+CVM (ligação glicosídica) na maltose apresenta-se como uma distribuição normal ao longo de toda a simulação. As constantes de força foram ajustadas empiricamente para permitir passos de integração de até 7 ps, sem o uso de restrições nas simulações.

Tabela 6. Tipo de ligação, constante de força, distância de equilíbrio e distância médias obtidas com o PITOMBA e mapeadas do modelo atomístico GROMOS 53A6_{GLYC} para a glicose.

Tipo de Ligação	K /kJ mol⁻¹ nm⁻⁴	b_0 /nm	Distância CG / nm	Distância GROMOS 53A6_{GLYC} / nm
CV1 – CV2	$1,110 \cdot 10^6$	0,280	$0,278 \pm 0,004$	$0,280 \pm 0,005$
CV1 – CVM	$5,850 \cdot 10^5$	0,332	$0,326 \pm 0,003$	$0,332 \pm 0,005$
CV1 – CV6	$9,900 \cdot 10^5$	0,279	$0,266 \pm 0,004$	$0,279 \pm 0,008$
CV2 – CVM	$9,800 \cdot 10^5$	0,298	$0,292 \pm 0,004$	$0,298 \pm 0,006$
CV2 – CV6	$5,730 \cdot 10^5$	0,441	$0,427 \pm 0,003$	$0,441 \pm 0,010$
CVM – CV6	$1,155 \cdot 10^6$	0,258	$0,249 \pm 0,004$	$0,258 \pm 0,012$
CV1 – +CVM*	$1,250 \cdot 10^6$	0,227	$0,227 \pm 0,003$	$0,227 \pm 0,007$

* O sinal “+” representa o resíduo seguinte na sequência de resíduos. Distância medida na molécula de maltose.

A **Tabela 6** mostra os valores das constantes de força, distância de equilíbrio e os valores medidos em simulações de 100 ns das moléculas de glicose e maltose nos modelos PITOMBA e GROMOS 53A6_{GLYC}. Como esperado a média das distribuições de distâncias estão dentro das distribuições de distâncias obtidas com o modelo GROMOS 53A6_{GLYC}, no entanto as distribuições do PITOMBA apresentam menores desvios comparados com as distribuições atomísticas, promovendo assim um anel de glicose menos flexível do que no modelo atomístico de referência.

Contribuições Angulares

De acordo com o campo de força GROMOS, os termos angulares têm forma

$$V_a = \frac{1}{2}K_a(\cos\theta_a - \cos\theta_{0,a})^2, \quad (24)$$

em que K_a é a constante de força de ligação, θ_a e $\theta_{0,a}$ representam o ângulo e ângulo de equilíbrio do a -ésimo termo angular do campo de força, respectivamente.

Os ângulos de equilíbrio entre esferas de diferentes resíduos foram medidos entre esferas mapeadas de 100 ns de simulações atomísticas de maltose e celobiose, cujo campo de força foi descrito como GROMOS 53A6_{GLYC}. As constantes de força foram ajustadas empiricamente para reproduzir os valores angulares médios (**Tabela 7**).

Tabela 7. Tipo de ângulo, constante de força, ângulo de equilíbrio, ângulos médios e posições relativa dos ângulos do PITOMBA e mapeadas do modelo atomístico GROMOS 53A6_{GLYC} para o resíduo de maltose e celobiose. Os erros representam o desvio padrão.

Tipo de Ângulo*	K /kJ mol ⁻¹	θ_0 /°	Posição	Ângulo CG / °	Ângulo GROMOS 53A6 _{GLYC} / °
CV1 – CV2 – CVM	0	0	Anel	70,0 ± 0,4	69,9 ± 1
CV1 – CV2 – CV6	0	0	Anel	37,8 ± 0,4	37,7 ± 2
CV1 – CVM – CV2	0	0	Anel	52,4 ± 0,3	52,5 ± 1
CV1 – CVM – CV6	0	0	Anel	54,7 ± 0,4	54,7 ± 2
CV1 – CV6 – CV2	0	0	Anel	38,0 ± 0,4	38,0 ± 2
CV1 – CV6 – CVM	0	0	Anel	76,3 ± 0,5	76,2 ± 3
CV2 – CV1 – CVM	0	0	Anel	57,5 ± 0,4	57,5 ± 1
CV2 – CV1 – CV6	0	0	Anel	104,2 ± 0,8	104,2 ± 3
CV2 – CVM – CV6	0	0	Anel	104,8 ± 0,8	104,7 ± 3
CV2 – CV6 – CVM	0	0	Anel	40,8 ± 0,4	40,8 ± 2

CVM – CV1 – CV6	0	0	Anel	49,0 ± 0,4	49,1 ± 3
CVM – CV2 – CV6	0	0	Anel	34,4 ± 0,4	34,4 ± 2
CV2 – CV1 – +CVM	850	95,7	1→4 α	96,8 ± 2	95,7 ± 4
CV6 – CV1 – +CVM	850	128,0	1→4 α	131,2 ± 3	128,0 ± 10
CV1 – +CVM – +CV1	850	147,4	1→4 α	150,4 ± 4	147,4 ± 6
CV1 – +CVM – +CV2	850	128,7	1→4 α	127,8 ± 3	128,7 ± 10
CV1 – +CVM – +CV6	850	106,7	1→4 α	107,1 ± 2	106,7 ± 10
CVM – CV1 – +CVM	850	121,4	1→4 α	120,3 ± 2	121,4 ± 5
CV2 – CV1 – +CVM	850	108,9	1→4 β	110,6 ± 2	108,9 ± 4
CV6 – CV1 – +CVM	850	142,1	1→4 β	144,5 ± 3	142,1 ± 6
CV1 – +CVM – +CV1	850	150,4	1→4 β	152,0 ± 5	150,4 ± 7
CV1 – +CVM – +CV2	850	107,7	1→4 β	108,0 ± 3	107,7 ± 6
CV1 – +CVM – +CV6	850	130,8	1→4 β	132,3 ± 4	130,8 ± 10
CVM – CV1 – +CVM	850	165,5	1→4 β	165,2 ± 3	164,5 ± 4

* O sinal “+” representa o resíduo seguinte na sequência de resíduos.

Devido ao uso de ligações em pipa, os ângulos internos do anel da glicose foram definidos como tendo constante de força zero, sendo apenas listados na **Tabela 7** por causa do programa `pdb2gmx` do GROMACS.²⁷ No entanto, estes podem ser removidos da topologia para execução das simulações.

Contribuições de Ângulos Impróprios

De acordo com o campo de força GROMOS os termos de diedro impróprio tem forma

$$V_i = \frac{1}{2} K_i (\xi_i - \xi_{0,i})^2, \quad (25)$$

em que K_i é a constante de força de ligação, ξ_i e $\xi_{0,i}$ representam o ângulo e ângulo de equilíbrio do i -ésimo termo de diedro impróprio do campo de força, respectivamente.

No modelo PITOMBA há dois tipos diferentes de diedros impróprios. O primeiro tipo é usado para descrever conformações de anel e do grupo hidroximetil (O5-C5-C6-O6), enquanto que o segundo é utilizado para descrever os anômeros da glicose. O modelo de quatro esferas gera uma dependência entre os termos diédricos impróprios e as distâncias de ligação do modelo PITOMBA, desta forma, similarmente ao protocolo adotado nos termos ligados, os ângulos de equilíbrio foram definidos a partir do *ensemble* com o hidroximetil na conformação gg. As constantes de força determinadas empiricamente para que as flutuações do modelo PITOMBA fossem semelhantes às do modelo atomístico (**Tabela 8**).

Tabela 8. Tipo de ângulo impróprio, constante de força, ângulo de equilíbrio, ângulos médios e posições relativa dos ângulos do PITOMBA e mapeadas do modelo atomístico GROMOS 53A6GLYC para o resíduo de maltose e celobiose.

Tipo de Ângulo Impróprio*	K /kJ mol ⁻¹ ° ⁻²	θ_0 /°	Posição	Ângulo CG / °	Ângulo GROMOS 53A6 _{GLYC} / °
CV1 – CVM – CV2 – CV6	600	16,4	Anel	16,4 ± 3	16,4 ± 5
CV2 – CV6 – +CVM – CV1	600	-47,2	1→4 α	-47,9 ± 2	-47,2 ± 5
CV2 – CV6 – +CVM – CV1	600	7,0	1→4 β	7,2 ± 2	7,0 ± 12

* O sinal “+” representa o resíduo seguinte na sequência de resíduos.

O campo de força GROMOS 53A6_{GLYC} é uma re-parametrização que descreve a distribuição (100%) de conformações cadeira ⁴C₁ da glicose em água a 298 K. Portanto, os valores dos ângulos impróprios descritos na **Tabela 8** estão associados à conformação ⁴C₁.

Contribuições de Ângulos Diédricos Próprios

Como descrito anteriormente, os termos diédricos foram os últimos termos definidos na parametrização. Isto ocorre não apenas porque o perfil de ângulo diédrico

depende da definição prévia de todos os outros termos, mas também porque estes termos tem a função garantir que as energias previstas pelo campo de força se ajustem as barreiras rotacionais das moléculas.

De acordo com o campo de força GROMOS, os termos de ângulos próprios têm forma

$$V_d = K_d(1 + \cos \delta_d \cdot \cos n_d \phi_d), \quad (26)$$

em que K_d é a constante de força de ligação, δ_i é o deslocamento de fase, n_d é a multiplicidade e ϕ_d é o ângulo diédrico do d -ésimo termo de diedro próprio do campo de força.

Há dois ângulos diédricos (ϕ e ψ) ao longo da ligação glicosídica do tipo 1→4 α e β (**Figura 5**). Este diedros são definidos no campo de força atomístico GROMOS, como sendo O5-C1-O4-C4 e C1-O4-C4-C3, respectivamente. No entanto, no PITOMBA um único ângulo diédrico (χ) define o perfil rotacional entre monossacarídeos adjacentes. Independentemente das esferas CG que o definam, este ângulo contém informações de ambos os diedros atomísticos, bem como dos ângulos ao longo da ligação glicosídica. Aqui, χ foi definido entre as esferas CV2 – CV1 – +CVM – +CV2, em que o sinal “+” representa o resíduo seguinte.

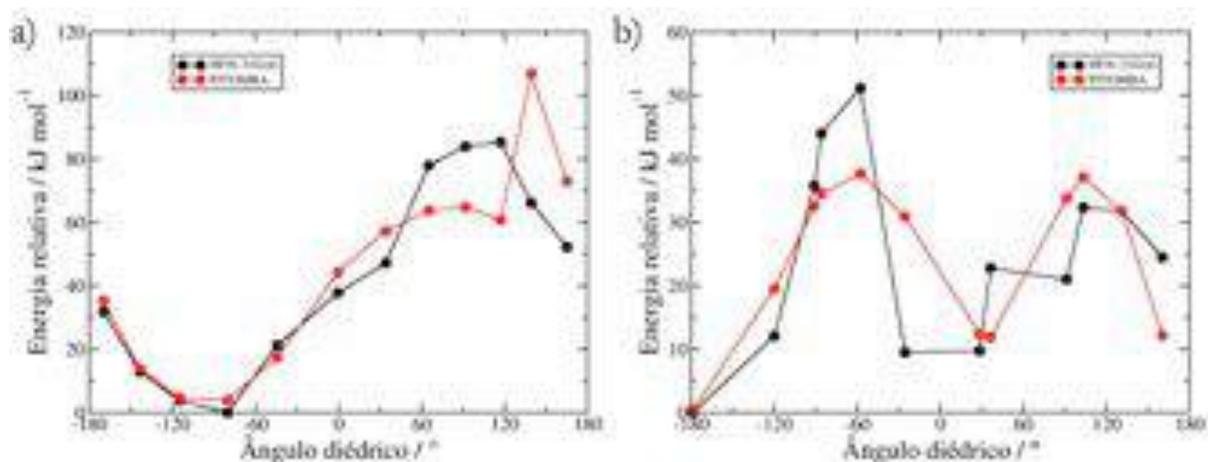


Figura 21. Perfis rotacionais dos ângulos ϕ (GROMOS 53A6_{GLYC}) e χ (CG PITOMBA) para a molécula de maltose (a) e celobiose (b). O perfil atomístico foi calculado com o nível de teoria HF/6-31G(d). Os ângulos estão definidos na representação CG.

A parametrização do perfil diedral χ seguiu o mesmo protocolo de parametrização dos campos de força GROMOS. Desta forma, os valores de energia do perfil rotacional de χ foram ajustados ao perfil gerado por química quântica (HF/6-31G(d)) das moléculas de maltose e celobiose. Os valores dos ângulos atomísticos foram mapeados para CG em cada ponto (**Figura 21**).

Tabela 9. Tipo de ângulo próprio, constante de força, deslocamento de fase, multiplicidade para ligação glicosídica em dissacarídeos presentes no PITOMBA.

Tipo de Ângulo Próprio*	K /kJ mol ⁻¹	$\delta / ^\circ$	n	Ligação glicosídica
CV2 – CV1 – +CVM – +CV2	32,4	80	1	1→4 α
CV2 – CV1 – +CVM – +CV2	6,8	0	0	1→4 β
CV2 – CV1 – +CVM – +CV2	-9,8	0	2	1→4 β
CV2 – CV1 – +CVM – +CV2	12,9	120	2	1→4 β

* O sinal “+” representa o resíduo seguinte na sequência de resíduos.

Quase todos os valores de menor energia para a superfície rotacional da maltose apresentaram ψ próximo de 96°. Desta forma, o perfil de χ foi ajustado fixando-se ψ em 96° (**Figura 21a**). Para a celobiose o perfil rotacional (**Figura 21b**) foi obtido com valor de ψ fixado em 94°. Os valores dos parâmetros ajustados do perfil diedral de χ para a maltose e celobiose estão listado na **Tabela 9**.

O perfil ajustado para a maltose no PITOMBA (**Figura 21a**) não é capaz de reproduzir exatamente o valor de energia de todos os pontos do perfil de química quântica, principalmente na região da barreira. No PITOMBA a barreira de rotação tem energia de 103 kJ mol⁻¹ e ângulo de 140°, enquanto que no perfil de referência o valor é de 85 kJ mol⁻¹, e ângulo de 118°. No entanto, o formato da curva, no geral, não é muito diferente e o mínimo de energia está devidamente ajustado. Apesar das diferenças nestes

perfis e do uso de apenas um termo de correção do ângulo diédrico, o campo de força CG PITOMBA apresenta uma flutuação do ângulo diédrico similar à maltose atomística no modelo GROMOS 53A6_{GLYC} durante 100 ns de simulação (**Figura 22a**).

