

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
CENTRO DE BIOTECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

**CARACTERIZAÇÃO ESTRUTURAL E CONFORMACIONAL DE  
PROSTAGLANDINA ENDOPERÓXIDO SINTASES**

Liana Guimarães Sachett

Porto Alegre – Brasil  
Janeiro de 2011

**CARACTERIZAÇÃO ESTRUTURAL E CONFORMACIONAL DE PROSTAGLANDINA  
ENDOPERÓXIDO SINTASES**

*Liana Guimarães Sachett*

Dissertação de mestrado elaborada no Grupo de Bioinformática Estrutural do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul sob orientação do professor doutor:

Hugo Verli

Porto Alegre – Brasil

Janeiro de 2011

**CARACTERIZAÇÃO ESTRUTURAL E CONFORMACIONAL DE PROSTAGLANDINA  
ENDOPERÓXIDO SINTASES**

*Liana Guimarães Sachett*

Dissertação submetida ao Programa de Pós-Graduação em Biologia Celular e Molecular do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul como parte dos requisitos necessários para a obtenção do grau de Mestre em Biologia Celular e Molecular.

Banca Examinadora:

---

Hugo Verli (Centro de Biotecnologia - UFRGS) (Presidente)

---

Charley Staats (Centro de Biotecnologia - UFRGS)

---

José Cláudio Fonseca Moreira (Instituto de Ciências Básicas da Saúde – UFRGS)

---

Mario Tyago Murakami (Laboratório Nacional de Luz Síncroton)

---

Fernanda Stanisçuaski (Centro de Biotecnologia - UFRGS) (Suplente)

Esta dissertação foi realizada sob a orientação do Professor Doutor Hugo Verli, com o apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) como requisito para obtenção do grau de Mestre em Biologia Celular e Molecular, junto ao Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul.

**FICHA CATALOGRÁFICA**

SACHETT, Liana Guimarães

**CARACTERIZAÇÃO ESTRUTURAL E CONFORMACIONAL DE PROSTAGLANDINA  
ENDOPERÓXIDO SINTASES**

Rio Grande do Sul, UFRGS, Centro de Biotecnologia, 2011.

Dissertação: Mestre em Ciências (Biologia Celular e Molecular)

- |                       |   |
|-----------------------|---|
| 1. Dinâmica Molecular | 2. Prostaglandina Endoperóxido Sintases |
| 3. Glicosilação       | 4. Ácido Araquidônico                   |

I. Hugo Verli

II. Universidade Federal do Rio Grande do Sul – Centro de Biotecnologia

III. Títulos

## **AGRADECIMENTOS**

Ao Prof. Hugo Verli, por ter me recebido junto ao Grupo de Bioinformática Estrutural e ter se dedicado a me orientar.

Ao Programa de Pós-Graduação em Biologia Celular e Molecular, pela oportunidade de realização deste trabalho.

Aos alunos do grupo de Bioinformática Estrutural, pelo companheirismo e convivência, tanto no laboratório quando fora dele.

Ao Luciano e Sílvia, pela competência e atendimento sempre atencioso.

Aos meus pais, pelo apoio e incentivo.

Desde já aos membros da banca, por terem aceitado o convite.

*“Um pouco de ciência nos afasta de Deus.  
Muito, nos aproxima.”*

Louis Pasteur.

**SUMÁRIO**

LISTA DE ABREVIATURAS .....	x
RESUMO .....	xi
ABSTRACT .....	xii
ÍNDICE DE FIGURAS .....	xiii
ÍNDICE DE TABELAS .....	xvi
1 Introdução.....	1
1.1 Prostaglandina endoperóxido sintases (PGHSs).....	1
1.2 Estrutura terciária e glicosilação de PGHSs .....	5
1.3 Mecanismo de ação de PGHSs.....	10
1.4 Inibidores de PGHSs .....	13
1.5 Caracterização conformacional de biomoléculas através de dinâmica molecular (DM) .....	16
2 Objetivos.....	20
3 Metodologia .....	21
3.1 Programas utilizados .....	21
3.2 Proposta do modelo de glicosilação pra PGHSs .....	21
3.3 Simulações por DM .....	22
3.3.1 Protocolo de simulação .....	22
3.3.2 Sistemas simulados.....	23
4 Resultados.....	25
4.1 Preâmbulo .....	25
4.2 Parte I .....	26
4.3 Parte II .....	31
4.4 Parte III.....	54

5	Discussão Geral .....	70
5.1	Influência da glicosilação na dinâmica de PGHSs.....	70
5.2	Caracterização da dinâmica do substrato das enzimas PGHSs em suas diferentes orientações .....	73
6	Conclusões.....	75
7	Perspectivas.....	76
8	Referências Bibliográficas .....	77
9	Anexos .....	90
9.1	Topologia utilizada para o AA.....	90
9.2	Arquivos .mdp.....	91
	Curriculum Vitae .....	100

## LISTA DE ABREVIATURAS

AA – ácido araquidônico

AINES – anti-inflamatórios não esteroidais

COX – ciclooxigenase

DM – dinâmica molecular

EGF – fator de crescimento epidermal

GlcNAc – N-acetil-glicosamina

Man – manose

MBD – membrane binding domain (domínio de ligação à membrana)

MD – molecular dynamics

NOE - Efeito Nuclear Overhauser

PDB – Protein Data Bank

PGG<sub>2</sub> – prostaglandina G<sub>2</sub>

PGH<sub>2</sub> – prostaglandina H<sub>2</sub>

PGHS – prostaglandina endoperóxido sintase

POX – peroxidase

RE – retículo endoplasmático

## RESUMO

Prostaglandina Endoperóxido Sintases (PGHSs) são enzimas homodiméricas integrais de membrana, N-glicosiladas, localizadas no lúmen do retículo endoplasmático, onde catalisam o primeiro passo na síntese de prostanóides, produzindo prostaglandina H<sub>2</sub> a partir do ácido araquidônico (AA). Constituem-se, assim, em um dos principais alvos terapêuticos no tratamento de doenças inflamatórias. PGHSs apresentam duas isoformas: PGHS-1, expressa constitutivamente em diversos tipos celulares, e PGHS-2, usualmente expressa por indução. Ambas isoformas são glicosiladas nos resíduos Asn68, Asn144 e Asn410, modificações necessárias para a atividade e função adequadas destas enzimas. A reação ciclooxygenase necessita que o AA esteja numa orientação catalítica. Adicionalmente a esta orientação, o banco de dados Protein Data Bank contém estruturas cristalográficas que apresentam uma orientação alternativa para o AA, proposta como não-catalítica. Considerando a importância da glicosilação para a função da enzima, o presente trabalho tem como objetivo a caracterização estrutural e conformacional de PGHSs em suas formas glicosiladas, bem como analisar as diferentes orientações adotadas pelo AA em dados cristalográficos prévios. Assim, modelos de PGHS-1 e PGHS-2 glicosiladas foram construídos combinando diferentes fontes de dados experimentais (bioquímicos, cristalográficos e espectrométricos) e foram submetidos a simulações de dinâmica molecular (DM) usando o pacote de simulação do GROMACS e campo de força GROMOS9643a1. Parâmetros desenvolvidos pelo grupo foram empregados para descrever a porção sacarídica dos sistemas. Os dados obtidos indicam uma diminuição na flexibilidade da proteína na presença de glicosilação no domínio EGF em ambas as isoformas. Além disso, PGHS-1 e PGHS-2 ligadas às duas diferentes orientações do AA foram submetidas a estudos de DM, indicando que a orientação alternativa não é estável, possivelmente sendo uma conformação minoritária em solução. Tal orientação alternativa, contudo, pode sugerir novas abordagens para o planejamento de novos agentes anti-inflamatórios baseado na complementaridade à enzima-alvo.

## ABSTRACT

Prostaglandin Endoperoxide Synthases (PGHSs) are homodimeric integral membrane enzymes, N-glycosylated, located in the endoplasmic reticulum lumen, where they catalyze the first step in the prostanoids synthesis, generating prostaglandin H<sub>2</sub> from arachidonic acid (AA), so constituting one of the main therapeutic targets for treating inflammatory diseases. PGHSs present two isoforms: PGHS-1, constitutively expressed in several cell types, and PGHS-2, usually expressed by induction. Both isoforms are glycosylated at residues Asn68, Asn144 and Asn410, modifications necessary for adequate PGHSs enzymatic activities and function. The cyclooxygenase reaction requires AA to be in a catalytic orientation. Additionally to this orientation, the Protein Data Bank contains crystallographic structures presenting an alternative orientation for AA, proposed as non-catalytic. Considering the importance of glycosylation for the enzyme's function, the current work intends to structurally and conformationally characterize the role of the sacharidic moiety of PGHSs, as well as to analyze the different orientations adopted by AA found in previous crystallographic data. Thus, glycosylated models for PGHS-1 and PGHS-2 were built combining different sources of experimental data (biochemical, crystallographic and spectrometric) and submitted to molecular dynamics (MD) using the GROMACS simulation suite and GROMOS9643a1 force field. Parameters developed by the group were employed in order to describe the carbohydrate moiety. The obtained data indicate a decrease in protein flexibility in the presence of glycosylation on the EGF domain of both isoforms. Also, PGHS-1 and PGHS-2 bound to the two different orientations of AA were submitted to MD studies indicating that the alternative orientation is not stable, possibly being a minority orientation in solution. This orientation, however, may suggest new approaches for structure-based design of new anti-inflammatory agents.