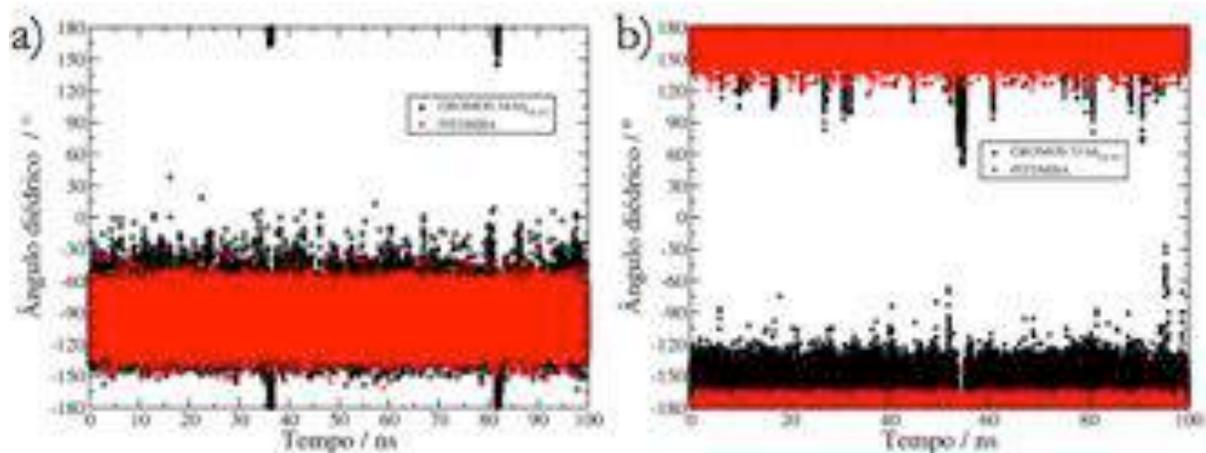


Figura 22. Valores do ângulo diédrico da ligação glicosídica da a) maltose e b) celobiose, no modelo PITOMBA, durante 100 ns de simulação CG PITOMBA e atomísticas GROMOS 53A6_{GLYC}. As moléculas atomísticas foram mapeadas para CG.

Similarmente à maltose, o perfil ajustado da celobiose (**Figura 21b**) também não é capaz de reproduzir o valor de energia de todos os pontos do perfil de referência. Mas a posição dos pontos de máxima energia das duas barreiras rotacionais (-58° e 103°) e o ponto de menor energia (-180°) são os mesmos. O perfil HF/6-31G(d) apresenta barreira de energia de 51 e 32 kJ mol⁻¹, enquanto que no PITOMBA as barreiras são 38 e 37 kJ mol⁻¹, respectivamente. Desta forma o perfil rotacional do PITOMBA é 13 kJ mol⁻¹ menor do que a barreira mais energética e 5 kJ mol⁻¹ maior do que a barreira menos energética. Apesar da existência de dois mínimos no perfil rotacional da celobiose em ambas os níveis de cálculo (HF/6-31G(d) e PITOMBA) (**Figura 21**), as simulações atomística, com o campo de força GROMOS 53A6_{GLYC}, e CG PITOMBA concentram a amostragem na região de -180° (**Figura 22b**).

Parametrização dos Termos Não-Ligados

Interações Eletrostáticas

O PITOMBA é o único modelo CG desenvolvido para carboidratos, até o momento, que inclui explicitamente interações eletrostáticas. Esta adição é feita através da inclusão de cargas pontuais às esferas CG de modo a ajustar o momento de dipolo médio do resíduo de D-glicose no campo de força GROMOS 53A6_{GLYC}. As cargas elétricas das esferas foram obtidas matematicamente das equações de momento de dipolo e de conservação de energia (**Tabela 10**). Devido à dependência do momento de dipolo do resíduo de glicose com as posições das esferas, para manter a consistência do modelo, apenas o *ensemble* com o grupo hidroximetil na conformação gg foi utilizado para ajustar as cargas do resíduo de glicose.

Tabela 10. Valores das cargas elétricas atribuídas a cada esfera CG no PITOMBA.

Esfera	Carga / e
CV1	-0,531
CV2	0,109
CVM	0,106
CV6	0,316

Embora as cargas tenham sido ajustadas apenas para parte do conjunto de configurações assumidas da glicose, a **Figura 23** mostra o momento de dipolo do resíduo de glicose nas representações PITOMBA e GROMOS 53A6_{GLYC} ao longo de todos os 100 ns simulação (inclui as conformações gg, gt e tg).

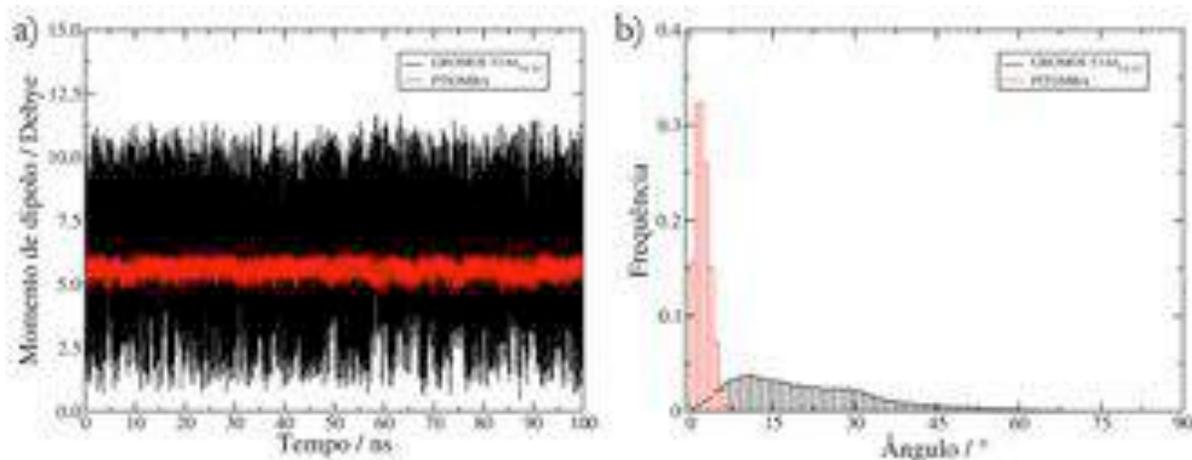


Figura 23. a) Momento de dipolo do resíduo de glicose nos modelos PITOMBA e GROMOS 53A6_{GLYC} em função do tempo ao longo de 100 ns de simulação; b) Histograma do ângulo que entre o momento de dipolo do resíduo de glicose e o seu valor médio nos modelos PITOMBA e GROMOS 53A6_{GLYC} ao longo de 100 ns de simulação.

Da **Figura 23a** percebe-se que o modelo PITOMBA tem a mesma média para a intensidade de dipolo do resíduo de glicose do modelo GROMOS 53A6_{GLYC}, mesmo para o conjunto de estruturas que apresentam o hidroximetil nas conformações gt e tg. Como esperado, as flutuações da magnitude do momento de dipolo do PITOMBA (**Figura 23a**) e as flutuações da direção do dipolo (**Figura 23b**) são bem menores do que as previstas pelo modelo GROMOS 53A6_{GLYC}, visto que a grande parte das flutuações advém da rotação dos diversos grupos hidroxila da glicose (**Figura 24**). A diferença média entre as direções do momento de dipolo do modelo PITOMBA e GROMOS 53A6_{GLYC} é em torno de 21° (**Figura 23b**).



Figura 24. Momento de dipolo do resíduo de glicose nos modelos GROMOS 53A6_{GLYC} (vermelho) e PITOMBA (azul). A posição do vetor atomístico foi deslocado para melhor visualização do vetor CG, mas ambos são sobrepostos.

Interações de van der Waals

As interações de van der Waals no campo de força GROMOS são calculadas através da função de Lennard-Jones na forma

$$V_j = \frac{C_{j,12}}{r_j^{12}} - \frac{C_{j,6}}{r_j^6}, \quad (27)$$

em que $C_{j,12}$, $C_{j,6}$ e r_j são as constantes de repulsão, constantes de atração e a distância entre as partículas na j -ésima interação de Lennard-Jones do campo de força.

Seguindo a filosofia GROMOS, os termos de van der Waals foram ajustados através de cálculos de energia livre de hidratação. Neste caso, para cada esfera CG individualmente na água SPC CG GROMOS, utilizando-se a metodologia BAR. Porque cada esfera do modelo SPC CG GROMOS representa cinco moléculas de água SPC, assumiu-se uma relação entre os termos de van der Waals de ambas as representações. Como o termo de atração é basicamente dispersão, e esta por sua vez, é basicamente aditiva, o termo C_6 foi definido como sendo a somada das dispersões dos átomos de compõe cada esfera CG multiplicado por 6,72. Isto acontece porque a soma dos termos C_6 de cinco águas SPC é 6,72 menor do que o valor do mesmo termo na água CG SPC GROMOS. Assim, o termo C_{12} passa a ser o único termo ajustável, em cada esfera CG, para ajustar a energia de hidratação. Este termo foi ajustado para reproduzir a energia de hidratação da molécula que compõe cada esfera CG, com valência completada por hidrogênio. Por exemplo, a energia de hidratação da esfera CV2 foi ajustada para representar a energia de hidratação do 1,2-etanodiol e a esfera CV1, o metanol. A **Tabela 11** mostra os valores das energias de hidratação experimental e calculadas para cada esfera CG e molécula de referência (referência da parametrização) e **Tabela 12** mostra os valores dos termos de Lennard-Jones.

Tabela 11. Energia de hidratação em água SPC CG calculada e escalonada de cada esfera CG do PITOMBA. Os valores experimentais estão associado as moléculas que compõe cada esfera CG, com valência completa por hidrogênio.

Esfera CG	Molécula de referência	Número de Coordenação	ΔG^0 Hidratação Modelo CG / kJ mol^{-1}	ΔG^0 Hidratação escalonado / kJ mol^{-1}	Valor Experimental* / kJ mol^{-1}
CV1	Metanol	2	-32,1	-20,6	-21,3
CV2	1,2-etanodiol	6	-2,4	-36,4	-32,2
CVM	Metanol	4	1,9	-21,1	-21,3
CV6	Etanol	2	-10,0	-22,4	-20,9

* Valores extraídos de ^{82,83}.

A diferença entre as energias de hidratação de uma esfera SPC CG GROMOS comparada a cinco moléculas de água SPC é de $68,9 \text{ kJ mol}^{-1}$, desta forma uma equivalência deve ser feita para comparar os valores de energia livre do modelo PITOMBA em água SPC CG GROMOS com os valores experimentais.⁷⁴ Assumiu-se que esta diferença está associada às doze ligações de hidrogênio não completas das cinco águas SPC que compõe cada esfera SPC CG GROMOS. Desta forma, as energias de hidratação de cada esfera CG foi escalonada pelo número de ligações de hidrogênio não completas por cada esfera. A contribuição de cada ligação de hidrogênio para a energia de hidratação foi tomada como sendo $68,9/12 = 5,74 \text{ kJ mol}^{-1}$.

O número de ligações de hidrogênio em falta entre cada esfera CG e a água foi calculada usando-se o primeiro pico da função de distribuição radial entre cada molécula (numa simulação atomística) e a água e arredondada para o inteiro mais próximo (**Tabela 11**).

O campo de força CG apresenta um erro no cálculo da energia de hidratação maior que o modelo GROMOS 53A6 (aproximadamente 6% vs. 0.8%, respectivamente).^{22,23} No entanto, é esperado que esta perda de precisão tenha um impacto menor nos modelos CG, pois a superfícies termodinâmicas dos modelos CG são mais suaves.

Tabela 12. Parâmetros de Lennard-Jones (C_6 e C_{12}) para cada esfera CG no PITOMBA.

Esfera CG	$C_6 / \text{kJ mol}^{-1} \text{nm}^6$	$C_{12} / \text{kJ mol}^{-1} \text{nm}^{12}$
CV1	$5,59800434 \cdot 10^{-2}$	$0,98528029 \cdot 10^{-3}$
CV2	$1,11960087 \cdot 10^{-1}$	$0,19705606 \cdot 10^{-2}$
CVM	$5,59800434 \cdot 10^{-2}$	$0,98528029 \cdot 10^{-3}$
CV6	$1,06167802 \cdot 10^{-1}$	$0,13249361 \cdot 10^{-2}$

Desempenho Computacional do Campo de Força PITOMBA

Uma das grandes características dos modelos CG é o ganho computacional quando comparados aos modelos atomísticos. O desempenho das representações CG é diretamente atrelado ao desempenho do modelo de água utilizado, pois a maior parte do tempo de simulação é gasto calculando-se as interações do solvente. O cálculo das interações do modelo de água SPC CG GROMOS é 1-2 ordens de grandeza mais rápido do que a da água SPC. O PITOMBA é entre 1-2 ordens de grandeza mais rápido do que campo de força atomístico GROMOS 53A6_{GLYC} (**Figura 25**). A **Figura 25** sugere que quanto maior o sistema, maior é o ganho computacional do modelo PITOMBA.



Figura 25. Ganho computacional do campo de força PITOMBA comparado ao campo de força GROMOS 53A6_{GLYC}, para as moléculas de glicose, γ -ciclodextrina e amilose A,

respectivamente. Os cálculos foram realizado num computador com arquitetura Intel i7, com quatro núcleos de processamento 2,9 GHz e 8 GB de memória RAM e sistema operacional OS X 10.9.1.

Validação do Campo de Força PITOMBA

O PITOMBA foi construído para reproduzir propriedades de açúcares em solução, não sendo válida a comparação dos dados do modelo em outras condições. Portanto, as propriedades estruturais e de solvatação do PITOMBA foram validadas para oligossacarídeos de glicose contendo ligações do tipo 1→4 α em solução. As análises do açúcares com ligações 1→4 β em água não serão apresentadas, pois os dados associados a estes sistemas são restritos a fase cristalina.

É importante ter em mente que os dados experimentais para carboidratos em solução são em geral escassos e impedem a comparação direta entre dados experimentais e simulados. Desta forma, alguns dados do PITOMBA foram comparados com simulações atomísticas com o campo de força GROMOS 53A6_{GLYC}.

Efeito do Passo de Integração no Campo de Força PITOMBA

As constantes de força do PITOMBA para os termos de ligação foram definidas de modo a permitir passos de integração de até 7 fs, sem o uso de qualquer potencial de restrição. Para investigar a dependência das propriedades do modelo com o tamanho do passo de integração, realizou-se três simulações da maltose com diferentes passos de integração, 2, 4 e 7 fs. Devido à substituição da função harmônica da versão modificada do GROMACS (Apêndice C), não é possível aplicar restrições nas ligações e avaliar o campo de força PITOMBA com passos de integração maiores. É importante salientar que mesmo que os modelos CG permitam valores para o passo de integração muito acima dos permitidos pelos modelo atomísticos, o passo de integração não deve ser muito grande. Pois algumas propriedades, como por exemplo, tempo de relaxação da água, influenciam na dinâmica dos sistemas. Os modelos de água atomísticas SPC e SPC/E apresentam tempo de relaxação de 5 e 7 ps, respectivamente.⁸⁴ Tais valores não foram medidos para a água SPC CG GROMOS e podem ser menores.

Do ponto de vista energético, como esperado, os valores médio de energia de cada termo do PITOMBA nos diferentes passos de integração não apresentam diferenças

significativas (**Tabela 13**). Os demais termos energéticos e propriedades do sistema, tais como, energia de interação de eletrostática, de Lennard-Jones, energia do campo de reação, pressão, temperatura, densidade também não apresentam diferenças significativas (dados não apresentados por concisão).

Tabela 13. Energia média, de 1000 passo de simulação, de cada termo do PITOMBA em função do passo de integração para diferentes simulações da maltose em solução.

Propriedade / kJ mol ⁻¹	Passo de integração		
	2 fs	4 fs	7 fs
Energia de Ligação	2,6 ± 0,1	2,4 ± 0,1	3,0 ± 0,1
Energia de Ângulo	4,9 ± 0,1	4,89 ± 0,03	4,99 ± 0,03
Energia de Diedro Próprio	1,18 ± 0,03	1,22 ± 0,03	1,17 ± 0,02
Energia de Diedro Impróprio	1,9 ± 0,1	1,82 ± 0,05	1,88 ± 0,03
Energia Potencial	-12344 ± 10	-12351 ± 8	-12354 ± 4
Energia Total	-8442 ± 11	-8449 ± 11	-8453 ± 4

Do ponto de vista geométrico, os diferentes passos de integração não promovem diferenças na estrutura da maltose. A **Figura 26** mostra o ângulo diédrico da ligação glicosídica da maltose para os primeiros 500 passos de simulação com diferentes passos de integração. Tanto para os valores dos ângulos diédrico, bem como para os demais termos do campo de força para a maltose, não há diferenças significativas nos valores médios (dados não mostrados por concisão).

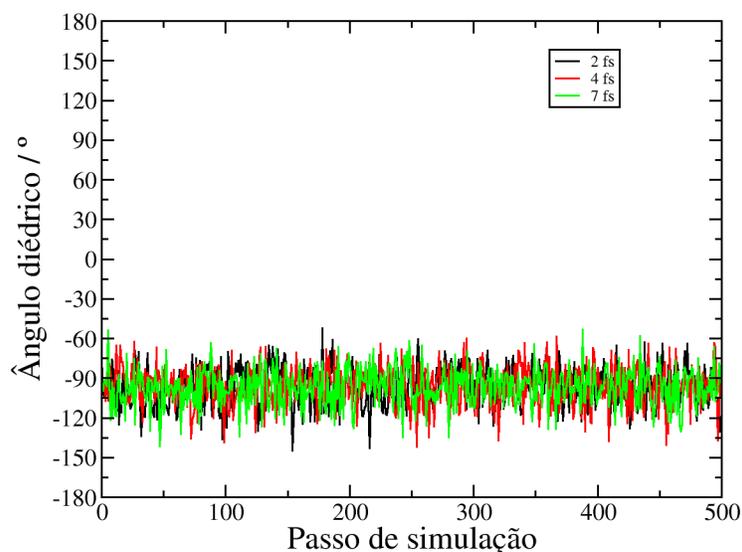


Figura 26. Valores do ângulo diédrico da ligação glicosídica da maltose, no modelo PITOMBA, durante 500 passos de simulação no CG PITOMBA.

Hélices Simples e Duplas com Ligações 1-4 α de Glicose

A fita única composta por monômeros de glicose 1 \rightarrow 4 α é denominada amilose V, sendo bastante flexível em água.⁸⁵ Co-cristalizada com diferentes compostos, tais como álcoois, DMSO e iodo revela-se como sendo uma hélice de mão esquerda contendo de seis a oito resíduos por volta.⁸⁶ No PITOMBA, o mesmo número de resíduos é observado para três diferentes comprimentos de amilose V (10, 18 e 36 monômeros) em água. A distância entre os resíduos numa volta é de ~ 20 Å, o que não é diretamente comparado com os dados experimentais de ~ 8 Å em cristais com hélices em simetria 6_5 de amilose V, em solvente não polares. Esta diferença demonstra que os dados obtidos em solução com o PITOMBA não pode ser diretamente comparado com dados cristalinos

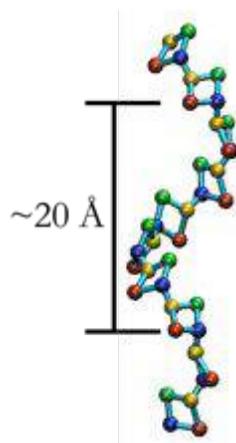


Figura 27. Representação da amilose V em um passo de um trajetória de 100 ns em solução no PITOMBA. A distância entre dois resíduos numa volta da amilose V é ~ 20 Å.

As cadeias simples em solução e em contato direto com outras cadeias de polissacarídeos exibem uma grande flexibilidade, apresentando-se no estado desenovelado. Para verificar esta hipótese, o raio de giro experimental em fase gel para uma fita contendo 10 monômeros de amilose V de 10,1 Å e 10,5 Å foi comparado com os obtidos com o PITOMBA de 11,9 Å (Figura 28).^{87,88}

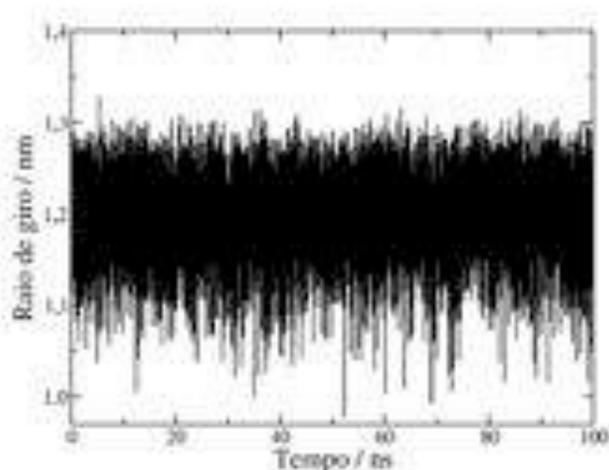


Figura 28. Raio de giro da amilose V ao longo de 100 ns de simulação no campo de força PITOMBA em solução. Valor médio de 11,9 Å.

É importante perceber que os valores para o raio de giro obtidos, em solução, com o PITOMBA se aproximam muito mais dos dados obtidos em fase gel do que os valores de distância de resíduos numa volta comparados a dados cristalográficos. Esta melhora

indica que dados previstos pelo PITOMBA tendem a ser mais parecidos com os dados experimentais em fase líquida, do que cristalina.

As fitas duplas paralelas de glicose 1→4 α são chamadas de amilose A e as fitas duplas anti-paralelas são chamadas de amilose B. Enquanto ambas apresentam-se como hélices com seis resíduos por volta, não há um consenso aparente na literatura a respeito da quiralidade das hélices (mão direita ou esquerda).⁸⁻¹³

Para a validação simulou-se uma amilose A construída com hélice de mão esquerda na qual o dado cristalográfico (clama ser de mão esquerda) mostra uma distância entre resíduos numa volta de 21,4 Å.¹¹ Para este sistema o PITOMBA prediz um uma distância média de 24,1 Å. A amilose A apresenta uma estimativa de distância muito melhor ao dado experimental do que a amilose V (fita única). Isto acontece porque na amilose A, os contatos açúcar-açúcar mantêm o oligossacarídeo no estado enovelado (**Figura 29**).

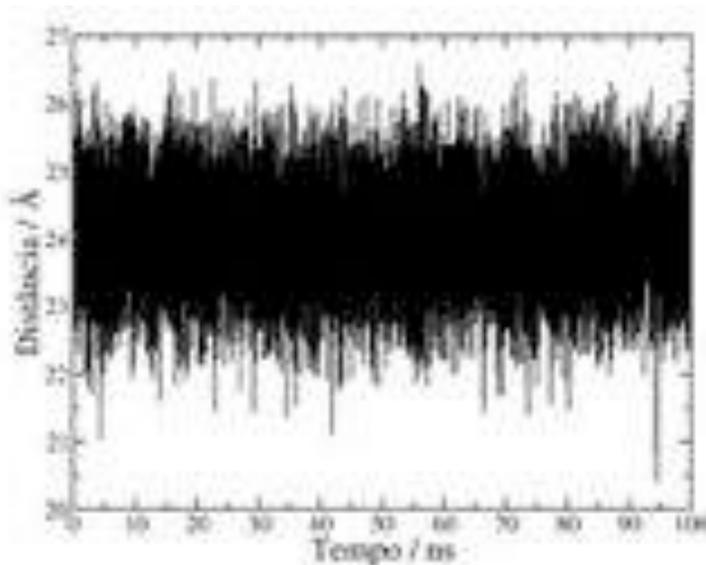


Figura 29. Distância entre resíduos numa volta de amilose A ao longo de 100 ns de simulação no campo de força PITOMBA em solução. Valor médio de 24,1 Å.

α -, β - e γ -Ciclodextrinas

As ciclodextrinas (CD) são oligossacarídeos homoquirais compostos de seis a treze monômeros de glicose ligados com ligação glicosídica do tipo 1→4 α . Dados

cristalográficos de raio-X mostram que cada monômero apresenta um arranjo de anel do tipo 4C_1 .⁸⁹ As CDs mais comumente utilizadas são compostas por seis, sete e oito monômeros são denominadas α -, β - e γ -ciclodextrinas, respectivamente (**Figura 6**).¹⁷

Uma grande quantidade de compostos podem ser inseridos parcialmente ou completamente dentro de suas cavidades. Geralmente, o tamanho das ciclodextrinas determinam os tipos de complexo formado, sendo moléculas hidrofóbicas melhor ligantes do que as hidrofílicas.¹⁸⁻²⁰ Na ausência de ligantes, as CDs são encontradas, em raio-X de cristal, hidratadas. A quantidade de moléculas de água varia com o tamanho da CD e com o experimento, por exemplo, a quantidade média de água encontrada na cavidade da β -ciclodextrina é de 6,5, mas até 12 moléculas podem ser encontradas na cavidade.^{90,91}

Acredita-se que o processo de dessolvatação de ambos, ligante e CD, são os efeitos dominantes para a formação de complexos com as β - e γ -CDs.⁹²⁻⁹⁶ No entanto, foi-se proposto que as interações polares entre a α -CD e seu ligante sejam as forças dominantes da complexação, visto que devido ao tamanho do ligante e da cavidade da α -CD ambos são fracamente solvatados.⁹²

Tabela 14. Número médio de moléculas de água dentro das cavidades das α -, β - e γ -ciclodextrinas durante 100 ns de simulação com os campos de força GROMOS 53A6_{GLYC} e PITOMBA.

Sistema	Número de moléculas de água	
	GROMOS 53A6 _{GLYC} *	PITOMBA [†]
α -ciclodextrina	5	0
β - ciclodextrina	8	1,8 (9)
γ - ciclodextrina	11	2,2 (11)

* Água SPC

[†] Água SPC CG GROMOS. Em parênteses o número correspondente de molécula de água SPC

Para avaliar o balanço das interações CD-água, o número de moléculas de água dentro das cavidades da α -, β - e γ -CDs foram monitorados durante 100 ns de simulação

nos modelos GROMOS 53A6_{GLYC} (referência) e CG PITOMBA. A **Tabela 14** e **Figura 31** mostram o número médio de moléculas de água SPC dentro das cavidades de cada ciclodextrina. O número equivalente de moléculas de água SPC dentro das cavidades no PITOMBA a razão de 5:1 de moléculas de água SPC CG GROMOS por cinco.

O número de moléculas de água dentro de cada CD foi medido da mesma maneira em todos os casos. Contou-se o número de moléculas de água que continham pelo menos um átomo dentro de um cilindro com 4 Å de altura e com raio dependente da CD. O raio do cilindro da α -, β - e γ -CD foram baseados no tamanho da cavidade na representação atomística como sendo 4,47; 5,00 e 5,47 Å, respectivamente. A **Figura 30** mostra o cilindro utilizando na contagem das águas para a γ -CD.

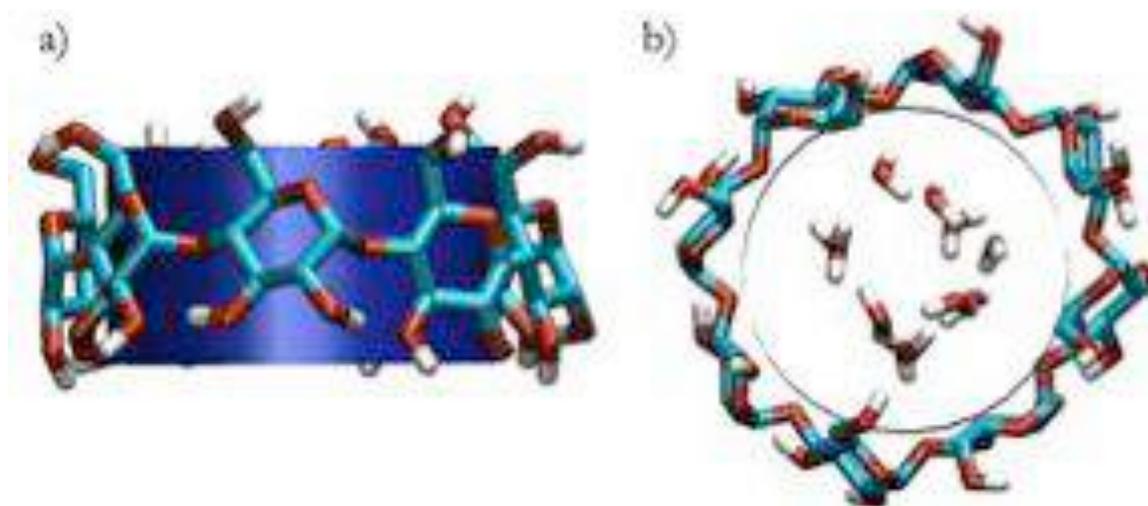


Figura 30. Representação esquemática da γ -ciclodextrina no campo de força GROMOS 53A6_{GLYC} e visões lateral (a) e de cima (b) do cilindro (azul) utilizado na contagem das moléculas de água.

A falta de moléculas de água SPC CG GROMOS dentro da α -CD pode ser explicado pelo fato de que o raio da esfera que representa as cinco moléculas de água na representação CG (formato esférico) é maior do que o raio do cilíndrico adotado pelas cinco moléculas de água SPC dentro da cavidade das CDs (**Figura 30**). No entanto, o número de moléculas de água dentro da cavidade das β - e γ -CDs concorda em ambas as representações (**Tabela 14** e **Figura 31**). O número de moléculas de água dentro das cavidades atomísticas da β - e γ -CDs é 8 e 11, respectivamente. Na representação

PITOMBA o número médio de moléculas de água SPC CG GROMOS dentro das mesmas cavidades são 1,8 e 2,2, o que corresponde a 8 e 11 moléculas de água atomísticas, respectivamente.