## ÍNDICE DE FIGURAS

<b>Figura 1:</b> Alinhamento das sequências de aminoácidos das PGHSs 1 e 2 de murinos (mu), ovinos (ov) e humanos (hu).....	4
<b>Figura 2: A.</b> Estrutura terciária do monômero de PGHS. Domínios EGF (verde, resíduos 32 a 73), MBD (vermelho, resíduos 74 a 121) e catalítico (azul, resíduos 122 a 584). O substrato AA (esferas) está inserido no sítio COX, enquanto o grupamento heme localiza o sítio POX. (OBS: Resíduos 1 a 31, correspondentes ao peptídeo sinal, não mostrados). <b>B.</b> Imagem ilustrativa de um dímero de PGHS inserido em membrana (domínios coloridos como em A.).....	5
<b>Figura 3:</b> Cavidade do sítio COX para A. PGHS-1 e B. PGHS-2. Os resíduos que diferem entre as isoformas estão destacados pelos círculos vermelhos. (Adaptado de Roche, 2009).....	7
<b>Figura 4:</b> Coordenação o resíduo de His388 de PGHSs com o átomo de ferro do grupamento heme. O resíduo His207 também é mostrado.....	8
<b>Figura 5:</b> Localização dos sítios de N-glicosilação (esferas laranjas) nas PGHSs. Asn68 (domínio EGF) e Asn 144 e Asn410 (domínio catalítico). Domínios EGF, MBD e catalítico coloridos conforme figura 2. ....	9
<b>Figura 6:</b> Orientação catalítica do substrato AA (verde) no sítio COX (azul). ....	11
<b>Figura 7:</b> Representação esquemática do mecanismo de ação COX e a formação de PGG <sub>2</sub> a partir de AA. Adaptado de Malkowski <i>et al.</i> , 2000.....	12
<b>Figura 8:</b> Orientação alternativa do substrato AA (verde) no sítio COX (azul).....	13
<b>Figura 9:</b> Estrutura do substrato das PGHSs, AA, e dos principais inibidores de PGHSs (AINES). ....	16
<b>Figura 10:</b> Funções de energia que compõe campos de força. De cima para baixo estão representadas as equações que descrevem o estiramento de ligações	

químicas, ângulos de ligação, diedros e interações intermoleculares (eletrostáticas/Coulômbicas e van der Waals, também identificadas como potencial de Lennard-Jones). Adaptado de Verli, 2005, van Gunsteren <i>et al.</i> , 2006 e Fauth, 2009. ....	18
<b>Figura 11:</b> Estrutura sacarídica utilizada nos sítios de glicosilação das PGHSs (Asn68, Asn144 e Asn410).....	22
<b>Figura 12:</b> Esquema das etapas que compõe as simulações de DM a 310K. De 0ps a 30ps, uma série de etapas, de 5ps cada, aquece lentamente o sistema em incrementos de 50K. Após 30ps, o sistema está a 310K e se mantém nesta temperatura pelo resto da simulação (Verli, 2005).....	23
<b>Figura 13:</b> Análise de RMSD para o monômero de PGHS-1 nas formas não glicosilada (preto) e glicosilada (vermelho). Os sítios de glicosilação estão localizados nos domínios EGF (Asn68) e catalítico (Asn144 e Asn410). ....	27
<b>Figura 14:</b> Análise de RMSD para o monômero de PGHS-2 nas formas não glicosilada (preto) e glicosilada (vermelho). Os sítios de glicosilação estão localizados nos domínios EGF (Asn68) e catalítico (Asn144 e Asn410). ....	28
<b>Figura 15:</b> Dados de tRMSF para PGHS-1 nas formas não-glicosilada e glicosilada. Áreas em vermelho/laranja representam as regiões de maior flexibilidade, enquanto áreas azuis representam regiões estáveis. Os três domínios da enzima tem suas regiões delimitadas pelas linhas tracejadas. ....	29
<b>Figura 16:</b> Dados de tRMSF para PGHS-2 nas formas não-glicosilada e glicosilada. Áreas em vermelho/laranja representam as regiões de maior flexibilidade, enquanto área azuis representam regiões estáveis. Os três domínios da enzima tem suas regiões delimitadas pelas linhas tracejadas. ....	30

**Figura 17:** Sobreposição dos passos da DM (a cada 10 ns) para as oligomanoses nos três sítios de N-glicosilação das PGHSs, demonstrando a alta flexibilidade desta porção sacarídica. .... 72

## ÍNDICE DE TABELAS

Tabela 1: Preferências conformacionais dos dissacarídeos presentes nas PGHSs quando simulados livres em solução (Fernandes <i>et al.</i> , 2010).....	22
---	----

# 1 Introdução

## 1.1 *Prostaglandina endoperóxido sintases (PGHSs)*

Prostaglandina Endoperóxido Sintases (PGHSs), também conhecidas como ciclooxigenases (COXs), são enzimas pertencentes à superfamília das mieloperoxidases (Daiyasu & Toh, 2000) que atuam na síntese de prostaglandinas a partir de seu precursor, o ácido araquidônico (AA). As prostaglandinas são mediadores lipídicos que, dentre suas várias funções, atuam como sinalizadores em processos como dor, febre, inchaço e inflamação. Por conseguinte, a inibição de sua biossíntese (usualmente por fármacos denominados anti-inflamatórios não esteroidais ou AINES, conforme descrito adiante) causa diversas alterações fisiopatológicas no organismo, como diminuição da síntese de prostaciclina e tromboxanos envolvidos na proteção da mucosa gástrica e prevenção de sangramentos inapropriados, respectivamente. Além disso, existem evidências de que estas enzimas têm um papel direto em outras patologias além da inflamação, como trombose (Patrignani *et al.*, 1994), vários tipos de câncer (Kalgutkar & Zhao, 2001; Song *et al.*, 2002; Thun *et al.*, 2002) e desordens neurológicas como as doenças de Alzheimer (McGeer & McGeer, 1999) e Parkinson (Smythies, 1996).

Existem duas isoformas bem caracterizadas de PGHSs: PGHS-1, conhecida como a isoforma constitutiva e PGHS-2, a isoforma majoritariamente induzida. Estas isoformas compartilham 60% de identidade de sequência (Smith & DeWitt, 1996) (Figura 1), são homodímeros, N-glicosiladas e encontradas no lúmen do retículo endoplasmático (RE) e envelope nuclear (DeWitt *et al.*, 1981). Ambas possuem atividade ciclooxigenase (COX) e peroxidase (POX). Existe ainda uma terceira isoforma, PGHS-3, encontrada em córtex cerebral de caninos, que é uma variante de *splicing* do gene da PGHS-1 (Chandrasekharan *et al.*, 2002), postulada como envolvida na ação de fármacos como dipirona e acetaminofeno. Esta terceira isoforma consiste de um RNAm de PGHS-1 que retém o íntron 1. Apesar de este íntron possuir tamanho semelhante em todas as espécies, em humanos e roedores ele acarreta em um deslocamento de leitura na sequência de códons, levando a formação de uma proteína muito diferente da PGHS-1. Portanto, para fazer uma proteína funcional seriam necessários sítios de *splicing* alternativos, *frameshifting* ribossomal ou edição do RNA, o que ainda não foi descrito em humanos

(Chandrasekharan *et al.*, 2002; Dinchuk *et al.*, 2003; Dou *et al.*, 2004; Qin *et al.*, 2005).

A PGHS-1, com uma massa molecular de 71 kDa, foi isolada pela primeira vez em 1976 a partir de vesículas seminais de ovelhas (Hemler & Lands, 1976). Ela é considerada a isoforma constitutiva de PGHS, possuindo funções fisiológicas do tipo *housekeeping*, sintetizando prostaglandinas que regulam a atividade celular normal (DeWitt & Meade, 1993). Por exemplo, sua ativação leva a produção de prostaciclina que, quando liberada na mucosa gástrica, tem atividade citoprotetora (Whittle *et al.*, 1978). Além disso, a PGHS-1 acarreta na produção tromboxana A<sub>2</sub> nas plaquetas, causando agregação plaquetária e impedindo, assim, sangramentos inapropriados (Funk *et al.*, 1991). Em indivíduos com diminuição da perfusão renal, as prostaglandinas agem como importante mecanismo compensatório, aumentando a taxa de filtração glomerular (Brater *et al.*, 2001). Estas funções estão envolvidas, respectivamente, na gastrotoxicidade de AINES, no emprego terapêutico destes compostos na prevenção de acidentes tromboembólicos e nos sérios efeitos adversos renais em indivíduos que apresentam função renal dependente de prostaglandinas, incluindo falência renal aguda.

Em 1991 uma segunda isoforma de PGHS, também com massa molecular de 71 kDa, a forma induzível, foi identificada através de estudos de divisão celular (Xie *et al.*, 1991). Esta isoforma, chamada PGHS-2, está normalmente presente nas células em níveis muito baixos e o aumento de sua expressão é controlado por diversos fatores como citocinas, fatores de crescimento, mensageiros intracelulares e disponibilidade de substrato. Assim, estímulos inflamatórios aumentam a síntese de prostaglandinas devido à modulação da expressão de PGHS-2 (DeWitt & Meade, 1993).

A estrutura primária das enzimas PGHS-1 e PGHS-2 nascentes apresenta cerca de 600 (de 599 a 602, dependendo da espécie) e 604 resíduos de aminoácidos respectivamente (Smith & DeWitt, 1996). Ambas isoformas são então processadas por peptidases do lúmen do RE até as formas maduras pela remoção dos peptídeos sinais. Adicionalmente, o alto grau de identidade de sequência entre as isoformas processadas e entre espécies permite que sejam feitas comparações diretas (Smith *et al.*, 1996; Smith *et al.*, 2000). Assim, para auxiliar estas comparações, convencionou-se que os resíduos de PGHS seriam, usualmente,