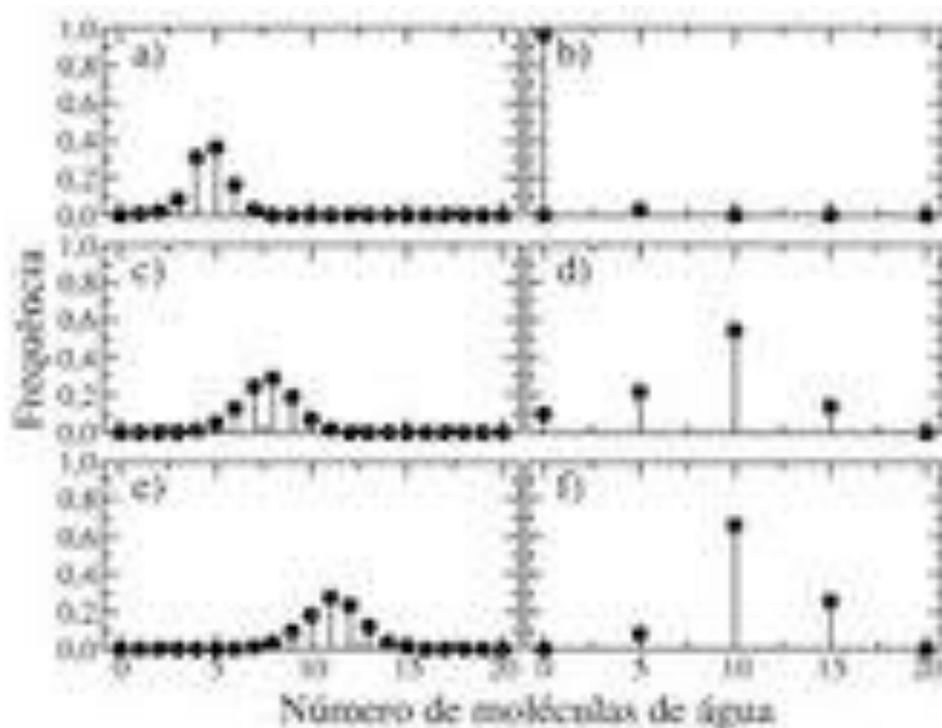


Figura 31. Histograma do número de moléculas de água dentro das cavidades atomísticas no modelo GROMOS 53A6_{GLYC} da (a) α -, (c) β - e (e) γ -CDs, e do número de moléculas de água atomísticas dentro das cavidades CG PITOMBA do (b) α -, (d) β - e (f) γ -CDs. O número de moléculas de água atomísticas no modelo CG foi obtido ao multiplicar por cinco o número de esferas dentro das cavidades.

Considerações Finais

O novo campo de força CG para carboidratos, denominado PITOMBA foi desenvolvido para simulações em fase condensada e é compatível com o modelo de água SCP CG GROMOS. O resíduo de glicose é representado por quatro esferas e um único ângulo diédrico é empregado ao longo da ligação glicosídica 1→4 α e β . A estratégia de parametrização deste ângulo diédrico seguiu o mesmo protocolo adotado nos campos de força GROMOS 45A4 e GROMOS 53A6_{GLYC}. Este último foi utilizado nas simulações atômica das molécula glicose, maltose e celobiose, utilizadas para definir os demais termos do PITOMBA. As interações eletrostáticas são calculadas explicitamente ao adicionar cargas pontuais às esferas CG e foram determinadas a partir do momento de dipolo médio do resíduo de glicose. Os termos de Lennard-Jones foram definidos a partir dos parâmetros atômicos e escalonados para ajustar-se às energias livres de hidratação experimentais.

No modelo PITOMBA todos os anéis dos açúcares são assumidos nas conformações 4C_1 , visto que esta é a conformação adotada pela glicose em solução experimental e no modelo GROMOS 53A6_{GLYC}. O modelo foi capaz de descrever propriedades estruturais de polissacarídeos como, amilose V e A e das α -, β - e γ -ciclodextrinas em solução. O número de moléculas de água dentro das cavidades das β - e γ -ciclodextrinas mostra concordância quando comparados com simulações atômicas utilizando o modelo GROMOS 53A6_{GLYC}, solvatado em água SPC, demonstrando a compatibilidade do PITOMBA com o modelo SPC CG GROMOS.

Apenas os sistemas contendo ligações 1→4 α e β foram parametrizados, sendo os primeiros os únicos validados em solução, na forma de maltose, amilose V, A e B e α -, β - e γ -ciclodextrinas. O PITOMBA foi construído para ser fácil de expandir e incluir outros tipos de ligação glicosídica, tais como 1→1 α e β , 1→6 α e β , 1→2 α e β e 1→3 α e β .

Perspectivas

O campo de força PITOMBA foi desenvolvido para simulações em fase condensada com a água SPC CG GROMOS e foi testado para alguns sistemas com ligação glicosídica do tipo 1→4 α e β . O modelo de parametrização torna fácil expandir e incluir outros tipos de ligação glicosídica, tais como 1→1 α e β , 1→6 α e β , 1→2 α e β e 1→3 α e β . No entanto, o PITOMBA pode avançar em outras direções, tais como incluir outras esferas associadas a derivados de açúcares, tais como quitina e quitosana; incluir outros tipos de biomoléculas, tais como lipídeos. E estudar a termodinâmica do PITOMBA em água SPC. Pode-se ainda verificar a viabilidade do modelo em sistemas híbridos, pois é possível numa mesma simulação ter molécula de água expressas como no modelo SPC e no modelo SPC CG GROMOS.

Do ponto de vista das ferramentas associadas ao PITOMBA é necessário expandir o programa `g_map` para fazer o particionamento dos grupos do arquivo de índices do GROMACS (*index file*) em função de ligações e moléculas e expandir o mapeamento para CG para, virtualmente, qualquer tipo de sistema. Além de implementar o mapeamento inverso, ou seja, CG para atomístico.

Referências Bibliográficas

Rever 25, 35

- (1) Marradi, M.; Martín-Lomas, M.; Penadés, S. Glyconanoparticles: Polyvalent Tools to Study Carbohydrate-Based Interactions. In *Advances in Carbohydrate Chemistry and Biochemistry*; Horton, D., Ed.; Advances in Carbohydrate Chemistry and Biochemistry; Academic Press, 2010; Vol. Volume 64, pp. 211–290.
- (2) Hollingsworth, R. I.; Wang, G. Toward a Carbohydrate-Based Chemistry: Progress in the Development of General-Purpose Chiral Synthons From Carbohydrates. *Chem. Rev.* **2000**, *100*, 4267–4282.
- (3) Dwek, R. A.; Butters, T. D. Introduction: Glycobiology Understanding the Language and Meaning of Carbohydrates. *Chem. Rev.* **2002**, *102*, 283–284.
- (4) McReynolds, K. D.; Gervay-Hague, J. Chemotherapeutic Interventions Targeting HIV Interactions with Host-Associated Carbohydrates. *Chem. Rev.* **2007**, *107*, 1533–1552.
- (5) Murrey, H. E.; Hsieh-Wilson, L. C. The Chemical Neurobiology of Carbohydrates. *Chem. Rev.* **2008**, *108*, 1708–1731.
- (6) McNaught, A. D. Nomenclature of Carbohydrates. *Pure Appl. Chem.* **1996**, *68*, 1919–2008.
- (7) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*, Fourth Edition; Freeman, Ed.; Fourth Edition. 2004.
- (8) Imberty, A.; Pérez, S. A Revisit to the Three-Dimensional Structure of B-Type Starch. *Biopolymers* **1988**, *27*, 1205–1221.
- (9) Hinrichs, W.; Saenger, W. Crystal and Molecular Structure of the Hexasaccharide Complex (P-Nitrophenyl .Alpha.-Maltohexaoside)₂.Cntdot.Ba(I3)2.Cntdot.27H2O. *J. Am. Chem. Soc.* **2013**, *112*, 2789–2796.
- (10) Hsein-Chih, H. W.; Sarko, A. The Double-Helical Molecular Structure of Crystalline B-Amylose. *Carbohydrate Research* **1978**, *61*, 7–25.
- (11) Imberty, A.; Chanzy, H.; Pérez, S.; Bulèon, A.; Tran, V. The Double-Helical Nature of the Crystalline Part of a-Starch. *J Mol Biol* **1988**, *201*, 365–378.
- (12) Hsien-Chih, H. W.; Sarko, A. The Double-Helical Molecular Structure of Crystalline a-Amylose. *Carbohydrate Research* **1978**, *61*, 27–40.
- (13) Wu, H.-C. H.; Sarko, A. The Crystal Structure of a-Starch: Is It Double Helical? *Carbohydrate Research* **1977**, *54*, C3–C6.
- (14) Rappenecker, G.; Zugenmaier, P. Detailed Refinement of the Crystal Structure of Vh-Amylose. *Carbohydrate Research* **1981**, *89*, 11–19.
- (15) Brisson, J.; Chanzy, H.; Winter, W. T. The Crystal and Molecular Structure of VH Amylose by Electron Diffraction Analysis. *International Journal of Biological Macromolecules* **1991**, *13*, 31–39.
- (16) Zaslow, B.; Miller, R. L. Hydration of the “v” Amylose Helix 1. *J. Am. Chem. Soc.* **1961**, *83*, 4378–4381.
- (17) Venkatesh, G.; Thulasidhasan, J.; Rajendiran, N. A Spectroscopic and Molecular

- Modeling Studies of the Inclusion Complexes of Orciprenaline and Terbutaline Drugs with Native and Modified Cyclodextrins. *J Incl Phenom Macrocycl Chem* **2014**, *78*, 225–237.
- (18) Zhang, J.; Ma, P. X. Cyclodextrin-Based Supramolecular Systems for Drug Delivery: Recent Progress and Future Perspective. *Polysaccharide-based systems in drug and gene delivery* **2013**, *65*, 1215–1233.
- (19) Kahle, C.; Holzgrabe, U. Determination of Binding Constants of Cyclodextrin Inclusion Complexes with Amino Acids and Dipeptides by Potentiometric Titration. *Chirality* **2004**, *16*, 509–515.
- (20) Kolhapurkar, R.; Patil, K. Studies of Volumetric and Activity Behaviors of Binary and Ternary Aqueous Solutions Containing B-Cyclodextrin and Glucose. *Journal of Molecular Liquids* **2013**, *178*, 185–191.
- (21) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 1987.
- (22) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase. *J. Comput. Chem.* **2001**, *22*, 1205–1218.
- (23) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: the GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
- (24) Verlet, L. Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98–103.
- (25) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *J. Chem. Phys.* **1982**, *76*, 637–649.
- (26) Bennett, C. H. Efficient Estimation of Free Energy Differences From Monte Carlo Data. *Journal of Computational Physics* *22*, 245–268.
- (27) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (28) Stortz, C. A.; Johnson, G. P.; French, A. D.; Csonka, G. I. Comparison of Different Force Fields for the Study of Disaccharides. *Carbohydrate Research* **2009**, *344*, 2217–2228.
- (29) Lins, R. D.; Hünenberger, P. H. A New GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Comput. Chem.* **2005**, *26*, 1400–1412.
- (30) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: a Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29*, 622–655.
- (31) Allinger, N. L.; Yuh, Y. H.; Li, J. H. Molecular Mechanics. the MM3 Force Field for Hydrocarbons. 1. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
- (32) Duus, J. Ø.; Gotfredsen, C. H.; Bock, K. Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations. *Chem. Rev.* **2000**, *100*, 4589–4614.
- (33) Mechref, Y.; Novotny, M. V. Structural Investigations of Glycoconjugates at High Sensitivity. *Chem. Rev.* **2002**, *102*, 321–370.

- (34) Autieri, E.; Segal, M.; Pederiva, F.; Guella, G. Puckering Free Energy of Pyranoses: a NMR and Metadynamics-Umbrella Sampling Investigation. *J. Chem. Phys.* **2010**, *133*, 095104–095113.
- (35) Barnett, C. B.; Naidoo, K. J. Ring Puckering: a Metric for Evaluating the Accuracy of AM1, PM3, PM3CARB-1, and SCC-DFTB Carbohydrate QM/MM Simulations. *J. Phys. Chem. B* **2010**, *114*, 17142–17154.
- (36) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23*, 69–89.
- (37) Hehre, W. J. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- (38) Pol-Fachin, L.; Rusu, V. H.; Verli, H.; Lins, R. D. GROMOS 53A6GLYC, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 4681–4690.
- (39) Nishida, Y.; Ohruai, H.; Meguro, H. ¹H-NMR Studies of (6r)- and (6s)-Deuterated D-Hexoses: Assignment of the Preferred Rotamers About C5-C6 Bond of D-Glucose and D-Galactose Derivatives in Solutions. *Tetrahedron Letters* **2005**, *25*, 1575–1578.
- (40) Ohruai, H.; Nishida, Y.; Higuchi, H.; Hori, H.; Meguro, H. The Preferred Rotamer About the C5-C6 bond of D-Galactopyranoses and the Stereochemistry of Dehydrogenation by D-Galactose Oxidase. *Can. J. Chem.* **1987**, *65*, 1145–1153.
- (41) Nishida, Y.; Hori, H.; Ohruai, H.; Meguro, H. ¹H NMR Analyses of Rotameric Distribution of C5-C6 Bonds of D-Glucopyranoses in Solution. *Journal of Carbohydrate Chemistry* **1988**, *7*, 239–250.
- (42) Brochier-Salon, M.-C.; Morin, C. Conformational Analysis of 6-Deoxy-6-Iodo-D-Glucose in Aqueous Solution. *Magn. Reson. Chem.* **2000**, *38*, 1041–1042.
- (43) Thibaudeau, C.; Stenutz, R.; Hertz, B.; Klepach, T.; Zhao, S.; Wu, Q.; Carmichael, I.; Serianni, A. S. Correlated C–C and C–O Bond Conformations in Saccharide Hydroxymethyl Groups: Parametrization and Application of Redundant ¹H–¹H, ¹³C–¹H, and ¹³C–¹³C NMR J-Couplings. *J. Am. Chem. Soc.* **2004**, *126*, 15668–15685.
- (44) Spiwok, V.; Králová, B.; Tvaroška, I. Modelling of B-D-Glucopyranose Ring Distortion in Different Force Fields: a Metadynamics Study. *Carbohydrate Research* **2010**, *345*, 530–537.
- (45) Ercolessi, F.; Adams, J. B. Interatomic Potentials From First-Principles Calculations: the Force-Matching Method. *EPL (Europhysics Letters)* **1994**, *26*, 583–588.
- (46) Lyubartsev, A. P.; Laaksonen, A. Calculation of Effective Interaction Potentials From Radial Distribution Functions: a Reverse Monte Carlo Approach. *Phys. Rev. E* **1995**, *52*, 3730–3737.
- (47) Moritsugu, K.; Smith, J. C. Coarse-Grained Biomolecular Simulation with REACH: Realistic Extension Algorithm via Covariance Hessian. *Biophysical Journal* **2007**, *93*, 3460–3469.
- (48) Moritsugu, K.; Smith, J. C. REACH Coarse-Grained Biomolecular Simulation: Transferability Between Different Protein Structural Classes. *Biophysical Journal* **2008**, *95*, 1639–1648.
- (49) Moritsugu, K.; Kurkal-Siebert, V.; Smith, J. C. REACH Coarse-Grained Normal