numerados de acordo com as sequências de ovinos ou murinos da PGHS-1 (Smith & DeWitt, 1996) (Figura 1).

```

      10      20      30      40      50      60      70      80
huPGHS1  ---MSRSLLRFLFLFLLLPPLVLLADPAGAPTVPNPCCYYPCQHQGICVRFGLDRYQCDCTRTGYSGPNCTIPGLWTWL
ovPGHS1  --MSRQISILRFPLLLLLSPSPVFSADPGAPVPVNPCCYYPCQHQGICVRFGLDRYQCDCTRTGYSGPNCTIPELWTWL
muPGHS1  MSRRSLSLQFPPLLLLLPPPPVLLTDAGVPSVPVNPCCYYPCQNQGVCFVRFGLDHYQCDCTRTGYSGPNCTIPELWTWL
huPGHS2  -----MLARALLLCAVLALSHTANPCCSHPCQNRGVCMVSGFDQYKDCDTRTGFYGENCSTPEFLTRI
ovPGHS2  -----MLARALLLCAAVVCGAAN-PCCSHPCQNRGVCMVSGFDQYKDCDTRTGFYGENCSTPEFLTRI
muPGHS2  -----MLFRAVLLCAALGLSQAANPCCSNPCQNRGECMSTGFDQYKDCDTRTGFYGENCTTPEFLTRI
      : *      : :      .      * * * * * : * * : * : * : * : * : * : * : * : * : * : * :
      90      100     110     120     130     140     150     160
huPGHS1  RNSLRPSPSFTHFLLTHGRWFEVFN-ATFIREMLRLVLTVRSNLIPSPPTYNSAHDYISWESFSNVSYYTRILPSVPK
ovPGHS1  RTTLRPSPSFIHFMLTHGRWLWDFVN-ATFIRDTLMRLVLTVRSNLIPSPPTYNIAHDYISWESFSNVSYYTRILPSVPK
muPGHS1  RSSLRPSPSFTHFLLTHGYWFEVFN-ATFIREMLRLVLTVRSNLIPSPPTYNTAHDYISWESFSNVSYYTRILPSVPK
huPGHS2  KLFLKPTNTVHYILTHFKGVWNVNIPFLRNAIMSVLTSRSHLIDSPTYNADYGYKSWEAFSNLSYYTRALPPVPD
ovPGHS2  KLLKLPTDTVHYILTHFKGVWNIVNKISFLRNMIMRYVLTSRSHLIESSPTYNVHYSYKSWEAFSNLSYYTRALPPVPD
muPGHS2  KLLKLPTNTTVHYILTHFKGVWNIVNIPFLRSLIMKYVLTSRSYLIDSPTYNVHYGYKSWEAFSNLSYYTRALPPVAD
      : * : * : * : * * * * * : * : * * * * * : * * * * * * * * * * * : * * * * * : * * * * * : * * * * * : * * * * * :
      170     180     190     200     210     220     230     240
huPGHS1  DCPTMGTKGKQLPDAQLLRFLLRRFIPDPQGNLMFAFFAQHFTHQFFKTSGKMGPGTKALGHGVDLGHIYGDN
ovPGHS1  DCPTMGTKGKQLPDAELSRFLLRRFIPDPQGNLMFAFFAQHFTHQFFKTSGKMGPGTKALGHGVDLGHIYGDN
muPGHS1  DCPTMGTKGKQLPDIHLLQRLLRREFIPGQNTLVFAFAQHFTHQFFKTSGKMGPGTKALGHGVDLGHIYGDS
huPGHS2  DCPTPLGVKGKQLPDSNEIVEKLLRRFIPDPQGSNMMFAFAQHFTHQFFKTDHKRGPAFTNGLGHGVDLNHIYGET
ovPGHS2  DCPTPMGVKRKELPDSKEVKVLLRRRFIPDPQGNLMFAFAQHFTHQFFKTDIERGPAFTKGKNHGVDLSHVYGES
muPGHS2  DCPTPMGVKGNKELPDSKEVLEKVLLRRRFIPDPQGSNMMFAFAQHFTHQFFKTDHKRGPGTRGLGHGVDLNHIYGET
      * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * :
      250     260     270     280     290     300     310     320
huPGHS1  LERQYQLRLFKDGKLKYQVLDGEYPPSVEAPLMHYPRGIPPQSQMAVGQEVFGLLPGLMYATIWLREHNRVCDLLK
ovPGHS1  LERQYQLRLFKDGKLKYQMLDGEVPPSVEAPLMHYPRGIPPQSQMAVGQEVFGLLPGLMYATIWLREHNRVCDLLK
muPGHS1  LERQYHLRLFKDGKLKYQVLDGEVPPSVEASVLMRYPPGPVEPKQMAVGQEVFGLLPGLMLFSTIWLREHNRVCDLLK
huPGHS2  LARQRKLRLFKDGKMKYQIIDGEMYPPTVKDTQAEMIYPPQVPEHLRFAVGQEVFGLVPGLMYATIWLREHNRVCDVLK
ovPGHS2  LRQRNRLFKDGKMKYQMIGEMYPPTVKDTQVEMIYPPHIPEHLRFAVGQEVFGLVPGLMYATIWLREHNRVCDVLK
muPGHS2  LDRQHKLRLFFKDGKLKYQVIGEVYPPTVKDTQVEMIYPPHIPENLQFAVGQEVFGLVPGLMYATIWLREHNRVCDILK
      * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * :
      330     340     350     360     370     380     390     400
huPGHS1  AEHPTWDEQLFQTRLILIGETIKIVIEEYVQLSGYFLQLKFDPELLFGVQFYRNRIAEFNHLYHWHPLMPDSFKV
ovPGHS1  AEHPTWDEQLFQTARLLILIGETIKIVIEEYVQLSGYFLQLKFDPELLFGAQFYRNRIAEFNQLYHWHPLMPDSFRV
muPGHS1  EEHPTWDEQLFQTRLILIGETIKIIEEYVQHLSGYFLQLKFDPELLFRAQFYRNRIAEFNHLYHWHPLMPDSFQV
huPGHS2  QEHPEWDEQLFQTSRLLILIGETIKIVIEDYVQHLSGYHFKLKFDPELLFNKQFYQNRIAEFNTLYHWHPLLPDTFIQI
ovPGHS2  QEHPEWDEQLFQTSRLLILIGETIKIVIEDYVQHLSGYHFKLKFDPELLFNQQFYQNRIAEFNTLYHWHPLLPDVFIQI
muPGHS2  QEHPEWDEQLFQTSRLLILIGETIKIVIEDYVQHLSGYHFKLKFDPELLFNQQFYQNRIAEFNTLYHWHPLLPDTFNI
      * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * :
      410     420     430     440     450     460     470     480
huPGHS1  GSQEYSYEQLFNTSMLVDYGVEALVDAFSRQIAGRIGGRNMDHILHVAVDVIRESREMRLQPFNEYRKRFGMKPYTS
ovPGHS1  GPQDYSYEQLFNTSMLVDYGVEALVDAFSRQPAGRIGGRNIDHILHVAVDVIKESRVLRLQPFNEYRKRFGMKPYTS
muPGHS1  GSQEYSYEQLFNTSMLVDYGVEALVDAFSRQRAGRIGGRNFDYHVLHVAEDVIKESREMRLQSFNEYRKRFGLKPYTS
huPGHS2  HDQKYNYQFIYNNSILLEHGITQFVESFTRQIAGRVAGRNVPPAVQKVSQASIDQSRQMKYQSFNEYRKRFEMLKPYES
ovPGHS2  DGQEYNYQFIYNNSVLLEHGVTQFVESFTRQIAGRVAGRNLPAAVEKVSKASLDQSREMKYQSFNEYRKRFLLKPYES
muPGHS2  EDQEYSFKQLFYNNSILLEHGLTQFVESFTRQIAGRVAGRNVPIAVQAVAKASIDQSREMKYQSLNEYRKRFSLKPYTS
      * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * : * * :
      490     500     510     520     530     540     550     560
huPGHS1  FQELVGEKEMAELEELYGDIDALEFYPGLLLEKCHPNSIFGESMIEIGAPFSLKGLLGNPICSPEYWKPSTFGGEVGFN
ovPGHS1  FQELTGEKEMAELEELYGDIDALEFYPGLLLEKCHPNSIFGESMIEMGAPFSLKGLLGNPICSPEYWKASTFGGEVGFN
muPGHS1  FQEFTGEKEMAELEELYGDIDALEFYPGLMLEKCQPNSLFGESMIEMGAPFSLKGLLGNPICSPEYWKPSTFGGDVGFN
huPGHS2  FEELTGEKEMSAELEALYGDIDAVELYPALLVEKPRDAIFGETMVEVGAPFSLKGLMGNVICSPAYWKPSTFGGEVGFQ
ovPGHS2  FEELTGEKEMAELEALYGDIDAMELYPALLVEKPADAIFGETMVEAGAPFSLKGLMGNPICSPEYWKPSTFGGEVGFK
muPGHS2  FEELTGEKEMAELKALYSDIDVMELYPALLVEKPRDAIFGETMVELGAPFSLKGLMGNPICSPQYWKPSTFGGEVGFK
      * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * :
      570     580     590     600     610     620
huPGHS1  IVKTATLKKLVCLNTKTCPYVSFRVPDASQDDGPAVERPSTEL-----
ovPGHS1  IVKTATLKKLVCLNTKTCPYVSFHVPDPRQEDRPGVERPTEL-----
muPGHS1  IVNTASLKKLVCLNTKTCPYVSFRVPDYPGDDGSVFVRPSTEL-----
huPGHS2  IINTASIQSLICNNVKGCPFTSFSVPDPELIKTVTINASSRSGLDDINPTVLLKERSTEL
ovPGHS2  IINTASIQSLICNNVKGCPFTSFSVQDAHLTKTVTINASSHSGLDDINPTVLLKERSTEL
muPGHS2  IINTASIQSLICNNVKGCPFTSFNVQDPQTKTATINASASHSRLDDINPTVLIKRSTEL
      : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * * :

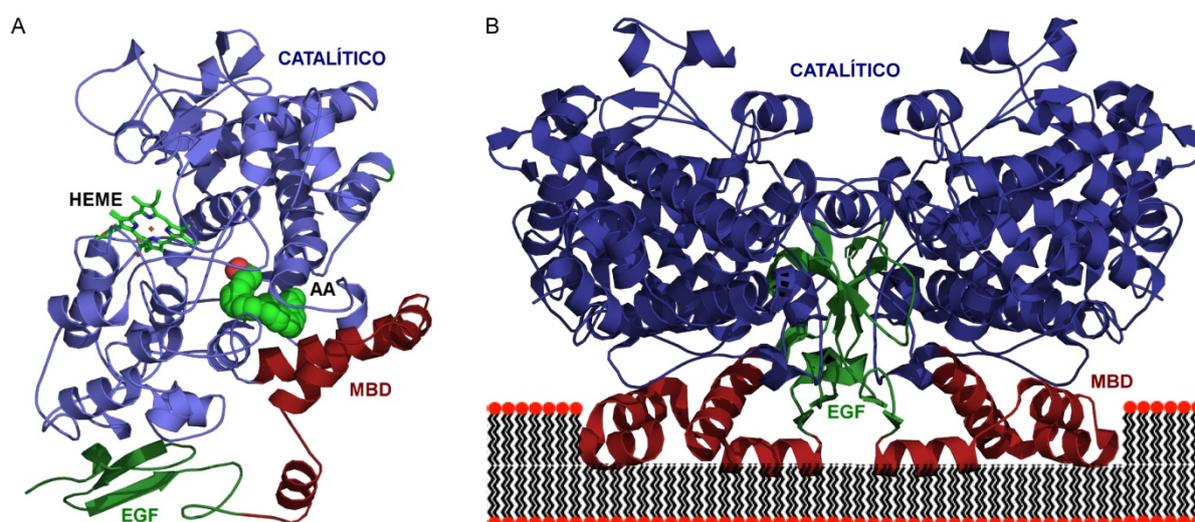
```

Identity (\*) : 322 is 51.85 %  
 Strongly similar (:) : 111 is 17.87 %  
 Weakly similar (.) : 47 is 7.57 %  
 Different : 141 is 22.71 %

**Figura 1:** Alinhamento das sequências de aminoácidos das PGHSs 1 e 2 de murinos (mu), ovinos (ov) e humanos (hu).

## 1.2 Estrutura terciária e glicosilação de PGHSs

A primeira estrutura tridimensional de PGHS foi determinada em 1994 para a isoforma 1 a partir de vesícula seminal de ovinos (Picot *et al.*, 1994). Em 1996, a estrutura cristalográfica de PGHS-2 de murinos foi obtida (Kurumbail *et al.*, 1996). Por possuírem uma alta identidade, as estruturas das duas isoformas são praticamente sobreponíveis (RMS = 0,91 Å). A partir destas estruturas pôde-se caracterizar a forma monomérica das PGHSs como composta por três domínios: um domínio N-terminal tipo fator de crescimento epidermal (EGF), que funciona como domínio de dimerização; um domínio de ligação à membrana (MBD); e um domínio catalítico C-terminal. No domínio catalítico encontram-se os sítios ativos COX e POX que, embora adjacentes, são espacialmente distintos. Também neste domínio encontra-se o sítio de ligação ao cofator da enzima, o grupamento heme (Figura 2).



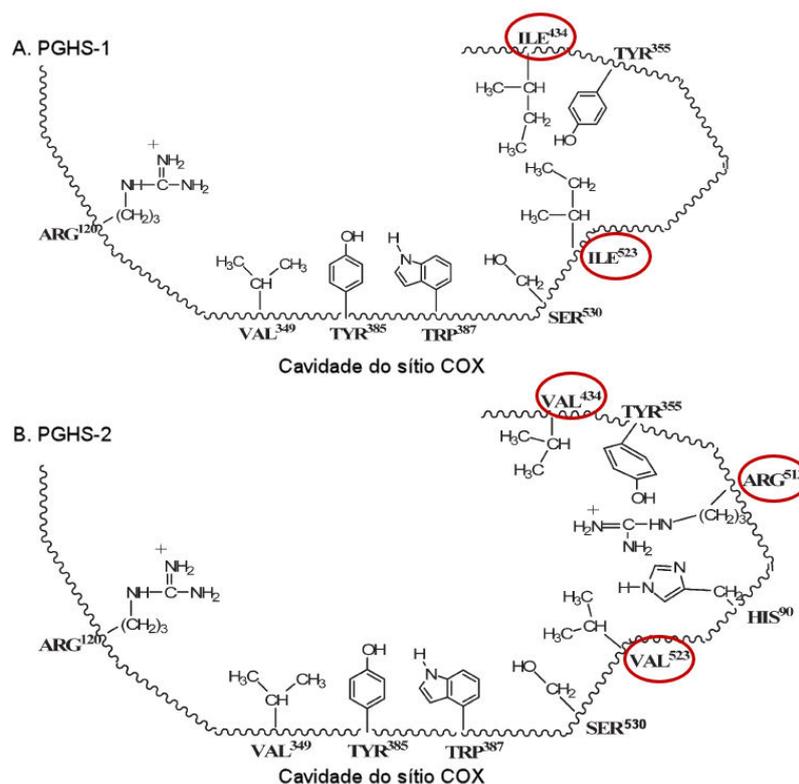
**Figura 2:** **A.** Estrutura terciária do monômero de PGHS. Domínios EGF (verde, resíduos 32 a 73), MBD (vermelho, resíduos 74 a 121) e catalítico (azul, resíduos 122 a 584). O substrato AA (esferas) está inserido no sítio COX, enquanto o grupamento heme localiza o sítio POX. (OBS: Resíduos 1 a 31, correspondentes ao peptídeo sinal, não mostrados). **B.** Imagem ilustrativa de um dímero de PGHS inserido em membrana (domínios coloridos como em A.).

O domínio EGF, considerado o domínio de dimerização nas PGHSs, é um domínio encontrado em outras famílias de proteínas de membrana e proteínas

secretadas (Campbell & Bork, 1993). Ocorre frequentemente na porção N-terminal a domínios de ancoramento em membrana. Nas PGHSs, acredita-se que este domínio desempenha um papel na inserção do domínio seguinte, o MBD, na bicamada lipídica (Garavito *et al.*, 1994).

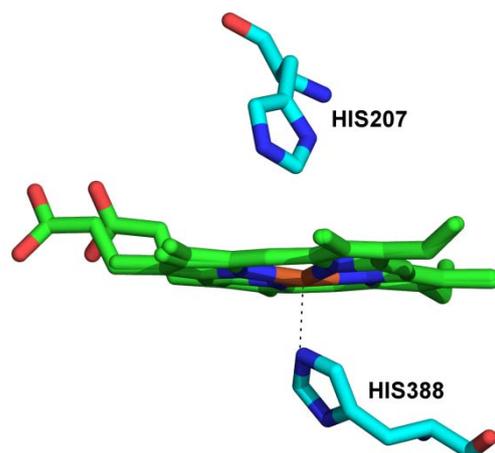
O MBD é um domínio formado por quatro hélices anfipáticas consecutivas que se inserem apenas em uma monocamada da bicamada lipídica (chamada inserção monotópica). Estas hélices são formadas por resíduos hidrofóbicos e aromáticos, o que permite que o substrato da enzima, AA, que é uma molécula hidrofóbica, ganhe acesso ao sítio ativo a partir do interior da bicamada lipídica através de uma solvatação hidrofóbica (Picot *et al.*, 1994).

O domínio catalítico das PGHSs é formado principalmente por estruturas de  $\alpha$ -hélices. O sítio COX é formado um canal longo e hidrofóbico que se estende a partir do MBD. O sítio ativo COX da PGHS-2 é levemente mais largo que o da PGHS-1 e postula-se que, assim, possa acomodar estruturas maiores do que aquelas capazes de encaixar-se no sítio da PGHS-1. Essa diferença de tamanho ocorre devido à diferença de três aminoácidos entre as isoformas 1 e 2: Ile523 para Val523, Ile434 para Val434 e His513 para Arg513 (Kurumbail *et al.*, 1996) (Figura 3).



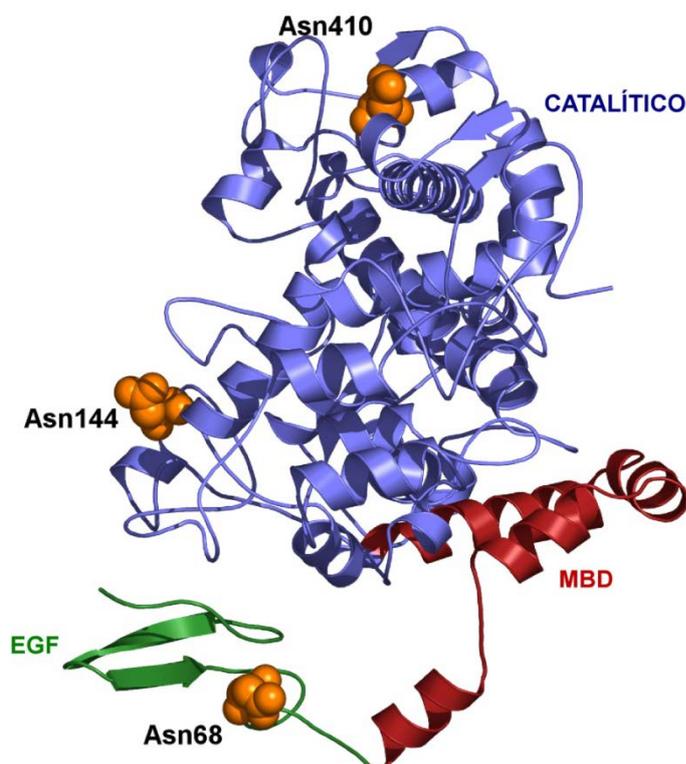
**Figura 3:** Cavidade do sítio COX para A. PGHS-1 e B. PGHS-2. Os resíduos que diferem entre as isoformas estão destacados pelos círculos vermelhos. (Adaptado de Roche, 2009).

O sítio POX, por sua vez, está em uma fenda larga no lado oposto ao MBD (Picot *et al.*, 1994). No sítio POX, a ligação ao grupamento heme ocorre com por coordenação da His388 proximal com o átomo de ferro do grupamento (Figura 4). Entretanto, esta histidina não forma uma ligação de hidrogênio ou iônica forte, o que a torna neutra. Já na face distal do grupamento heme, a His207 parece ser importante para a desprotonação do substrato peróxido (Seibold *et al.*, 2000).



**Figura 4:** Coordenação o resíduo de His388 de PGHSs com o átomo de ferro do grupamento heme. O resíduo His207 também é mostrado.

Ambas PGHSs são N-glicosiladas nos resíduos Asn68, Asn144 e Asn410 (Figura 5). A PGHS-2 contém ainda um quarto sítio, Asn594, inexistente na PGHS-1 e que é variavelmente N-glicosilado (Otto *et al.*, 1993) e parece regular o *turnover* desta isoforma (Sevigny *et al.*, 2006; Mbonye *et al.*, 2008). Os polissacarídeos encontrados nestes sítios são do tipo oligomanose, o que está coerente com a retenção de ambas as isoformas no RE (Mutsaers *et al.*, 1985; Nemeth *et al.*, 2001), já que a N-glicosilação deste tipo não sofre processos de maturação adicionais no complexo de Golgi (Dwek, 1996). A glicosilação das PGHSs faz parte do processo de controle de qualidade para seu correto enovelamento no lúmen do RE. O bloqueio do processo de glicosilação com tunicamicina, um inibidor de N-glicosilação (que bloqueia a adição do primeiro carboidrato por impedir a ação da enzima N-acetilglicosamina fosfotransferase), impede a observação da atividade para ambas isoformas. Contudo, o tratamento com glicosidase na proteína madura não afeta a atividade enzimática de forma significativa. Ainda, estudos mutacionais individuais indicam que os três sítios de glicosilação encontrados nas duas isoformas são essenciais para a atividade da enzima, provavelmente porque nenhum dos mutantes atingiu a conformação necessária para ligar o substrato enzimático ou o inibidor testado (Otto *et al.*, 1993).



**Figura 5:** Localização dos sítios de N-glicosilação (esferas laranjas) nas PGHSs. Asn68 (domínio EGF) e Asn 144 e Asn410 (domínio catalítico). Domínios EGF, MBD e catalítico coloridos conforme figura 2.

Alguns oligossacarídeos presentes em glicoproteínas têm sua estrutura, ou parte, dela depositada no banco de dados Protein Data Bank (PDB), juntamente com a estrutura da proteína. A maioria dessas estruturas é proveniente de cristalografia de raios-x ou ressonância magnética nuclear (RMN). Estas estruturas sacarídicas geralmente estão incompletas, pois a alta flexibilidade das glicanas impede tanto a cristalização quanto a aquisição de densidade eletrônica (Woods, 1998). Mesmo quando conseguem ser obtidas por cristalografia de raios-x, estas estruturas sacarídicas são de baixa qualidade, e estima-se que um terço dos dados para carboidratos encontrados no PDB contém erros significativos de estereoquímica, nomenclatura e concordância com mapas de densidade eletrônica (Lüteke *et al.*, 2004; Crispin *et al.*, 2007; Berman *et al.*, 2007). Da mesma maneira, as estruturas obtidas por RMN também têm dificuldades em descrever corretamente carboidratos, pois a alta flexibilidade das glicanas produz um conjunto complexo de conformações, com a possibilidade de coexistirem múltiplos confôrmeros simultaneamente em solução, concomitantemente a um baixo número ou ausência de sinais de NOE

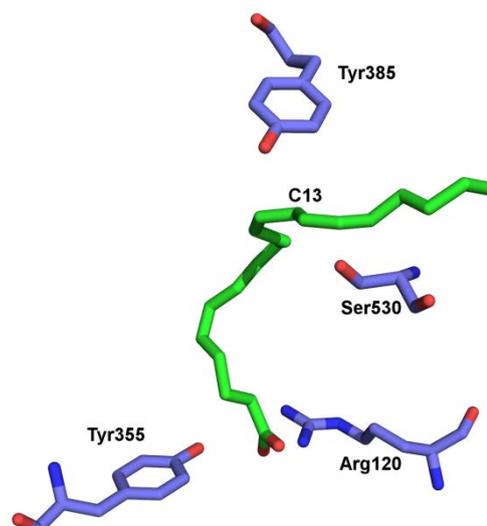
(Cumming & Carver, 1987; Woods, 1998; Wormald *et al.*, 2002; Pol-Fachin *et al.*, 2009). A espectrometria de massas, por sua vez, é a técnica que melhor descreve as estruturas dos oligossacarídeos (Pérez & Mulloy, 2005), embora não ofereça informações sobre a organização espacial, tridimensional, desta classe de biomoléculas.