- Mode Analysis of Protein Dimer Interaction Dynamics. *Biophysical Journal* **2009**, *97*, 1158–1167.
- (50) Molinero, V.; Goddard, W. A. M3B: a Coarse Grain Force Field for Molecular Simulations of Malto-Oligosaccharides and Their Water Mixtures. *J. Phys. Chem. B* **2004**, *108*, 1414–1427.
- (51) Liu, P.; Izvekov, S.; Voth, G. A. Multiscale Coarse-Graining of Monosaccharides. *J. Phys. Chem. B* **2007**, *111*, 11566–11575.
- (52) Bu, L.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. The Energy Landscape for the Interaction of the Family 1 Carbohydrate-Binding Module and the Cellulose Surface Is Altered by Hydrolyzed Glycosidic Bonds. *J. Phys. Chem. B* **2009**, *113*, 10994–11002.
- (53) López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210.
- (54) Wohler, J.; Berglund, L. A. A Coarse-Grained Model for Molecular Dynamics Simulations of Native Cellulose. *J. Chem. Theory Comput.* **2011**, *7*, 753–760.
- (55) Hynninen, A.-P.; Matthews, J. F.; Beckham, G. T.; Crowley, M. F.; Nimlos, M. R. Coarse-Grain Model for Glucose, Cellobiose, and Cellotetraose in Water. *J. Chem. Theory Comput.* **2011**, *7*, 2137–2150.
- (56) Srinivas, G.; Cheng, X.; Smith, J. C. A Solvent-Free Coarse Grain Model for Crystalline and Amorphous Cellulose Fibrils. *J. Chem. Theory Comput.* **2011**, *7*, 2539–2548.
- (57) Glass, D. C.; Moritsugu, K.; Cheng, X.; Smith, J. C. REACH Coarse-Grained Simulation of a Cellulose Fiber. *Biomacromolecules* **2012**, *13*, 2634–2644.
- (58) Bellesia, G.; Chundawat, S. P. S.; Langan, P.; Redondo, A.; Dale, B. E.; Gnanakaran, S. Coarse-Grained Model for the Interconversion Between Native and Liquid Ammonia-Treated Crystalline Cellulose. *J. Phys. Chem. B* **2012**, *116*, 8031–8037.
- (59) Markutsya, S.; Kholod, Y. A.; Devarajan, A.; Windus, T. L.; Gordon, M. S.; Lamm, M. H. A Coarse-Grained Model for B-D-Glucose Based on Force Matching. *Theor Chem Acc* **2012**, *131*, 1162–1176.
- (60) Markutsya, S.; Devarajan, A.; Baluyut, J. Y.; Windus, T. L.; Gordon, M. S.; Lamm, M. H. Evaluation of Coarse-Grained Mapping Schemes for Polysaccharide Chains in Cellulose. *J. Chem. Phys.* **2013**, *138*, 214108–214117.
- (61) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: a Generic Force Field for Molecular Simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.
- (62) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. Effective Force Fields for Condensed Phase Systems From Ab Initio Molecular Dynamics Simulation: a New Method for Force-Matching. *J. Chem. Phys.* **2004**, *120*, 10896–10912.
- (63) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (64) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (65) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: a Program for Macromolecular Energy, Minimization,

- and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (66) Kuttel, M.; Brady, J. W.; Naidoo, K. J. Carbohydrate Solution Simulations: Producing a Force Field with Experimentally Consistent Primary Alcohol Rotational Frequencies and Populations. *J. Comput. Chem.* **2002**, *23*, 1236–1243.
- (67) MacKerell, A. D.; Bashford, D.; Bellott; 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. K.; 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. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (68) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (69) Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal Structure and Hydrogen-Bonding System in Cellulose I β From Synchrotron X-Ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082.
- (70) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. Crystal Structure and Hydrogen Bonding System in Cellulose I A from Synchrotron X-Ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2003**, *125*, 14300–14306.
- (71) Matthews, J. F.; Skopec, C. E.; Mason, P. E.; Zuccato, P.; Torget, R. W.; Sugiyama, J.; Himmel, M. E.; Brady, J. W. Computer Simulation Studies of Microcrystalline Cellulose I β . *Carbohydrate Research* **341**, 138–152.
- (72) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2013**, *111*, 7812–7824.
- (73) López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9*, 1694–1708.
- (74) Riniker, S.; van Gunsteren, W. F. A Simple, Efficient Polarizable Coarse-Grained Water Model for Molecular Dynamics Simulations. *J. Chem. Phys.* **2011**, *134*, 084110.
- (75) Berendsen, H.; Postma, J.; van Gunsteren, W. F.; Hermans, J. Intermolecular Forces. *Pullman, B., Ed.; Reidel Publishing Company: Dordrecht* **1981**, 331–342.
- (76) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101–014107.
- (77) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (78) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: a Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (79) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A Generalized Reaction Field Method for Molecular Dynamics Simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (80) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding Singularities and Numerical Instabilities in Free Energy Calculations

- Based on Molecular Simulations. *Chemical Physics Letters* **222**, 529–539.
- (81) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09 Revision D.01.
- (82) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Universal Quantum Mechanical Model for Solvation Free Energies Based on Gas-Phase Geometries. *J. Phys. Chem. B* **1998**, *102*, 3257–3271.
- (83) Gonçalves, P. F. B.; Stassen, H. Calculation of the Free Energy of Solvation From Molecular Dynamics Simulations. *Pure Appl. Chem.* **2004**, *76*, 231–240.
- (84) van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C. A Systematic Study of Water Models for Molecular Simulation: Derivation of Water Models Optimized for Use with a Reaction Field. *J. Chem. Phys.* **1998**, *108*, 10220–10230.
- (85) Mark, H. X-Ray Investigation of Carbohydrates. *Chem. Rev.* **2013**, *26*, 169–186.
- (86) Le Bail, P.; Bizot, H.; Pontoire, B.; Buléon, A. Polymorphic Transitions of Amylose-Ethanol Crystalline Complexes Induced by Moisture Exchanges. *Starch/Stärke* **1995**, *47*, 229–232.
- (87) Müller, J. J.; Gernat, C.; Schulz, W.; Müller, E.-C.; Vorwerg, W.; Damaschun, G. Computer Simulations of X-Ray Scattering Curves: Gelation and Crystallization Process in Amylose Solutions. *Biopolymers* **1995**, *35*, 271–288.
- (88) Leloup, V. M.; Colonna, P.; Ring, S. G.; Roberts, K.; Wells, B. Microstructure of Amylose Gels. *Carbohydrate Polymers* **1992**, *18*, 189–197.
- (89) Kurkov, S. V.; Loftsson, T. Cyclodextrins. *Poorly Soluble Drugs* **2013**, *453*, 167–180.
- (90) Lindner, K.; Saenger, W. β -Cyclodextrin Dodecahydrate: Crowding of Water Molecules Within a Hydrophobic Cavity. *Angew. Chem. Int. Ed. Engl.* **1978**, *17*, 694–695.
- (91) Ponce Cevallos, P. A.; Buera, M. P.; Elizalde, B. E. Encapsulation of Cinnamon and Thyme Essential Oils Components (Cinnamaldehyde and Thymol) in β -Cyclodextrin: Effect of Interactions with Water on Complex Stability. *Journal of Food Engineering* **2010**, *99*, 70–75.
- (92) Song, L. X.; Bai, L.; Xu, X. M.; He, J.; Pan, S. Z. Inclusion Complexation, Encapsulation Interaction and Inclusion Number in Cyclodextrin Chemistry. *Coordination Chemistry Reviews* **2009**, *253*, 1276–1284.
- (93) Gelb, R. I.; Schwartz, L. M.; Cardelino, B.; Fuhrman, H. S.; Johnson, R. F.; Laufer, D. A. Binding Mechanisms in Cyclohexaamylose Complexes. *J. Am. Chem. Soc.* **2013**, *103*, 1750–1757.

- (94) Inoue, Y.; Hakushi, T.; Liu, Y.; Tong, L.; Shen, B.; Jin, D. Thermodynamics of Molecular Recognition by Cyclodextrins. 1. Calorimetric Titration of Inclusion Complexation of Naphthalenesulfonates with .Alpha.-, .Beta.-, and .Gamma.-Cyclodextrins: Enthalpy-Entropy Compensation. *J. Am. Chem. Soc.* **2013**, *115*, 475–481.
- (95) Rekharsky, M. V.; Goldberg, R. N.; Schwarz, F. P.; Tewari, Y. B.; Ross, P. D.; Yamashoji, Y.; Inoue, Y. Thermodynamic and Nuclear Magnetic Resonance Study of the Interactions of .Alpha.- and .Beta.-Cyclodextrin with Model Substances: Phenethylamine, Ephedrines, and Related Substances. *J. Am. Chem. Soc.* **2013**, *117*, 8830–8840.
- (96) Liu, Y.; Cao, R.; Chen, Y.; He, J.-Y. Effect of B-Cyclodextrin Charge Type on the Molecular Recognition Thermodynamics of Reactions with (Ferrocenylmethyl)Dimethylaminium Derivatives. *J. Phys. Chem. B* **2013**, *112*, 1445–1450.
- (97) Rusu, V. H.; Horta, V. A. C.; Horta, B. A. C.; Lins, R. D.; Baron, R. MDWiZ: a Platform for the Automated Translation of Molecular Dynamics Simulations. *Journal of Molecular Graphics and Modelling* **2014**, *48*, 80–86.

Apêndice A

Um conjunto de ferramentas foi desenvolvida ao longo do processo de criação do campo de força PITOMBA. As ferramentas são gerais o suficiente para que possam ser utilizadas para expandir o campo de força ou ser aplicada em outros campos de força CG. Estas ferramentas foram escritas em linguagem C como ferramentas de contribuição do GROMACS, ou seja, elas fazem uso das bibliotecas do GROMACS e podem ser utilizadas em conjunto com os arquivos compatível do referido programa. Instalação, uso e ajuda seguem o padrão de ferramentas GROMACS. Todos os programas estão disponíveis gratuitamente em: <http://dqfnet.ufpe.br/biomat/>. Algumas ferramentas são disponibilizadas para *download* e serem utilizadas localmente e outras são de uso remoto acessíveis apenas por interfaces *web*.

A.1. Programas Disponíveis para *Download*

Dentre as ferramentas desenvolvidas disponíveis para *download*, destaca-se `g_map`, que faz o mapeamento, por centro de massa, de trajetórias na representação atômica para CG e o `g_cg` que ajusta analiticamente as cargas das esferas CG para reproduzir o modelo de dipolo de referência.

A.1.1. `g_map`

O `g_map` mapeia por centro de massa de grupos de átomos em esferas CG. É necessário fornecer alguns arquivos e definir alguns parâmetros para o mapeamento. Os arquivos e opções de entrada do programa estão listados abaixo, no modelo GROMACS.

Option	Filename	Type	Description
-f	traj.xtc	Input	Trajectory: xtc trr trj gro g96 pdb cpt
-n	index.ndx	Input	Index file
-s	topol.tpr	Input	Run input file: tpr tpb tpa
-o	traj.xtc	Output	Trajectory: xtc trr trj gro g96 pdb cpt

Option	Type	Value	Description
--------	------	-------	-------------

```

-[no]h      bool    no      Print help info and quit
-[no]version bool    no      Print version info and quit
-nice       int     0       Set the nicelevel
-b          time    0       First frame (ps) to read from trajectory
-e          time    0       Last frame (ps) to read from trajectory
-dt         time    0       Only use frame when t MOD dt = first time (ps)
-[no]pbc    bool    yes     Takes into account the pbc in calculations
-box        vector -1 -1 -1   fixes the box size to specific values

```

O programa é flexível o suficiente para utilizar ou desprezar condições periódicas de contorno (-[no]pbc), filtrar alguns passos da trajetória (-dt), modificar o tamanho da caixa de simulação (-box), escolher o intervalo temporal do mapeamento (-b e -e) e determinar a prioridade no uso dos recursos do sistema então este estiver rodando (-nice).

A quantidade de grupos a serem mapeados e o nome de cada esfera CG são definidos no arquivo de índices. O programa mapeia todos os grupos do arquivo de índice em esfera CG e nomeia a esfera CG de acordo com o nome do grupo de origem (quatro primeiras letras se existirem).

A.1.2. g_cg dipolefitter

O g_cg dipolefitter ajustam analiticamente as cargas das esferas CG para reproduzir o modelo de dipolo de referência. É necessário fornecer alguns arquivos e definir alguns parâmetros para o mapeamento. Os arquivos e opções de entrada do programa estão listados abaixo, no modelo GROMACS.

Option	Filename	Type	Description
-f	traj.xtc	Input	Trajectory: xtc ttr trj gro g96 pdb cpt
-s	topol.tpr	Input	Run input file: tpr tpb tpa
-d	dipole.dat	Input	Generic data file
-n	index.ndx	Input, Opt.	Index file
-o	charge.xvg	Output	xvgr/xmgr file

Option	Type	Value	Description
-[no]h	bool	yes	Print help info and quit
-[no]version	bool	no	Print version info and quit
-nice	int	0	Set the nicelevel
-b	time	0	First frame (ps) to read from trajectory
-e	time	0	Last frame (ps) to read from trajectory
-dt	time	0	Only use frame when t MOD dt = first time (ps)
-xvg	enum	xmgrace	xvg plot formatting: xmgrace, xmgr or none

```
-[no]pbc      bool   yes      Takes into account the pbc in calculations
-box         vector -1 -1 -1   fixes the box size to specific values
```

O programa é flexível o suficiente para utilizar ou desprezar condições periódicas de contorno (-[no]pbc), filtrar alguns passos da trajetória (-dt), modificar o tamanho da caixa de simulação (-box), escolher o intervalo temporal do mapeamento (-b e -e), determinar a prioridade no uso dos recursos do sistema então este estiver rodando (-nice) e escolher o formato do arquivo de saída contendo as cargas das esferas CG (-xvg).

Devido à limitação imposta pela solução das equações de dipolo e conservação de carga, o `g_cg`dipolefitter requer como entrada quatro grupos no arquivo de índices que representem o mapeamento CG. O arquivo `dipole.dat`, (opção -d) contém a lista de dipolos a serem ajustados pelo programa e o arquivo de topologia (opção -s) contém a topologia CG.

A.2. Programas Disponíveis por Interface *Web*

Dentre as ferramentas desenvolvidas como aplicações *web*, destaca-se o Rotational Profiler, que ajusta perfis de torção de mecânica molecular a perfis de química quântica e o MDWiZ que converte arquivos de entrada de programas de simulação.

A.2.1. Rotational Profiler

O Rotational Profiler (<http://dqfnet.ufpe.br/biomat/rotprof>), foi desenvolvido e usado para determinar as constantes de força dos perfis de mecânica molecular atomísticos e CG que melhor ajustam aos perfis de química quântica. Esta ferramenta foi utilizada tanto na parametrização do modelo PITOMBA bem como do campo de força atomístico GROMOS 53A6_{GLYC}.³⁸ Os termos torcionais são definidos pelas equações de torção do campo de força GROMOS, mas o ajuste pode ser aplicado a qualquer perfil de torção que seja caso particular do GROMOS, como nos campos de força AMBER e OPLS.⁹⁷

O Rotational Profiler requer dois arquivos de entrada contendo os ângulos, em graus, e energias (relativas) de cada perfil. O primeiro arquivo descreve o perfil de torção

de referência (por exemplo, química quântica) e o segundo o perfil a ser ajustado (por exemplo, mecânica molecular). O usuário é responsável por determinar a quantidade e valores dos parâmetros ajustáveis dos cossenos (**Equação 26**). A interface permite ajustar até sete cossenos simultaneamente e alterar os parâmetros de multiplicidade e deslocamento de fase, além de contar com uma ferramenta gráfica que plota os perfis de referência, a ser ajustado e os ajustados.

A.2.2. MDWiZ

O MDWiZ (<http://barongroup.medchem.utah.edu/tools/mdwiz/>) visa auxiliar a transposição de arquivos de entrada entre programas de simulação.⁹⁷ Atualmente o requer um arquivo de parâmetro do GROMACS (tpr) para fazer a conversão. Sendo, atualmente, capaz de converter o arquivo tpr de entrada do GROMACS para o LAMMPS, nos campos de força AMBER, GROMOS e OPLS e modelos de água SPC, SPC/E, TIP3P e TIP4P. A adição de outros campos de força e programas está sendo realizada para expandir as suas capacidades.