Existem dados de RMN para as estruturas de glicosilação presentes na PGHS-1 (Mutsaers *et al.*, 1985), que detectaram apenas resíduos de N-acetilglicosamina e manose (6 a 9 manoses) e também dados de espectrometria de massas para as glicanas presentes na PGHS-2 (Nemeth *et al.*, 2001), que detectaram três glicofomas, contendo também apenas resíduos de N-acetilglicosamina e manose (6 a 10 manoses). Estes dados demonstram que existem glicofomas heterogêneas presentes nestas enzimas, todas do tipo oligomanose, mas com uma variação no número de manoses presente. Além disso, os dados presentes nas estruturas cristalográficas para PGHSs depositadas no PDB também apresentam apenas resíduos de N-acetilglicosamina e manose, embora todos estejam incompletos, já que a estrutura que mais contém manoses apresenta apenas três resíduos deste açúcar (ver item 4.3, tabela S2, material suplementar do artigo).

### **1.3 Mecanismo de ação de PGHSs**

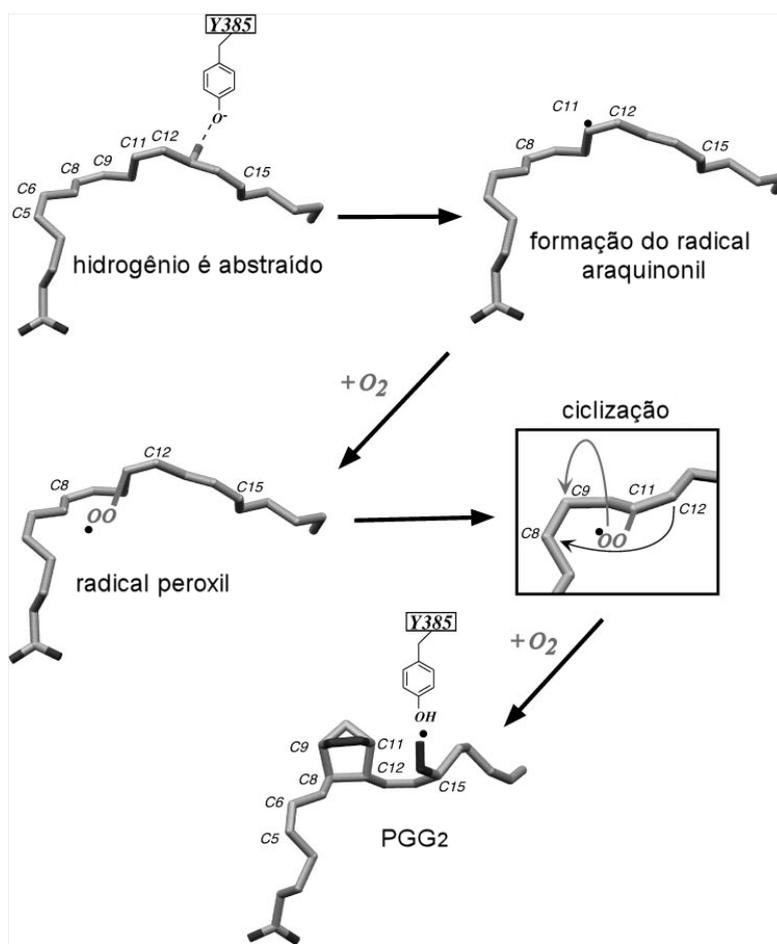
As PGHSs catalisam o primeiro passo na síntese de prostanóides. Inicialmente, convertem o substrato AA, duas moléculas de  $O_2$  e dois elétrons até prostaglandina  $G_2$  ( $PGG_2$ ) no sítio COX. Este intermediário segue então para o sítio POX, onde é reduzido a prostaglandina  $H_2$  ( $PGH_2$ ).

A atividade COX necessita que o AA esteja em uma conformação em forma de “L” e em uma orientação produtiva, com o grupamento carboxílico posicionado para interagir com o grupamento guanidínico da Arg120 (através de uma ponte salina ou ligação de hidrogênio) e o oxigênio fenólico da Tyr355. Além disso, uma curvatura do substrato entre os carbonos C7 e C14 circunda o resíduo de Ser530, que tem um papel estrutural, portanto, de manter a geometria de ataque da reação. O carbono C13 do substrato está próximo do oxigênio fenólico da Tyr385, o que favorece a abstração de um hidrogênio do AA quando o radical tirosil é formado neste resíduo (Malkowski *et al.*, 2000; Rouzer & Marnett, 2003) ([Figura 6](#)).



**Figura 6:** Orientação catalítica do substrato AA (verde) no sítio COX (azul).

A reação COX se inicia quando o radical tirosil no resíduo Tyr385 abstrai um hidrogênio do carbono C13 do substrato AA. A ativação do radical tirosil, por sua vez, envolve o grupamento heme do sítio POX. O radical assim gerado converte o substrato AA ao radical araquidonil. Na presença de oxigênio, o radical araquidonil é oxigenado no carbono C11 a um radical peroxil. Este radical ataca o carbono C9 para formar o endoperóxido, resultando na isomerização do carbono C8. Ocorre então uma reconfiguração do substrato que o posiciona para uma nova oxigenação no carbono C15, formando finalmente PGG<sub>2</sub>. O radical no carbono C15, que está agora posicionado abaixo da Tyr385, é quem faz a doação a este resíduo para completar o ciclo (Malkowski *et al.*, 2000) ([Figura 7](#)).

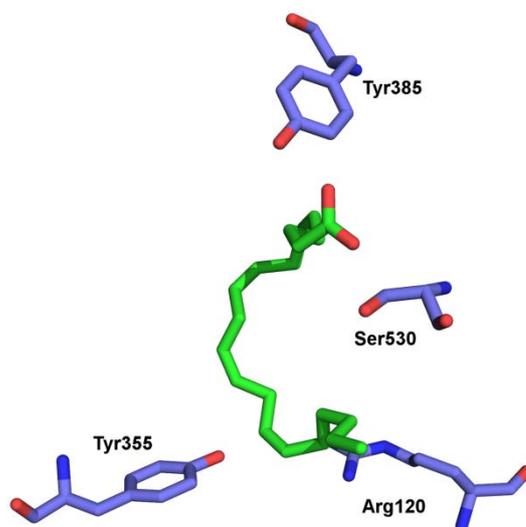


**Figura 7:** Representação esquemática do mecanismo de ação COX e a formação de PGG<sub>2</sub> a partir de AA. Adaptado de Malkowski *et al.*, 2000.

O intermediário PGG<sub>2</sub> formado segue para o sítio POX, onde é reduzido a PGH<sub>2</sub>. A atividade COX é dependente da oxidação do cofator heme, ou seja, depende da atividade POX, mas esta atividade POX não necessita ser contínua, pois o radical tirosil é regenerado a cada ciclo da atividade COX (Koshkin & Dunford, 1999). A PGH<sub>2</sub> é a prostaglandina da qual derivam todas as demais prostaglandinas e seus isômeros. Assim, sintetases específicas atuam na PGH<sub>2</sub>, via isomerização, oxidação ou redução, formando outras prostaglandinas, prostaciclina e tromboxanos.

Adicionalmente à orientação catalítica do substrato AA, existem estruturas cristalográficas que apresentam uma orientação alternativa para esta molécula, descrita como orientação não-produtiva ou não-catalítica (Kiefer *et al.*, 2000; Vecchio *et al.*, 2010). Nelas, o grupamento carboxílico do AA está interagindo com os

resíduos Tyr385 e Ser530 (Figura 8). Esta orientação, apesar de não ser favorável à catálise, é interpretada de maneira controversa na literatura (Kiefer *et al.*, 2000; Vecchio *et al.*, 2010). Por conseguinte, o papel da dinâmica desta orientação para o conhecimento da biologia estrutural e função biológica de PGHSs, assim como sua potencial relevância para o planejamento de novos agentes bioativos é, em grande parte, ausente (Furse *et al.*, 2006).



**Figura 8:** Orientação alternativa do substrato AA (verde) no sítio COX (azul).

#### 1.4 Inibidores de PGHSs

Os AINES, de maneira geral, atuam inibindo as enzimas PGHSs por competirem com o substrato, AA. A única exceção é a aspirina (ácido acetilsalicílico), que atua por modificação covalente, acetilando o resíduo Ser530 no sítio COX de maneira irreversível, bloqueando assim a formação de produtos nestas enzimas (Van der Ouderra *et al.*, 1980). A inibição por AINES é responsável pelas ações analgésica, antipirética e anti-inflamatória destes fármacos, bem como por seus efeitos gastrotóxicos e antitrombóticos. Para a maioria do AINES, os efeitos antitrombóticos são reversíveis assim que o fármaco é eliminado mas, no caso do ácido acetilsalicílico, que inibe as enzimas de forma irreversível, esse efeito só é revertido quando novas plaquetas são formadas, o que leva cerca de dez dias (Vane *et al.*, 1998).

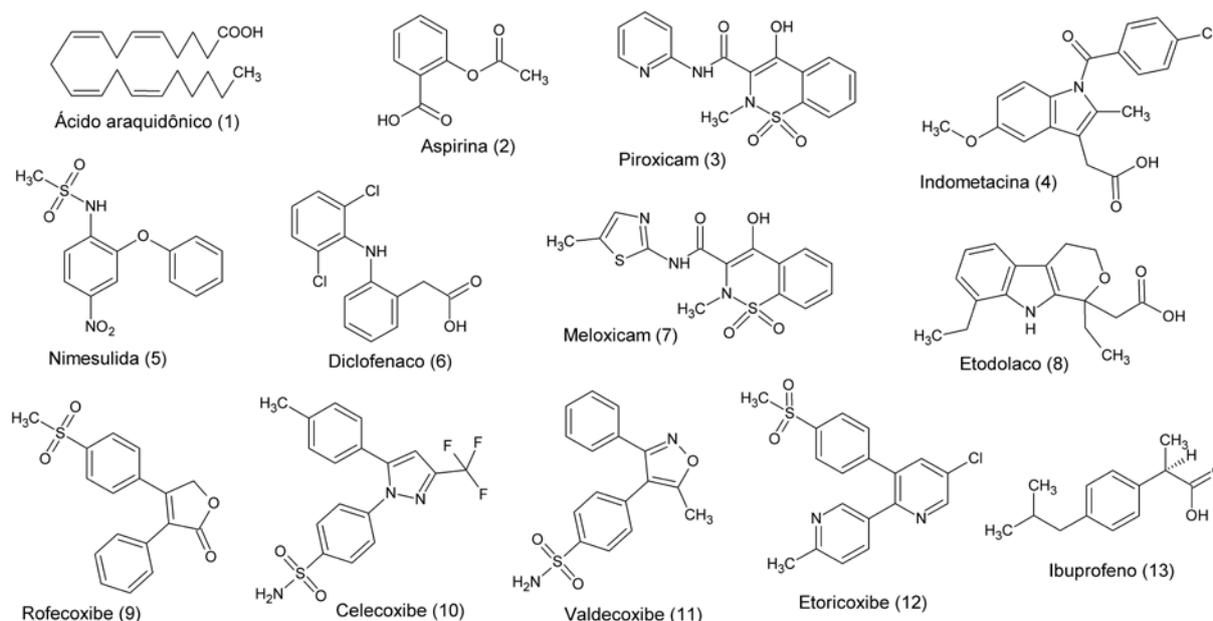
Devido ao fato de os AINES causarem efeitos colaterais, os mais severos no trato gastrointestinal, como sangramentos e úlceras, e com a descoberta da PGHS-2 como a isoforma induzível, passou-se a estudar as diferenças na farmacologia das duas enzimas (Mitchell *et al.*, 1993). Os fármacos ácido acetilsalicílico, piroxicam, indometacina e ibuprofeno são mais seletivos para PGHS-1 do que para PGHS-2, e também são estes AINES os que causam os maiores danos a mucosa estomacal (Meade *et al.*, 1993; Vane & Botting, 1995). Observou-se assim uma relação direta entre os efeitos colaterais gastrointestinais e a taxa de seletividade PGHS-1/PGHS-2. Os fármacos nimesulida, etodolaco e meloxicam foram descritos na década de 80 como potentes anti-inflamatórios com baixa atividade ulcerogênica. Após a descoberta da PGHS-2, observou-se que estes três fármacos apresentam preferência por esta segunda isoforma.

Em 1999 foram introduzidos fármacos planejados buscando seletividade para a PGHS-2, chamados coxibes, apresentando assim maior especificidade que compostos anteriormente no mercado. Os primeiros a serem introduzidos foram o celecoxibe, que contém um grupo sulfonamida, e o rofecoxibe, que contém uma metilsulfona (Figura 9). Estes fármacos são capazes de se ligar ao sítio COX da PGHS-2, mas interagem fracamente com a PGHS-1 (Kurumbail *et al.*, 1996). O celecoxibe tem uma seletividade de 7,6 e o rofecoxibe de 35 vezes a favor da PGHS-2. Os coxibes de segunda geração, como valdecoxibe (sucessor do celecoxibe) e etoricoxibe (sucessor do rofecoxibe) tem seletividade de 30 e 106 vezes a favor de PGHS-2, respectivamente (Figura 9) (Riendeau *et al.*, 2001). Estes fármacos possuem menores efeitos gastrotóxicos do que os outros AINES (Silverstein *et al.*, 2000). Entretanto, no caso do rofecoxibe, foi constatado um aumento na incidência de infarto do miocárdio nos pacientes que usavam este medicamento (Bombardier *et al.*, 2000), o que levou a retirada voluntária deste composto do mercado em 2004. Este efeito provavelmente ocorre porque estes fármacos, ao inibirem PGHS-2, inibem a síntese da prostaglandina antitrombótica (prostaciclina) nas células endoteliais (Marnett, 2009).

Apesar dos efeitos adversos dos AINES e coxibes, estes fármacos tem se mostrado promissores como agentes antitumorais, especialmente os seletivos para PGHS-2, por atuarem restaurando a apoptose normal, como observado em estudos em pólipos adenomatosos colorretais humanos e em várias linhagens celulares de câncer (Sumitani *et al.*, 2001). Além disso, foi observado que estes fármacos inibem

a angiogênese em culturas de células e modelos de angiogênese em roedores (Jones *et al.*, 1999). Estudos relacionam o uso contínuo de AINES com o baixo risco de câncer colorretal, pólipos adenomatosos e outros cânceres (Paganini-Hill, 1995). Estudos clínicos com pacientes portadores de pólipos adenomatosos familiar demonstraram que o uso de dois inibidores de PGHSs, sulindaco e celecoxibe, inibiram o crescimento dos pólipos adenomatosos e causaram a regressão dos pólipos existentes (Agarwal *et al.*, 1999; Steinbach *et al.*, 2000). Entretanto, questões ainda não resolvidas sobre eficácia, segurança e regime do tratamento limitam o uso destes medicamentos para esta função, sendo necessários cuidados e estudos de risco/benefício associados ao seu uso para prevenção de cânceres (Thun *et al.*, 2002).

Atualmente uma nova estratégia vem sendo desenvolvida para inibir as PGHSs, tendo em vista os efeitos colaterais dos AINES e dos coxibes. AINES liberadores de óxido nítrico estão sendo vistos como uma classe promissora para este fim, já que o óxido nítrico é uma molécula que apresenta efeitos citoprotetores no trato gastrointestinal, estimulando a secreção de muco pelas células gástricas e, além disso, possui propriedades relaxantes na musculatura lisa vascular. Desta maneira, estes compostos são potencialmente capazes de aumentar a segurança cardiovascular e também reduzir a toxicidade gastrointestinal (Keeble & Moore, 2002; Koç & Küçükgülzel, 2009).



**Figura 9:** Estrutura do substrato das PGHSs, AA, e dos principais inibidores de PGHSs (AINES).

### 1.5 Caracterização conformacional de biomoléculas através de dinâmica molecular (DM)

O uso de simulações de dinâmica molecular (DM), utilizado para biomoléculas desde a década de 70 (McCammon *et al.*, 1977), é hoje uma ferramenta amplamente difundida, utilizada para investigação da estrutura e dinâmica de biomoléculas em geral, abrangendo estudos de interações entre compostos e suas proteínas-alvo, desnaturação e enovelamento protéicos (Ponder & Case, 2003), entre outros.

Na DM, sucessivas configurações do sistema são geradas pela integração da equação do movimento de Newton:

$$\frac{d^2r_i}{dt^2} = \frac{F_i}{m_i}$$

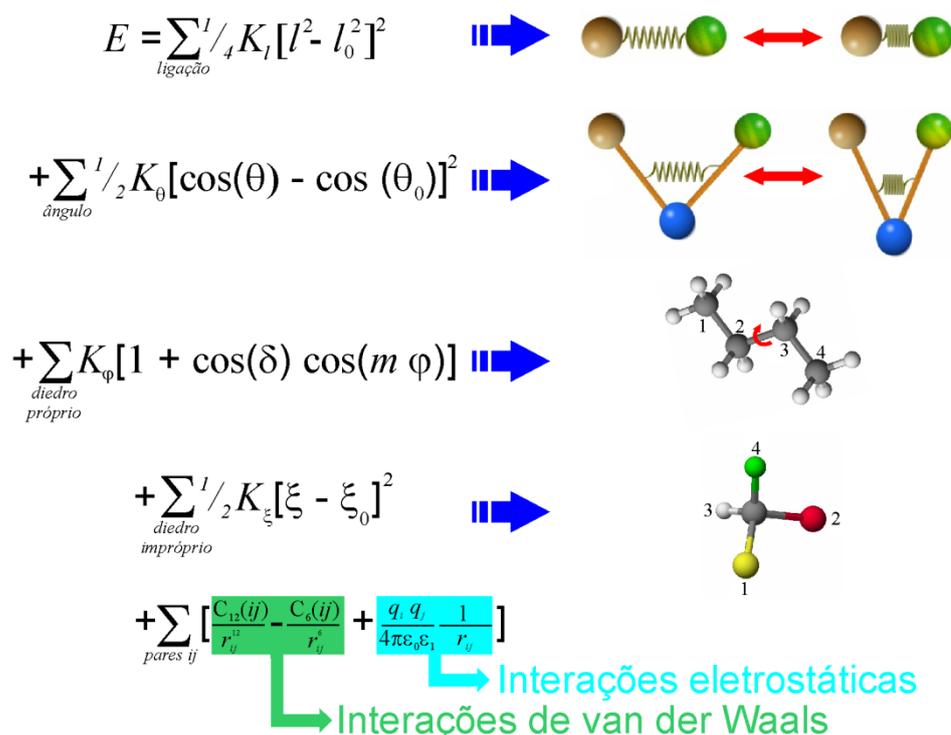
Nesta equação,  $d^2r_i/dt^2$  é a aceleração,  $m_i$  é a massa e  $F_i$  é a força sobre um determinado átomo  $i$ . O resultado desta integração empregada sobre todos os átomos do sistema é uma trajetória que especifica como as posições e velocidades

das partículas do sistema variam com o tempo relativo à aceleração (Leach, 2001). Entretanto, esta equação não consegue determinar a magnitude (módulo) e a direção da força  $F_i$  sobre os átomos do sistema, nem sua relação com as características químicas de cada molécula em estudo. Estes parâmetros serão calculados em função de mudanças de energia potencial ( $V$ ) entre a posição atual e a posição seguinte (aquela que representará o próximo passo da simulação) sobre cada átomo (coordenadas atômicas  $r_i$  a  $r_n$ ):

$$F_i = \frac{-\partial V(r_i, \dots, r_n)}{\partial r_i}$$

Esta superfície de energia potencial representa a energia de cada molécula, sendo descrita pelo denominado Campo de Força (Leach, 2001).

O campo de força pode ser definido como um conjunto de funções e parametrizações usadas em cálculos de mecânica molecular (de Sant'Anna, 2002) que definem as energias associadas a interações ligadas e não-ligadas. Por exemplo, entre átomos ligados covalentemente o estiramento de ligação e distorção do ângulo de ligação (tanto de valência quanto de diedro) de uma molécula é descrito em relação à sua conformação não tensionada (caracterizada pelos valores-padrão de comprimentos e de ângulos de ligação) (Figura 10). Juntamente a termos dedicados à descrição da interação entre átomos não-ligados, como forças eletrostáticas, ligação de hidrogênio e forças de van der Waals, dentre outros, os campos de força expressam o somatório das funções de energia potencial de cada átomo e calculam a energia dos sistemas em função das posições dos núcleos dos átomos (Leach, 2001).