Apêndice B

Este apêndice tem como objetivo apresentar o modelo de água SPC CG GROMOS utilizado no modelo PITOMBA, apresentar as condições de simulação deste modelo, comprar algumas propriedades da água para este modelo com alguns modelos de água CG e com dados experimentais. Além de avaliar estas mesmas propriedades na implementação no GROMACS 4.6.x.²⁷

Modelo de Água SPC CG GROMOS

O modelo de água SPC CG GROMOS foi desenvolvido pela Sereina Riniker no grupo de Wilfred van Gunsteren e apresenta grande concordância com resultados experimentais para algumas propriedades macroscópicas da água em fase líquida a 298 K e 1 atm Tabela III da Referência ⁷⁴. O modelo de água é composto de uma esfera CG de massa de 90,077 com dois sítios de interação representando cinco moléculas de água (**Figura 32**) e é baseado no modelo de água SPC.

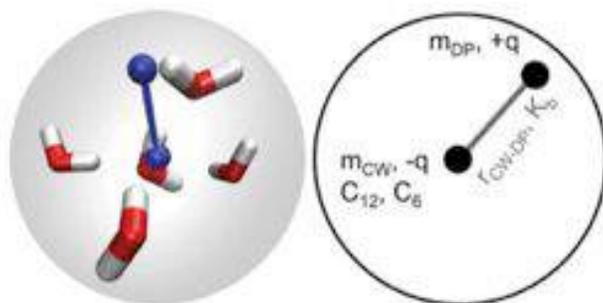


Figura 32. Modelo de água SPC CG GROMOS. Figura adaptada de Riniker, S.; van Gunsteren, W. F. *J. Chem. Phys.* **2011**, *134*, 084110, Copyright (2011) com permissão da AIP Publishing LLC.

Os dois sítios de interação elétrica são denominados CW (“*Center of van der Waals*”) e DP (“*Dummy Particle*”) e são ligados por uma função quártica (**Equação 23**) com constante de força de $2,0 \cdot 10^6 \text{ kJ mol}^{-1} \text{ nm}^{-4}$ e distância de equilíbrio de 2 Å. O sítio CW tem carga $-0,575 \text{ e}$, massa de 60,077 u.m.a. e parâmetros de van der Waals $C_6 = 0,088 \text{ kJ mol}^{-1} \text{ nm}^6$ e $C_{12} = 1,298 \cdot 10^{-3} \text{ kJ mol}^{-1} \text{ nm}^{12}$, enquanto que o sítio DP apresenta massa de 30,0 u.m.a e carga de $+0,575 \text{ e}$. Devido a presença dos dois sítios carregados o

modelo pode reagir a campos elétricos rotacionando ao redor do sítio CW e ajustando a distância da ligação CW–DP.

Validade do Modelo

Todo modelo apresenta um domínio de aplicação, neste caso o modelo foi testado para diferentes temperaturas na faixa de 258 K a 338 K sob condições N,p,T. Em 258 K (-15 °C) o modelo apresenta sinais de congelamento. A **Tabela 15** mostra as condições de simulação do modelo de água SPC CG GROMOS.⁷⁴

Tabela 15. Parâmetros requeridos pelo modelo de água SPC CG GROMOS.

Parâmetro	Valor
Modelo de correção de longa distância para as interações de van der Waals	Raio de corte
Raio de corte para interações de longa distância e van der Waals	20 Å
Modelo de correção de longa distância para as interações elétricas	Campo de reação
Permissividade relativa do campo de reação (ϵ_r)	78,5
Permissividade relativa dentro do raio de corte do campo de reação (ϵ_{cs})	2,5
Constante de compressibilidade de água	$7,51 \cdot 10^{-4} / (\text{kJ mol}^{-1} \text{ nm}^{-3})^{-1}$

Propriedades do Modelo Implementado

Devido à necessidade de implementação do termo quártico de ligação requerido pelo modelo de água SPC CG GROMOS no programa GROMACS 4.6.x, as propriedades do modelo de água foram revalidadas. E para tal, as propriedades de densidade, energia potencial por esfera CG e difusão das partículas foram escolhidas. A **Tabela 16** mostra os resultados destas propriedades apresentados por Riniker et. al. além dos obtidos após 100 ns de uma trajetória contendo 2500 moléculas de água CG.

Tabela 16. Densidade (ρ), energia potencial por esfera CG (V_{pot}) e coeficiente de difusão (D), em 298 K, para modelo SPC CG GROMOS implementado nesta tese, em comparação com diferentes modelos de água CG e com dados experimentais. Tabela adaptada de Riniker, S.; van Gunsteren, W. F. *J. Chem. Phys.* **2011**, *134*, 084110, Copyright (2011) com permissão da AIP Publishing LLC.

	ρ / g cm ⁻³	V_{pot} / kJ mol ⁻¹	D / 10 ⁻⁵ cm ² s ⁻¹
Experimental [citar referências]	0.997	-41.5	2.3
SPC [citar referências]	0.972	-41.3	4.2
SPC/E [citar referências]	0.994	-	2.4
Modelo de água polarizável do MARTINI	1.043	-	0.625
Modelo BMW	1.047	-	-
SPC CG GROMOS	0.995	-23.3	6.9
Modelo SPC CG GROMOS*	0.997	-23.5	6.7

* Modelo SPC CG GROMOS implementado no GROMACS 4.6.x

As diferenças entre os valores reportados por Riniker e van Gunsteren e implementados no GROMACS 4.6.x são muito pequenos e podem estar associados à amostragem, visto que na Ref ⁷⁴, apenas 1 ns de simulação foram utilizados na análise, enquanto que a implementação no GROMACS foi avaliada durante 100 ns de simulação.

Apêndice C

Este apêndice tem como objetivo demonstrar a matemática e implementação do potencial quártico de ligação do modelo de água SPC CG do GROMOS⁷⁴ no GROMACS.²⁷

C. 1. Potencial Quártico de Ligação

O modelo de água SPC CG GROMOS tem como função de energia de ligação o termo quártico na forma:

$$V_{ij}(|\vec{r}_{ij}|) = \frac{1}{2} K_{ij} (|\vec{r}_{ij}| - |\vec{r}_{0,ij}|)^4, \quad (28)$$

em que K_{ij} é a constante de força da ligação, \vec{r}_{ij} e $\vec{r}_{0,ij}$ representam a distância e a distância de equilíbrio entre os átomos i e j , respectivamente. De acordo com Riniker e van Gunsteren,⁷⁴ o uso desta função permite que a polarização da esfera CG aconteça de forma não linear e ainda evita a superpolarização quando duas moléculas distintas se aproximam.

C. 2. Força gerada por um Potencial Quártico

De acordo com a mecânica geral para sistemas conservativos, a força é o negativo do gradiente do potencial:

$$\vec{F}(\vec{r}) = -\frac{dV(\vec{r})}{d\vec{r}}. \quad (29)$$

Para o potencial quártico, portanto a força tem forma:

$$\vec{F}_{ij}(\vec{r}_{ij}) = -\frac{1}{2}K_{ij} \frac{d(|\vec{r}_{ij}| - |\vec{r}_{0,ij}|)^4}{d\vec{r}_{ij}}. \quad (30)$$

Pela regra da cadeia tem-se:

$$\frac{dV(\vec{r})}{d\vec{r}} = \frac{dV(\vec{r})}{d|\vec{r}|} \frac{d|\vec{r}|}{d\vec{r}}. \quad (31)$$

Como

$$\frac{dV(|\vec{r}_{ij}|)}{d|\vec{r}_{ij}|} = 4(|\vec{r}_{ij}| - |\vec{r}_{0,ij}|)^3. \quad (32)$$

E ainda como

$$\frac{d|\vec{r}|}{d\vec{r}} = \frac{\vec{r}}{|\vec{r}|}. \quad (33)$$

Tem-se portanto que a força é dada por:

$$\vec{F}_{ij}(\vec{r}_{ij}) = -2K_{ij}(|\vec{r}_{ij}| - |\vec{r}_{0,ij}|)^3 \frac{\vec{r}_{ij}}{|\vec{r}_{ij}|}. \quad (34)$$

C. 3. Potencial Quártico e Dependência com λ

Esta seção descreve a dependência do potencial quártico com o parâmetro λ usado em cálculos de energia livre. Nestes tipos de cálculo o potencial é interpolado entre dois estados A ($\lambda = 0$) e B ($\lambda=1$) e vice-versa. Para o potencial quártico esta dependência pode ser escrita na forma:

$$V_{ij}(|\vec{r}_{ij}|, \lambda) = \frac{1}{2} [(1 - \lambda)K_{ij}^A + \lambda K_{ij}^B] \cdot [|\vec{r}_{ij}| - (1 - \lambda)|\vec{r}_{0,ij}^A| - \lambda|\vec{r}_{0,ij}^B|]^4. \quad (35)$$

Desta forma a derivada potencial em relação a λ é:

$$\begin{aligned} \frac{\partial V_{ij}(|\vec{r}_{ij}|, \lambda)}{\partial \lambda} &= \\ &= -\frac{1}{2} (K_{ij}^B - K_{ij}^A) \cdot [|\vec{r}_{ij}| - (1 - \lambda)|\vec{r}_{0,ij}^A| - \lambda|\vec{r}_{0,ij}^B|]^4 \\ &+ 2(|\vec{r}_{0,ij}^A| - |\vec{r}_{0,ij}^B|) \cdot [(1 - \lambda)K_{ij}^A + \lambda K_{ij}^B] \\ &\cdot [|\vec{r}_{ij}| - (1 - \lambda)|\vec{r}_{0,ij}^A| - \lambda|\vec{r}_{0,ij}^B|]^3. \end{aligned} \quad (36)$$

Note que a derivada do potencial em relação a λ é uma grandeza escalar.

C. 4. Implementação em C no GROMACS 4.5.x e 4.6.x

Ao invés de adicionar o potencial quártico ao código, alterou-se o potencial harmônico para o potencial quártico. A implementação depende da modificação da função `harmonic(...)`, que é chamada pela função `bonds(...)`. Ambas as funções se encontram, dentro da pasta de instalação do GROMACS, no arquivo `src/gmxlib/bondfree.c`.

É importante destacar que na função `bonds(...)` acontece a chamada para a função `harmonic(...)`, mas o denominador da força (**Equação 29**) é calculada na primeira. Portanto, mesmo que a função `harmonic(...)` calcule a energia, derivada em relação a λ e a força a associada ao potencial quártico, este último termo não conterà seu denominador.

C. 4. 1. Função `bonds(...)`

```
real bonds(int nbonds,
```

```

    const t_iatom forceatoms[],
    const t_iparams forceparams[],
    const rvec x[], rvec f[], rvec fshift[],
    const t_pbc *pbc, const t_graph *g,
    real lambda, real *dvdlambda,
    const t_mdatoms *md, t_fcdata *fcd,
    int *global_atom_index)
{
    int i, m, ki, ai, aj, type;
    real dr, dr2, fbond, vbond, fij, vtot;
    rvec dx;
    ivec dt;

    vtot = 0.0;
    for (i = 0; (i < nbonds); )
    {
        type = forceatoms[i++];
        ai = forceatoms[i++];
        aj = forceatoms[i++];

        ki = pbc_rvec_sub(pbc, x[ai], x[aj], dx); /* 3 */
        dr2 = iprod(dx, dx); /* 5 */
        dr = dr2*gmX_invsqrt(dr2); /* 10 */

        *dvdlambda += harmonic(forceparams[type].harmonic.krA,
                               forceparams[type].harmonic.krB,
                               forceparams[type].harmonic.rA,
                               forceparams[type].harmonic.rB,
                               dr, lambda, &vbond, &fbond);

        if (dr2 == 0.0)
        {
            continue;
        }

        vtot += vbond; /* 1 */
        fbond *= gmX_invsqrt(dr2); /* 6 */

        if (g)
        {
            ivec_sub(SHIFT_IVEC(g, ai),
                    SHIFT_IVEC(g, aj), dt);
        }
    }
}

```

```

        ki = IVEC2IS(dt);
    }
    for (m = 0; (m < DIM); m++)          /* 15 */
    {
        fij                = fbond*dx[m];
        f[ai][m]           += fij;
        f[aj][m]           -= fij;
        fshift[ki][m]      += fij;
        fshift[CENTRAL][m] -= fij;
    }
}
return vtot;
}

```

C. 4. 2. Função harmonic (...)

```

real harmonic(real kA,
              real kB,
              real xA,
              real xB,
              real x,
              real lambda,
              real *V,
              real *F)
{
    const real half=0.5, two=2.0;
    real L1, kk, x0, dx, dx2, dx3, dx4;
    real v, f, dvdlambda;

    L1    = 1.0-lambda;
    kk    = L1*kA+lambda*kB;
    x0    = L1*xA+lambda*xB;

    dx    = x-x0;
    dx2   = dx*dx;
    dx3   = dx2*dx;
    dx4   = dx2*dx2;

    f     = -two*kk*dx3;
    v     = half*kk*dx4;
}

```

```
dvdlambda = half*(kB-kA)*dx4 + two*(xA-xB)*kk*dx3;  
  
*F      = f;  
*V      = v;  
  
return dvdlambda;  
}
```

Apêndice D

Este apêndice tem como objetivo listar os arquivos utilizados nas simulações desta tese. É importante notar que o PITOMBA devido à modificação do código para alterar na função harmônica para quártica a aplicação dos arquivos abaixo é apenas útil na versão modificada do programa, exceto pelos arquivos de parâmetros das dinâmicas. Os arquivos estão no formato do GROMACS 4.6.x.²⁷

D.1. Arquivos mdp das simulações

D.1.1. Simulações Atomísticas

As etapas de minimização foram realizadas com as opções

```
integrator           = steep
tinit                = 0.0
dt                   = 0.002
nsteps               = 500000
nstcomm              = 1
comm-grps            = Other SOL

nstxout              = 1000
nstvout              = 1000
nstfout              = 0
nstlog               = 1000
nstenergy            = 1000
nstxtcout            = 1000
xtc_precision        = 1000
xtc-grps             =
energygrps           = Other SOL

nstlist              = 1
ns_type              = grid
pbc                  = xyz
rlist                = 1.4

coulombtype          = reaction-field
rcoulomb              = 1.4
```

```

epsilon_rf           = 61

vdw_type            = cut-off
rvdw                = 1.4
DispCorr            = No

tcoupl              = v-rescale
tc-grps             = Other SOL
tau_t               = 0.2 0.2
ref_t               = 298 298

Pcoupl              = berendsen
Pcoupltype          = isotropic
tau_p               = 0.5
compressibility     = 4.5e-5
ref_p               = 1.013

constraints         = none

```

As dinâmicas foram realizadas com as opções

```

integrator          = md
tinit               = 0.0
dt                  = 0.002
nsteps              = 50000000
nstcomm             = 1
comm-grps           = Other SOL

nstxout             = 0
nstvout             = 0
nstfout             = 0
nstlog              = 1000
nstenergy           = 1000
nstxtcout           = 1000
xtc_precision       = 1000
xtc-grps            =
energygrps          = Other SOL

nstlist             = 1
ns_type             = grid
pbc                 = xyz
rlist               = 1.4