**Figura 10:** Funções de energia que compõe campos de força. De cima para baixo estão representadas as equações que descrevem o estiramento de ligações químicas, ângulos de ligação, diedros e interações intermoleculares (eletrostáticas/Coulômbicas e van der Waals, também identificadas como potencial de Lennard-Jones). Adaptado de Verli, 2005, van Gunsteren *et al.*, 2006 e Fauth, 2009.

Existem diversos campos de força disponíveis para simulações de DM, destacando-se AMBER (Case *et al.*, 2005), CHARMM (MacKerell *et al.*, 1998), CVFF (Kitson & Hagler, 1988), TRIPOS (Clark *et al.*, 1989), OPLS (Jorgensen & Tirado-Rives, 1988) e GROMOS96 (van Gunsteren *et al.*, 1996). Apesar de a maioria dos campos de força atuais serem capazes de lidar com proteínas, existe uma carência de parâmetros específicos para outras classes de biomoléculas, como ácidos nucleicos, lipídios, compostos sintéticos e carboidratos (Pol-Fachin, 2009). Particularmente nesse último caso, a presença de diversos grupos altamente polares, a elevada flexibilidade e as diferenças nas propriedades eletrônicas que podem ocorrer durante modificações conformacionais e configuracionais dificultam a parametrização de campos de força para estes compostos (Pérez *et al.*, 1998). Nesse sentido, o Grupo de Bioinformática Estrutural vem se dedicando ao uso e

desenvolvimento de parâmetros para campos de força a fim de possibilitar a descrição e caracterização de carboidratos em geral, centrando-se na parametrização do campo de força GROMOS96, particularmente na versão implementada no pacote GROMACS, que é uma ferramenta gratuita e que não exige grandes capacidades computacionais (Pol-Fachin, 2009). Assim, estes parâmetros podem ser utilizados para descrever as porções sacarídicas presentes em glicoproteínas.

## 2 Objetivos

A partir do exposto, o presente trabalho tem por objetivo a caracterização estrutural e conformacional de prostaglandina endoperóxido sintases através de técnicas de modelagem molecular. Nesse contexto, o trabalho visa comparar o comportamento conformacional destas enzimas quando em sua forma glicosilada e não glicosilada a fim de determinar o papel estrutural da porção sacarídica nestas moléculas. Tal comparação pode ser obtida através de simulações de DM, empregando para a porção sacarídica uma estratégia de parametrização de carboidratos previamente descrita pelo grupo, compatível com o pacote de simulação GROMACS, tendo em vista sua elevada velocidade e baixo custo computacional. Ainda, o trabalho pretende analisar o comportamento das duas possíveis orientações do substrato AA no sítio ativo de ambas as enzimas. Espera-se, a partir dos resultados obtidos, contribuir para o entendimento da influência da porção sacarídica na conformação da enzima, além de contribuir na interpretação das diferentes orientações que substrato e inibidores desta enzima são capazes de adotar no sítio ativo de forma a subsidiar futuras estratégias de planejamento racional de novos moduladores de PGHSs. Dessa forma, as seguintes metas foram estabelecidas:

- Caracterização da influência da presença de glicosilação em PGHS-1 e PGHS-2;
- Caracterização da dinâmica de PGHS-1 e PGHS-2 ligadas ao seu substrato AA em suas diferentes orientações possíveis.

### 3 Metodologia

#### 3.1 Programas utilizados

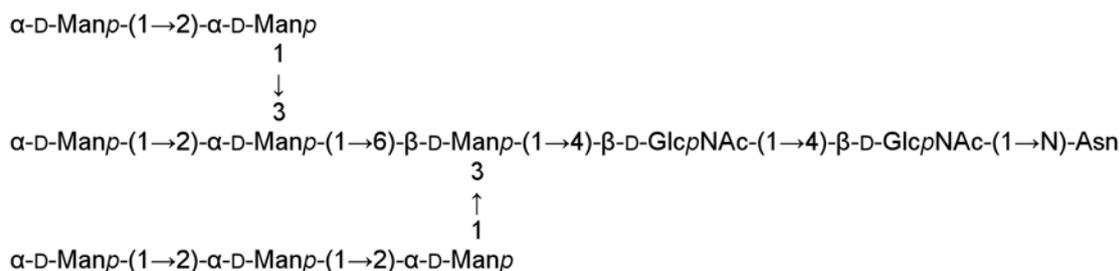
O presente trabalho foi baseado em uma série de metodologias de modelagem molecular, especialmente a DM. Os protocolos referentes a cada um destes métodos estão descritos em detalhes a seguir.

Os programas utilizados durante a presente dissertação incluem:

- Ferramentas de visualização de moléculas: VMD v1.8.6 (Humphrey *et al.*, 1996), PyMol v0.99 (DeLano, 2002) e MOLGEN (Schaftenaar & Noordik, 2000);
- Programa para simulações de dinâmica molecular: GROMACS v3.3.3 (van der Spoel *et al.*, 2005);
- Programa para geração de topologia: PRODRG *Beta* v2.5 (Schuettelkopf & van Aalten, 2004).

#### 3.2 Proposta do modelo de glicosilação pra PGHSs

Devido aos dados de RMN e espectrometria de massas indicarem uma heterogeneidade de glicofomas para os oligossacarídeos presentes nas PGHSs (Mutsaers *et al.*, 1985; Nemeth *et al.*, 2001), propusemos na presente dissertação um modelo de estrutura consenso a ser utilizado nos três sítios de glicosilação comuns às duas isoformas (Asn68, Asn144 e Asn410). Comparando-se as estruturas obtidas por espectrometria e aquelas depositadas no PDB (Tabelas S2 e S3, ver item 4.3 - material suplementar), foi necessário checar os dados cristalográficos para assegurar que os carboidratos estavam no estado anomérico apropriado e contendo as ligações glicosídicas corretas (Rao *et al.*, 1998). Para tanto, as estruturas do PDB foram submetidas ao software pdb-CARE (Lütteke & von der Lieth, 2004) para identificar e corrigir potenciais erros estruturais. O modelo putativo para PGHSs foi construído utilizando a ferramenta glycosciences (<http://www.glycosciences.de>) (Lütteke *et al.*, 2006) (Figura 11). As geometrias das ligações glicosídicas foram ajustadas de acordo com as conformações majoritárias obtidas previamente para os dissacarídeos simulados livres em solução (Fernandes *et al.*, 2010) (Tabela 1).



**Figura 11:** Estrutura sacarídica utilizada nos sítios de glicosilação das PGHSs (Asn68, Asn144 e Asn410).

Tabela 1: Preferências conformacionais dos dissacarídeos presentes nas PGHSs quando simulados livres em solução (Fernandes *et al.*, 2010)

Dissacarídeo	Ângulo (°)		
	$\phi$	$\psi$	$\Omega$
$\beta\text{-D-GlcpNAc-(1}\rightarrow\text{4)-}\beta\text{-D-GlcNAcp}$	$-80 \pm 24$	$115 \pm 16$	-
$\beta\text{-D-Manp-(1}\rightarrow\text{4)-}\beta\text{-D-GlcpNAc}$	$-160 \pm 21$	$90 \pm 18$	-
$\alpha\text{-D-Manp-(1}\rightarrow\text{3)-}\alpha\text{-D-Manp}$	$160 \pm 21$	$80 \pm 21$	-
$\alpha\text{-D-Manp-(1}\rightarrow\text{6)-}\beta\text{-D-Manp}$	$80 \pm 23$	$-160 \pm 30$	$60 \pm 17$
$\alpha\text{-D-Manp-(1}\rightarrow\text{2)-}\alpha\text{-D-Manp}$	$80 \pm 25$	$-135 \pm 36$	-

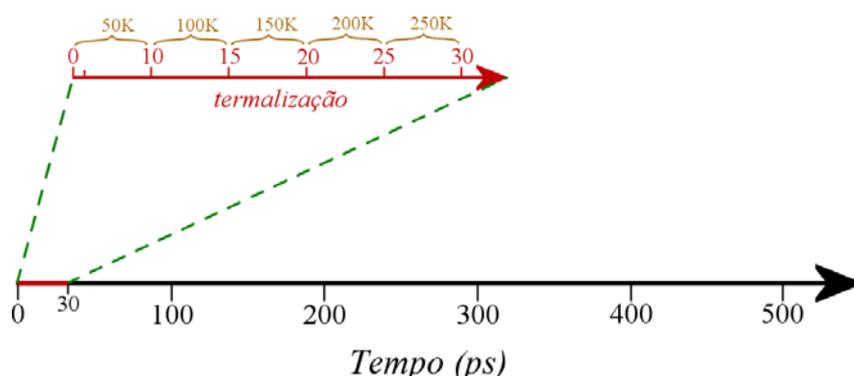
Dados expressos em média  $\pm$  desvio padrão de todos os valores de picos conformacionais principais obtidos a partir de dissacarídeos simulados livres em solução, que representam as geometrias de partida para estudos de DM de glicoproteínas.

### 3.3 Simulações por DM

#### 3.3.1 Protocolo de simulação

Todos os cálculos foram realizados usando o pacote de simulação GROMACS (Van der Spoel *et al.*, 2005) e o campo de força GROMOS96 (Van Gunsteren *et al.*, 1996). As enzimas foram solvatadas numa caixa triclinica utilizando condições periódicas de contorno, empregando o modelo de água SPC (Berendsen *et al.*,

1987). Contra-íons (cloreto ou sódio) foram adicionados, conforme a necessidade, de forma a neutralizar as cargas dos sistemas estudados. O método Lincs (Hess *et al.*, 1997) foi aplicado na restrição de ligações covalentes de forma a permitir um passo de integração de 2 fs, enquanto as interações eletrostáticas foram calculadas utilizando o método Particle-Mesh Ewald (Darden *et al.*, 1993). A temperatura e a pressão do sistema foram mantidas constantes através do acoplamento do soluto, íons e solvente a banhos externos de temperatura e pressão, utilizando constantes de acoplamento de, respectivamente,  $\tau = 0,1$  ps e  $\tau = 0,5$  ps (Berendsen *et al.*, 1984). A temperatura de referência de todos os sistemas foi ajustada para 310 K após a etapa de termalização, onde é feito um aquecimento gradativo de 50 até 310 K, em passos de 5 ps (conforme anexo, item 9.2), cada um elevando a temperatura em 50 K (Figura 12).



**Figura 12:** Esquema das etapas que compõe as simulações de DM a 310K. De 0ps a 30ps, uma série de etapas, de 5ps cada, aquece lentamente o sistema em incrementos de 50K. Após 30ps, o sistema está a 310K e se mantém nesta temperatura pelo resto da simulação (Verli, 2005).