```

```

coulombtype      = reaction-field
rcoulomb         = 1.4
epsilon_rf       = 61

vdw_type         = cut-off
rvdw             = 1.4
DispCorr        = No

tcoupl           = v-rescale
tc-grps         = Other SOL
tau_t           = 0.2 0.2
ref_t           = 298 298

Pcoupl           = berendsen
Pcoupltype      = isotropic
tau_p           = 0.5
compressibility  = 4.5e-5
ref_p           = 1.013

constraints      = h-bonds

```

D.1.2. Simulações “*coarse-grain*”

As etapas de minimização foram realizadas com as opções

```

integrator       = steep
tinit           = 0.0
dt              = 0.002
nsteps          = 50000
nstcomm         = 1
comm-grps       = Sugar SOL
nstcalcenergy   = 1

nstxout         = 0
nstvout         = 0
nstfout         = 0
nstlog          = 1000
nstenergy       = 1000
nstxtcout       = 1000
xtc_precision   = 1000
xtc-grps        =
energygrps      = System

```

```

nstlist           = 1
ns_type           = grid
pbc               = xyz
rlist            = 2.0

coulombtype       = reaction-field
rcoulomb          = 2.0
epsilon_rf        = 78.5
epsilon_r         = 2.5

vdw_type          = cut-off
rvdw              = 2.0
DispCorr         = No

tcoupl           = v-rescale
tc-grps          = System
tau_t            = 0.2
ref_t            = 298

Pcoupl           = berendsen
Pcoupltype       = isotropic
tau_p            = 0.5
compressibility   = 7.51e-4
ref_p            = 1.013

constraints       = none

```

As dinâmicas foram realizadas com as opções

```

integrator        = md
tinit            = 0.0
dt               = 0.002
nsteps           = 50000000
nstcomm          = 1
comm-grps        = Sugar SOL
nstcalcenergy    = 1

nstxout          = 0
nstvout          = 0
nstfout          = 0
nstlog           = 1000
nstenergy        = 1000

```

```

nstxtcout          = 1000
xtc_precision      = 1000
xtc-grps           =
energygrps         = System

nstlist            = 1
ns_type            = grid
pbc                = xyz
rlist              = 2.0

coulombtype        = reaction-field
rcoulomb           = 2.0
epsilon_rf         = 78.5
epsilon_r          = 2.5

vdw_type           = cut-off
rvdw               = 2.0
DispCorr           = No

tcoupl             = v-rescale
tc-grps            = System
tau_t              = 0.2
ref_t              = 298

Pcoupl             = berendsen
Pcoupltype         = isotropic
tau_p              = 0.5
compressibility    = 7.51e-4
ref_p              = 1.013

constraints        = none

```

As energias livres foram obtidas com as opções

```

integrator         = sd
tinit              = 0.0
dt                 = 0.002
nsteps             = 500000
nstcomm           = 1
comm-grps         = Sugar SOL

nstxout            = 0
nstvout           = 0

```

```

nstfout           = 0
nstlog           = 1000
nstenergy        = 1000
nstxtcout        = 0
xtc_precision    = 1000
xtc-grps         =
energygrps       = Sugar SOL

nstlist          = 1
ns_type          = grid
pbc              = xyz
rlist            = 2.0

coulombtype      = reaction-field
rcoulomb         = 2.0
epsilon_rf       = 78.5
epsilon_r        = 2.5

vdw_type         = cut-off
rvdw             = 2.0
DispCorr         = No

tcoupl           = v-rescale
tc-grps          = System
tau_t            = 0.2
ref_t            = 298

Pcoupl           = berendsen
Pcoupltype       = isotropic
tau_p            = 0.5
compressibility  = 7.51e-4
ref_p            = 1.013

constraints      = none

free-energy       = yes
couple-moltype   = Sugar
init-lambda      = $LAMBDA$

sc-power         = 1
sc-sigma         = 0.3
sc-alpha         = 0.5

```

```

couple-intramol      = no
couple-lambda0       = vdwq/vdw
couple-lambda1       = vdw/none
foreign-lambda       = $ALL_LAMBDAS$

```

D.2. Arquivos de topologia do PITOMBA

Por motivos de tamanho dos arquivos apenas alguns sistemas terão seus arquivos de topologia adicionados, sendo estes a α -D-Glicose, maltose, celobiose e o modelo de água SPC CG GROMOS.

D.2.1. α -D-Glicose

```

; Include forcefield parameters
#include "pitomba.ff/forcefield.itp"

[ moleculetype ]
; Name          nrexcl
Sugar           3

[ atoms ]
; nr      type  resnr residue  atom  cgnr      charge      mass  typeB
chargeB   massB
; residue 1 GLC rtp GLC  q  0.0
   1      CV1   1   GLC  CV1   1   -0.531      29   ; qtot -0.531
   2      CV2   1   GLC  CV2   1    0.109      60   ; qtot -0.422
   3      CVM   1   GLC  CVM   1    0.106      29   ; qtot -0.316
   4      CV6   1   GLC  CV6   1    0.316      44   ; qtot  0

[ bonds ]
; ai    aj  funct  cs
   1    2    2     gb_1
   1    3    2     gb_2
   1    4    2     gb_3
   2    3    2     gb_4
   2    4    2     gb_5
   3    4    2     gb_6

[ angles ]
; ai    aj    ak  funct  cs
   2    1    3    2     ga_7
   2    1    4    2     ga_8
   3    1    4    2     ga_11
   1    2    3    2     ga_1
   1    2    4    2     ga_2
   3    2    4    2     ga_12

```

```

1      3      2      2      ga_3
1      3      4      2      ga_4
2      3      4      2      ga_9
1      4      2      2      ga_5
1      4      3      2      ga_6
2      4      3      2      ga_10

[ dihedrals ]
; ai    aj    ak    al funct    cs
  4     1     2     3     1    gd_4
  2     1     3     4     1    gd_5
  2     1     4     3     1    gd_3
  1     2     3     4     1    gd_1
  1     2     4     3     1    gd_6
  2     3     4     1     1    gd_2

[ dihedrals ]
; ai    aj    ak    al funct    cs
  1     3     2     4     2    gi_1

```

```

; Include water topology
#include "pitomba.ff/spc-cg.itp"

```

```

[ system ]
; Name
alpha-D-Glucose

```

```

[ molecules ]
; Compound      #mols
Sugar           1

```

D.2.2. Maltose

```

; Include forcefield parameters
#include "pitomba.ff/forcefield.itp"

```

```

[ moleculetype ]
; Name          nrexcl
Sugar           3

```

```

[ atoms ]
; nr          type  resnr  residue  atom  cgnr    charge    mass  typeB
chargeB      massB
; residue    1  GLA  rtp  GLA  q  0.0
  1          CV1    1    GLA   CV1    1   -0.531    29   ; qtot -0.531
  2          CV2    1    GLA   CV2    1    0.109    60   ; qtot -0.422
  3          CVM    1    GLA   CVM    1    0.106    29   ; qtot -0.316
  4          CV6    1    GLA   CV6    1    0.316    44   ; qtot  0
; residue    2  GLA  rtp  GLA  q  0.0
  5          CV1    2    GLA   CV1    2   -0.531    29   ; qtot -0.531
  6          CV2    2    GLA   CV2    2    0.109    60   ; qtot -0.422
  7          CVM    2    GLA   CVM    2    0.106    29   ; qtot -0.316
  8          CV6    2    GLA   CV6    2    0.316    44   ; qtot  0

```

```

[ bonds ]
; ai    aj funct    cs

```

1	2	2	gb_1
1	3	2	gb_2
1	4	2	gb_3
1	7	2	gb_7
2	3	2	gb_4
2	4	2	gb_5
3	4	2	gb_6
5	6	2	gb_1
5	7	2	gb_2
5	8	2	gb_3
6	7	2	gb_4
6	8	2	gb_5
7	8	2	gb_6

[pairs]

;	ai	aj	funct
	2	5	1
	2	6	1
	2	8	1
	3	5	1
	3	6	1
	3	8	1
	4	5	1
	4	6	1
	4	8	1

[angles]

;	ai	aj	ak	funct	cs
	2	1	3	2	ga_7
	2	1	4	2	ga_8
	2	1	7	2	ga_13
	3	1	4	2	ga_11
	3	1	7	2	ga_18
	4	1	7	2	ga_14
	1	2	3	2	ga_1
	1	2	4	2	ga_2
	3	2	4	2	ga_12
	1	3	2	2	ga_3
	1	3	4	2	ga_4
	2	3	4	2	ga_9
	1	4	2	2	ga_5
	1	4	3	2	ga_6
	2	4	3	2	ga_10
	6	5	7	2	ga_7
	6	5	8	2	ga_8
	7	5	8	2	ga_11
	5	6	7	2	ga_1
	5	6	8	2	ga_2
	7	6	8	2	ga_12
	1	7	5	2	ga_15
	1	7	6	2	ga_16
	1	7	8	2	ga_17
	5	7	6	2	ga_3
	5	7	8	2	ga_4
	6	7	8	2	ga_9
	5	8	6	2	ga_5
	5	8	7	2	ga_6

```

        6      8      7      2      ga_10

[ dihedrals ]
; ai    aj    ak    al funct    cs
  4     1     2     3     1    gd_4
  2     1     3     4     1    gd_5
  2     1     4     3     1    gd_3
  2     1     7     6     1    gd_7
  1     2     3     4     1    gd_1
  1     2     4     3     1    gd_6
  2     3     4     1     1    gd_2
  8     5     6     7     1    gd_4
  6     5     7     8     1    gd_5
  6     5     8     7     1    gd_3
  5     6     7     8     1    gd_1
  5     6     8     7     1    gd_6
  6     7     8     5     1    gd_2

[ dihedrals ]
; ai    aj    ak    al funct    cs
  1     7     4     2     2    gi_2
  1     3     2     4     2    gi_1
  5     7     6     8     2    gi_1

; Include water topology
#include "pitomba.ff/spc-cg.itp"

[ system ]
; Name
Maltose

[ molecules ]
; Compound      #mols
Sugar           1

```

D.2.3. Cellobiose

```

; Include forcefield parameters
#include "pitomba.ff/forcefield.itp"

[ moleculetype ]
; Name          nrexcl
Sugar          3

[ atoms ]
; nr          type  resnr residue  atom    cgnr    charge    mass  typeB
chargeB      massB
; residue    1 GLB rtp GLB  q  0.0
  1          CV1   1   GLB   CV1   1   -0.531    29   ; qtot -0.531
  2          CV2   1   GLB   CV2   1    0.109    60   ; qtot -0.422
  3          CVM   1   GLB   CVM   1    0.106    29   ; qtot -0.316
  4          CV6   1   GLB   CV6   1    0.316    44   ; qtot 0
; residue    2 GLB rtp GLB  q  0.0
  5          CV1   2   GLB   CV1   2   -0.531    29   ; qtot -0.531
  6          CV2   2   GLB   CV2   2    0.109    60   ; qtot -0.422
  7          CVM   2   GLB   CVM   2    0.106    29   ; qtot -0.316

```

8 CV6 2 GLB CV6 2 0.316 44 ; qtot 0

[bonds]

```
; ai aj funct cs
  1  2  2  gb_1
  1  3  2  gb_2
  1  4  2  gb_3
  1  7  2  gb_7
  2  3  2  gb_4
  2  4  2  gb_5
  3  4  2  gb_6
  5  6  2  gb_1
  5  7  2  gb_2
  5  8  2  gb_3
  6  7  2  gb_4
  6  8  2  gb_5
  7  8  2  gb_6
```

[pairs]

```
; ai aj funct
  2  5  1
  2  6  1
  2  8  1
  3  5  1
  3  6  1
  3  8  1
  4  5  1
  4  6  1
  4  8  1
```

[angles]

```
; ai aj ak funct cs
  2  1  3  2  ga_7
  2  1  4  2  ga_8
  2  1  7  2  ga_19
  3  1  4  2  ga_11
  3  1  7  2  ga_24
  4  1  7  2  ga_20
  1  2  3  2  ga_1
  1  2  4  2  ga_2
  3  2  4  2  ga_12
  1  3  2  2  ga_3
  1  3  4  2  ga_4
  2  3  4  2  ga_9
  1  4  2  2  ga_5
  1  4  3  2  ga_6
  2  4  3  2  ga_10
  6  5  7  2  ga_7
  6  5  8  2  ga_8
  7  5  8  2  ga_11
  5  6  7  2  ga_1
  5  6  8  2  ga_2
  7  6  8  2  ga_12
  1  7  5  2  ga_21
  1  7  6  2  ga_22
  1  7  8  2  ga_23
  5  7  6  2  ga_3
```

```

5      7      8      2      ga_4
6      7      8      2      ga_9
5      8      6      2      ga_5
5      8      7      2      ga_6
6      8      7      2      ga_10

```

```

[ dihedrals ]
; ai    aj    ak    al funct    cs
  4     1     2     3     1    gd_4
  2     1     3     4     1    gd_5
  2     1     4     3     1    gd_3
  2     1     7     6     1    gd_11
  2     1     7     6     1    gd_8
  2     1     7     6     1    gd_9
  2     1     7     6     1    gd_10
  2     1     7     6     1    gd_12
  1     2     3     4     1    gd_1
  1     2     4     3     1    gd_6
  2     3     4     1     1    gd_2
  8     5     6     7     1    gd_4
  6     5     7     8     1    gd_5
  6     5     8     7     1    gd_3
  5     6     7     8     1    gd_1
  5     6     8     7     1    gd_6
  6     7     8     5     1    gd_2

```

```

[ dihedrals ]
; ai    aj    ak    al funct    cs
  1     7     4     2     2    gi_3
  1     3     2     4     2    gi_1
  5     7     6     8     2    gi_1

```

```

; Include water topology
#include "pitomba.ff/spc-cg.itp"

```

```

[ system ]
; Name
Celobiose

```

```

[ molecules ]
; Compound      #mols
Sugar           1

```

D.2.4. α -ciclodextrina

```

; Include forcefield parameters
#include "pitomba.ff/forcefield.itp"

```

```

[ moleculetype ]
; Name          nrexcl
Sugar          3

```

```

[ atoms ]

```

nr	type	resnr	residue	atom	cgnr	charge	mass	typeB
chargeB		massB						
; residue		1	GLA	rtp	GLA	q	0.0	
1	CV1	1	GLA	CV1	1	-0.531	29	; qtot -0.531
2	CV2	1	GLA	CV2	1	0.109	60	; qtot -0.422
3	CVM	1	GLA	CVM	1	0.106	29	; qtot -0.316
4	CV6	1	GLA	CV6	1	0.316	44	; qtot 0
; residue		2	GLA	rtp	GLA	q	0.0	
5	CV1	2	GLA	CV1	2	-0.531	29	; qtot -0.531
6	CV2	2	GLA	CV2	2	0.109	60	; qtot -0.422
7	CVM	2	GLA	CVM	2	0.106	29	; qtot -0.316
8	CV6	2	GLA	CV6	2	0.316	44	; qtot 0
; residue		3	GLA	rtp	GLA	q	0.0	
9	CV1	3	GLA	CV1	3	-0.531	29	; qtot -0.531
10	CV2	3	GLA	CV2	3	0.109	60	; qtot -0.422
11	CVM	3	GLA	CVM	3	0.106	29	; qtot -0.316
12	CV6	3	GLA	CV6	3	0.316	44	; qtot 0
; residue		4	GLA	rtp	GLA	q	0.0	
13	CV1	4	GLA	CV1	4	-0.531	29	; qtot -0.531
14	CV2	4	GLA	CV2	4	0.109	60	; qtot -0.422
15	CVM	4	GLA	CVM	4	0.106	29	; qtot -0.316
16	CV6	4	GLA	CV6	4	0.316	44	; qtot 0
; residue		5	GLA	rtp	GLA	q	0.0	
17	CV1	5	GLA	CV1	5	-0.531	29	; qtot -0.531
18	CV2	5	GLA	CV2	5	0.109	60	; qtot -0.422
19	CVM	5	GLA	CVM	5	0.106	29	; qtot -0.316
20	CV6	5	GLA	CV6	5	0.316	44	; qtot 0
; residue		6	GLA	rtp	GLA	q	0.0	
21	CV1	6	GLA	CV1	6	-0.531	29	; qtot -0.531
22	CV2	6	GLA	CV2	6	0.109	60	; qtot -0.422
23	CVM	6	GLA	CVM	6	0.106	29	; qtot -0.316
24	CV6	6	GLA	CV6	6	0.316	44	; qtot 0

[bonds]

ai	aj	funct	cs
1	2	2	gb_1
1	3	2	gb_2
1	4	2	gb_3
1	7	2	gb_7
2	3	2	gb_4
2	4	2	gb_5
3	4	2	gb_6
5	6	2	gb_1
5	7	2	gb_2
5	8	2	gb_3
5	11	2	gb_7
6	7	2	gb_4
6	8	2	gb_5
7	8	2	gb_6
9	10	2	gb_1
9	11	2	gb_2
9	12	2	gb_3
9	15	2	gb_7
10	11	2	gb_4
10	12	2	gb_5
11	12	2	gb_6
13	14	2	gb_1

```

13 15 2 gb_2
13 16 2 gb_3
13 19 2 gb_7
14 15 2 gb_4
14 16 2 gb_5
15 16 2 gb_6
17 18 2 gb_1
17 19 2 gb_2
17 20 2 gb_3
17 23 2 gb_7
18 19 2 gb_4
18 20 2 gb_5
19 20 2 gb_6
21 22 2 gb_1
21 23 2 gb_2
21 24 2 gb_3
22 23 2 gb_4
22 24 2 gb_5
23 24 2 gb_6
; periodic bonds
21 3 2 gb_7

```

```

[ pairs ]
; ai aj funct
1 11 1
2 5 1
2 6 1
2 8 1
3 5 1
3 6 1
3 8 1
4 5 1
4 6 1
4 8 1
5 15 1
6 9 1
6 10 1
6 12 1
7 9 1
7 10 1
7 12 1
8 9 1
8 10 1
8 12 1
9 19 1
10 13 1
10 14 1
10 16 1
11 13 1
11 14 1
11 16 1
12 13 1
12 14 1
12 16 1
13 23 1
14 17 1
14 18 1

```

```

14 20 1
15 17 1
15 18 1
15 20 1
16 17 1
16 18 1
16 20 1
18 21 1
18 22 1
18 24 1
19 21 1
19 22 1
19 24 1
20 21 1
20 22 1
20 24 1
; periodic pairs
17 3 1
21 7 1
22 1 1
22 2 1
22 4 1
23 1 1
23 2 1
23 4 1
24 1 1
24 2 1
24 4 1

[ angles ]
; ai aj ak funct cs
2 1 3 2 ga_7
2 1 4 2 ga_8
2 1 7 2 ga_13
3 1 4 2 ga_11
3 1 7 2 ga_18
4 1 7 2 ga_14
1 2 3 2 ga_1
1 2 4 2 ga_2
3 2 4 2 ga_12
1 3 2 2 ga_3
1 3 4 2 ga_4
2 3 4 2 ga_9
1 4 2 2 ga_5
1 4 3 2 ga_6
2 4 3 2 ga_10
6 5 7 2 ga_7
6 5 8 2 ga_8
6 5 11 2 ga_13
7 5 8 2 ga_11
7 5 11 2 ga_18
8 5 11 2 ga_14
5 6 7 2 ga_1
5 6 8 2 ga_2
7 6 8 2 ga_12
1 7 5 2 ga_15
1 7 6 2 ga_16

```

1	7	8	2	ga_17
5	7	6	2	ga_3
5	7	8	2	ga_4
6	7	8	2	ga_9
5	8	6	2	ga_5
5	8	7	2	ga_6
6	8	7	2	ga_10
10	9	11	2	ga_7
10	9	12	2	ga_8
10	9	15	2	ga_13
11	9	12	2	ga_11
11	9	15	2	ga_18
12	9	15	2	ga_14
9	10	11	2	ga_1
9	10	12	2	ga_2
11	10	12	2	ga_12
5	11	9	2	ga_15
5	11	10	2	ga_16
5	11	12	2	ga_17
9	11	10	2	ga_3
9	11	12	2	ga_4
10	11	12	2	ga_9
9	12	10	2	ga_5
9	12	11	2	ga_6
10	12	11	2	ga_10
14	13	15	2	ga_7
14	13	16	2	ga_8
14	13	19	2	ga_13
15	13	16	2	ga_11
15	13	19	2	ga_18
16	13	19	2	ga_14
13	14	15	2	ga_1
13	14	16	2	ga_2
15	14	16	2	ga_12
9	15	13	2	ga_15
9	15	14	2	ga_16
9	15	16	2	ga_17
13	15	14	2	ga_3
13	15	16	2	ga_4
14	15	16	2	ga_9
13	16	14	2	ga_5
13	16	15	2	ga_6
14	16	15	2	ga_10
18	17	19	2	ga_7
18	17	20	2	ga_8
18	17	23	2	ga_13
19	17	20	2	ga_11
19	17	23	2	ga_18
20	17	23	2	ga_14
17	18	19	2	ga_1
17	18	20	2	ga_2
19	18	20	2	ga_12
13	19	17	2	ga_15
13	19	18	2	ga_16
13	19	20	2	ga_17
17	19	18	2	ga_3
17	19	20	2	ga_4

```

18  19  20  2  ga_9
17  20  18  2  ga_5
17  20  19  2  ga_6
18  20  19  2  ga_10
22  21  23  2  ga_7
22  21  24  2  ga_8
23  21  24  2  ga_11
21  22  23  2  ga_1
21  22  24  2  ga_2
23  22  24  2  ga_12
17  23  21  2  ga_15
17  23  22  2  ga_16
17  23  24  2  ga_17
21  23  22  2  ga_3
21  23  24  2  ga_4
22  23  24  2  ga_9
21  24  22  2  ga_5
21  24  23  2  ga_6
22  24  23  2  ga_10
; periodic angles
22  21  3  2  ga_13
23  21  3  2  ga_18
24  21  3  2  ga_14
21  3  1  2  ga_15
21  3  2  2  ga_16
21  3  4  2  ga_17

```

```

[ dihedrals ]
; ai  aj  ak  al  funct  cs
  4  1  2  3  1  gd_4
  2  1  3  4  1  gd_5
  2  1  4  3  1  gd_3
  2  1  7  6  1  gd_7
  1  2  3  4  1  gd_1
  1  2  4  3  1  gd_6
  2  3  4  1  1  gd_2
  8  5  6  7  1  gd_4
  6  5  7  8  1  gd_5
  6  5  8  7  1  gd_3
  6  5  11 10  1  gd_7
  5  6  7  8  1  gd_1
  5  6  8  7  1  gd_6
  6  7  8  5  1  gd_2
 12  9  10 11  1  gd_4
 10  9  11 12  1  gd_5
 10  9  12 11  1  gd_3
 10  9  15 14  1  gd_7
  9  10  11 12  1  gd_1
  9  10  12 11  1  gd_6
 10  11  12  9  1  gd_2
 16  13  14 15  1  gd_4
 14  13  15 16  1  gd_5
 14  13  16 15  1  gd_3
 14  13  19 18  1  gd_7
 13  14  15 16  1  gd_1
 13  14  16 15  1  gd_6
 14  15  16 13  1  gd_2

```

```

20 17 18 19 1 gd_4
18 17 19 20 1 gd_5
18 17 20 19 1 gd_3
18 17 23 22 1 gd_7
17 18 19 20 1 gd_1
17 18 20 19 1 gd_6
18 19 20 17 1 gd_2
24 21 22 23 1 gd_4
22 21 23 24 1 gd_5
22 21 24 23 1 gd_3
21 22 23 24 1 gd_1
21 22 24 23 1 gd_6
22 23 24 21 1 gd_2
; periodic dihedrals
22 21 3 2 1 gd_7

```

```

[ dihedrals ]
; ai aj ak al funct cs
1 7 4 2 2 gi_2
1 3 2 4 2 gi_1
5 11 8 6 2 gi_2
5 7 6 8 2 gi_1
9 15 12 10 2 gi_2
9 11 10 12 2 gi_1
13 19 16 14 2 gi_2
13 15 14 16 2 gi_1
17 23 20 18 2 gi_2
17 19 18 20 2 gi_1
21 23 22 24 2 gi_1
; periodic improper
21 3 24 22 2 gi_2

```

```

; Include water topology
#include "pitomba.ff/spc-cg.itp"

```

```

[ system ]
; Name
alpha cyclodextrin

```

```

[ molecules ]
; Compound #mols
Sugar 1

```

D.2.5. Água SPC CG GROMOS

```

[ moleculetype ]
; molname nrexcl
SOL 1

[ atoms ]
;id type resnr residu atom cgnr charge
1 CW 1 SOL CW 1 -0.575 60.077
2 DP 1 SOL DP 1 0.575 30.000

[bonds]
; i j funct length force const.

```

1 2 1 0.20 2.0e6

[exclusions]

1 2

D.2. Arquivos de geometria do PITOMBA

D.2.1. α -D-Glicose

Alpha-D-Glucose

4
1GLC CV1 1 2.494 2.599 2.605
1GLC CV2 2 2.736 2.697 2.518
1GLC CVM 3 2.773 2.525 2.752
1GLC CV6 4 2.544 2.507 2.850
5.00000 5.00000 5.00000

D.2.2. Maltose

Maltose

8
1GLA CV1 1 2.494 2.599 2.605
1GLA CV2 2 2.736 2.697 2.518
1GLA CVM 3 2.773 2.525 2.752
1GLA CV6 4 2.544 2.507 2.850
2GLA CV1 5 2.292 2.296 2.236
2GLA CV2 6 2.327 2.555 2.146
2GLA CVM 7 2.466 2.517 2.402
2GLA CV6 8 2.366 2.304 2.491
5.00000 5.00000 5.00000

D.2.3. Celobiose

Celobiose

8
1GLB CV1 1 2.601 2.470 2.513
1GLB CV2 2 2.746 2.700 2.510
1GLB CVM 3 2.919 2.469 2.498
1GLB CV6 4 2.816 2.263 2.484
2GLB CV1 5 2.079 2.526 2.489
2GLB CV2 6 2.231 2.302 2.518
2GLB CVM 7 2.330 2.518 2.486
2GLB CV6 8 2.276 2.750 2.501
5.00000 5.00000 5.00000

D.2.4. α -ciclodextrina

Alpha cyclodextrin

24
1GLA CV1 1 3.481 2.968 3.396

1GLA	CV2	2	3.644	3.075	3.194
1GLA	CVM	3	3.374	3.188	3.178
1GLA	CV6	4	3.216	3.066	3.390
2GLA	CV1	5	3.704	3.147	3.860
2GLA	CV2	6	3.886	3.111	3.648
2GLA	CVM	7	3.612	3.069	3.553
2GLA	CV6	8	3.434	3.114	3.779
3GLA	CV1	9	3.745	3.682	3.959
3GLA	CV2	10	3.955	3.582	3.800
3GLA	CVM	11	3.768	3.370	3.864
3GLA	CV6	12	3.647	3.448	4.084
4GLA	CV1	13	3.474	4.014	3.620
4GLA	CV2	14	3.744	3.975	3.550
4GLA	CVM	15	3.674	3.805	3.778
4GLA	CV6	16	3.463	3.913	3.885
5GLA	CV1	17	3.202	3.833	3.181
5GLA	CV2	18	3.465	3.918	3.128
5GLA	CVM	19	3.414	3.927	3.416
5GLA	CV6	20	3.154	3.936	3.441
6GLA	CV1	21	3.220	3.303	3.049
6GLA	CV2	22	3.431	3.456	2.936
6GLA	CVM	23	3.263	3.622	3.107
6GLA	CV6	24	3.039	3.474	3.214
7.00000	7.00000	7.00000			

Publicações Científicas

Durante o doutorado foram desenvolvidos trabalhos referentes e não referentes a esta tese, sendo eles:

The Effect of Temperature, Cations, and Number of Acyl Chains on the Lamellar to Non-Lamellar Transition in Lipid-A Membranes: A Microscopic View, Frederico J. S. Pontes, Victor H. Rusu, Thereza A. Soares, and Roberto D. Lins, *Journal of Chemical Theory and Computation* 2012 8 (10), 3830-3838

GROMOS 53A6GLYC, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates, Laercio Pol-Fachin, Victor H. Rusu, Hugo Verli, and Roberto D. Lins, *Journal of Chemical Theory and Computation* 2012 8 (11), 4681-4690

MDWiZ: A platform for the automated translation of molecular dynamics simulations, Victor H. Rusu, Vitor A.C. Horta, Bruno A.C. Horta, Roberto D. Lins, Riccardo Baron, *Journal of Molecular Graphics and Modelling*, Volume 48, March 2014, Pages 80-86, ISSN 1093-3263, <http://dx.doi.org/10.1016/j.jmgm.2013.12.006>

The Molecular Structure and Conformational Dynamics of Chitosan Polymers: An Integrated Perspective from Experiments and Computational Simulations, The Complex World of Polysaccharides, Dr. Desiree Nedra Karunaratne (Ed.), Richard A. Cunha, Thereza A. Soares, Victor H. Rusu, Frederico J.S. Pontes, Eduardo F. Franca and Roberto D. Lins (2012). ISBN: 978-953-51-0819-1, InTech, DOI: 10.5772/51803

PITOMBA: Parameter Interface for Oligosaccharide Molecules Based on Atoms, Victor H. Rusu, Riccardo Baron, Roberto D. Lins, *Journal of Chemical Theory and Computation.*, submetido

The Impact of Clonal Complex, Polymorphisms in Fibronectin-Binding Protein A and Dynamic Fibronectin-Binding Phenotypes on the Pathogenesis of *Staphylococcus aureus* Endovascular Infections, Yan Q. Xiong, Batu Sharma Kuinkel, Vance Fowler Jr., Thomas Rude, Nadia N. Casillas-Ituarte, Roberto D. Lins, Alex C. DiBartola, Wessam

Abdelhady, Isabelle F. T. Viana, Victor H. Rusu, Steven K. Lower, Arnold S. Bayer,
Journal of Infectious Disease, submetido