### 3.3.2 Sistemas simulados

As isoformas PGHS-1 e PGHS-2 não glicosiladas na forma monomérica, bem como modelos de monômeros de PGHS-1 e -2 glicosilados foram submetidos ao protocolo descrito nos itens anteriores, totalizando um tempo de simulação de 50 ns para cada um dos quatro sistemas. Para tal foram utilizadas as estruturas cristalográficas de PGHS-1 de ovinos (PDB 1Q4G) e de PGHS-2 de murinos (PDB 1CVU). Para descrever a porção glicosídica destas enzimas foram usados parâmetros desenvolvidos no grupo. Estes parâmetros foram estabelecidos com

base na adequação do emprego de estados conformacionais de dissacarídeos, obtidos através de simulações de DM em solução aquosa, como guias para a construção de glicanas compondo glicoproteínas (conforme item 3.2). Tal capacidade preditiva foi analisada a partir da comparação de dados conformacionais obtidos por DM para tais dissacarídeos com a geometria dessas ligações glicosídicas obtidas por RMN.

As enzimas PGHS-1 e -2 glicosiladas, em sua forma monomérica, foram complexadas com as duas possíveis orientações do substrato AA, provenientes das estruturas cristalográficas depositadas no PDB (PDB 1DIY contendo a orientação catalítica e PDB 1CVU contendo a orientação alternativa), totalizando assim quatro sistemas, os quais foram submetidos ao protocolo de simulação descrito no item 3.3.1 com uma duração de 50 ns, utilizando restrições de posição de 1000 kJ/mol no MBD para mimetizar o ambiente da membrana. A topologia para o substrato AA foi gerada no programa PRODRG (Schuettelkopf & van Aalten, 2004) e as cargas foram corrigidas de maneira que o substrato se apresentasse na forma ionizada ( $\text{COO}^-$ ) (ver anexo, item 9.1).

## 4 Resultados

### 4.1 *Preâmbulo*

Os resultados obtidos durante a realização do trabalho serão apresentados a seguir. Estes trabalhos estão citados abaixo, assim como uma breve descrição sobre cada um.

i. Influência da glicosilação na conformação de PGHSs. Este item faz uma análise comparativa entre a dinâmica da enzima glicosilada e não glicosilada.

ii. Cláudia Lemelle Fernandes, **Liana Guimarães Sachett**, Laércio Pol-Fachin, Hugo Verli: GROMOS96 43a1 performance in predicting oligosaccharides conformational ensemble within glycoproteins. *Carbohydr. Res.*, **2010**, 345; 663-671.

Este trabalho analisa a adequação do uso de conformações obtidas por simulações de DM para dissacarídeos em solução aquosa como pontos de partida para a construção da estrutura de glicanas compondo glicoproteínas. A porção sacarídica das PGHS-1 e PGHS-2 glicosiladas foi analisada em comparação aos dissacarídeos simulados livres em solução.

iii. **Liana Guimarães Sachett**, Hugo Verli: Dynamics of Different Arachidonic Acid Orientations bound to Prostaglandin Endoperoxide Synthases. *Eur. J. Med. Chem.*, em preparação.

Este trabalho analisa a estabilidade e o comportamento conformacional das duas possíveis orientações encontradas em estruturas cristalográficas para o substrato das PGHSs, AA.

Os itens acima serão apresentados em detalhe a seguir.

## 4.2 Parte I

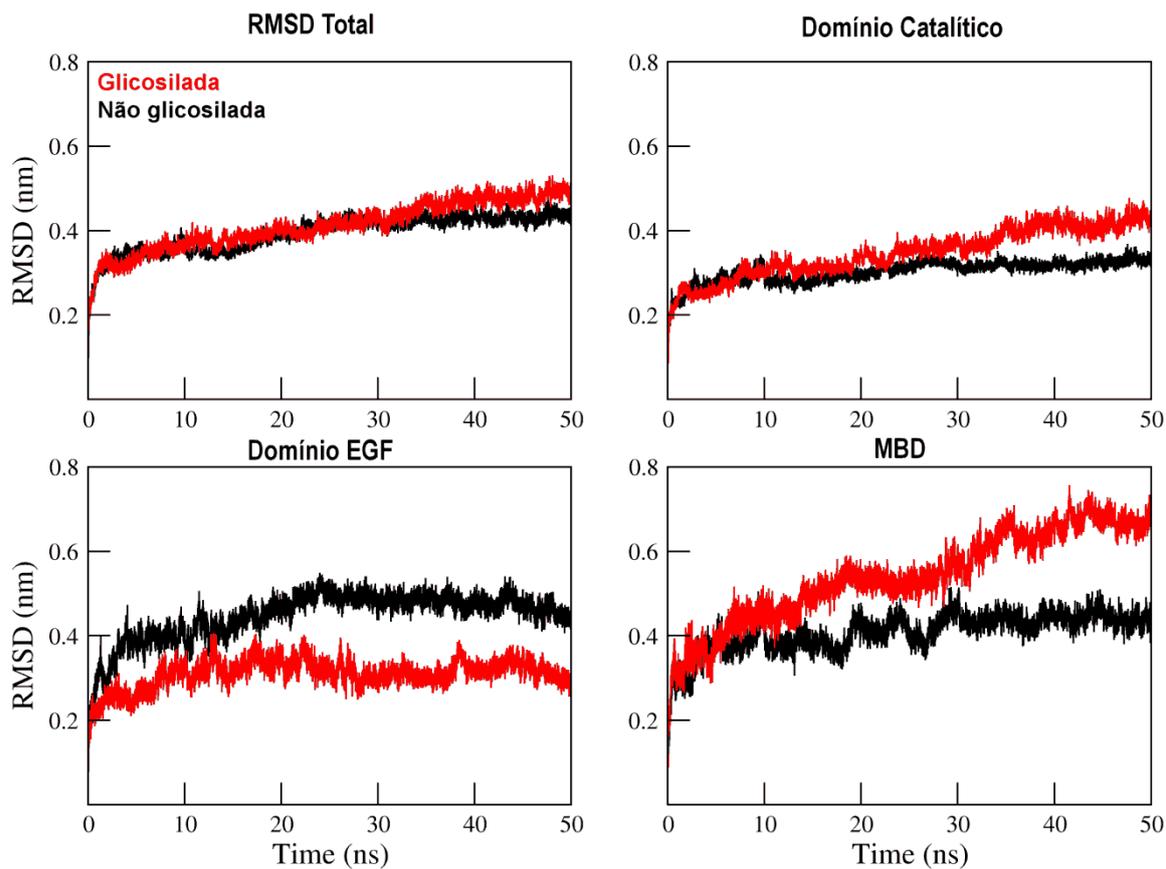
Nesta parte da presente dissertação, os seguintes sistemas foram simulados e analisados:

- PGHS-1 não glicosilada;
- PGHS-1 glicosilada;
- PGHS-2 não glicosilada;
- PGHS-2 glicosilada.

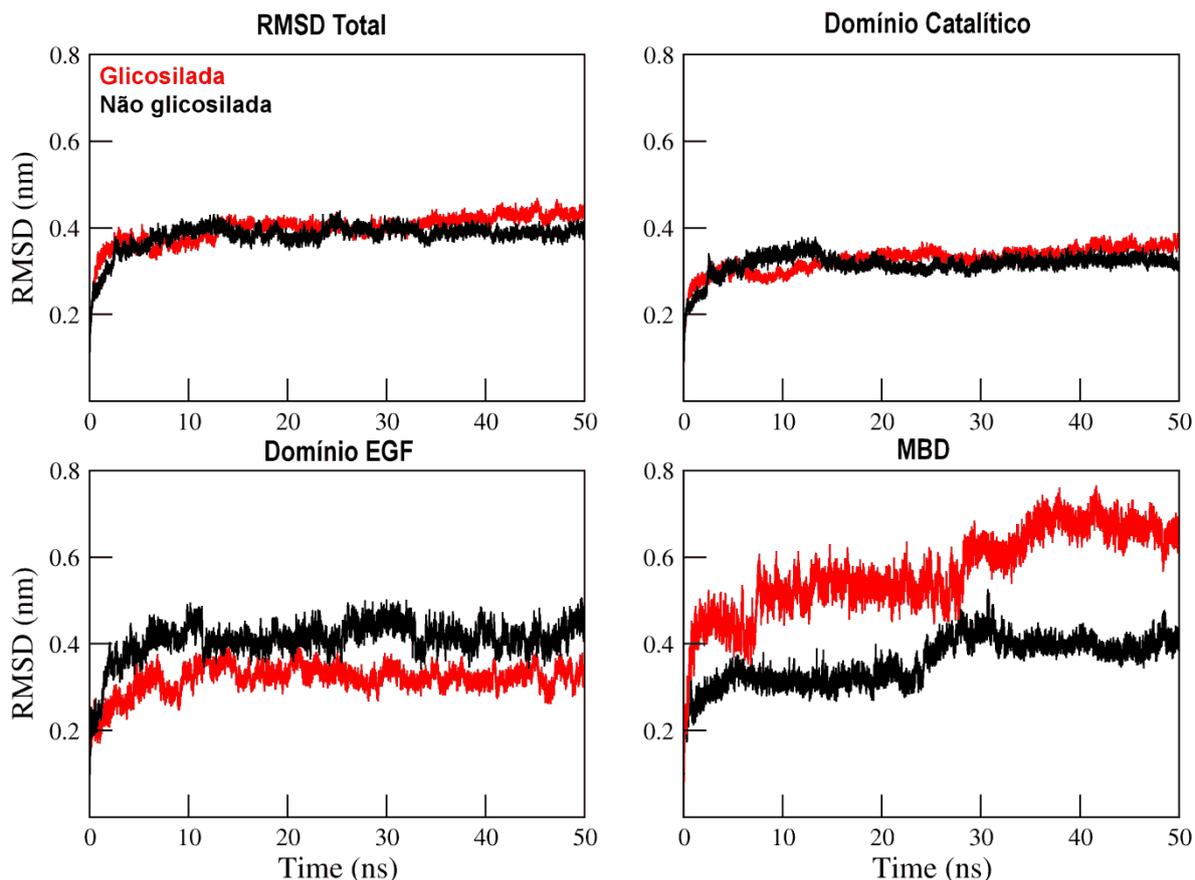
A análise comparativa das simulações das enzimas PGHS-1 e -2 em suas formas glicosilada e não glicosilada foi feita através da medida do desvio médio quadrático (*Root Mean Square Deviation* – RMSD), que fornece uma descrição da mudança da geometria molecular, em relação à estrutura cristalográfica de partida, como função do tempo e como uma média sobre todos os átomos da molécula, conforme a equação abaixo,  $r_i(t)$  é a posição do átomo  $i$  no tempo  $t$ .

$$RMSD(t_1, t_2) = \left[ \frac{1}{M} \sum_{i=1}^N m_i \|r_i(t_1) - r_i(t_2)\|^2 \right]^{1/2}$$

Adicionalmente, a dimensão da mudança da posição dos átomos da proteína, a partir da estrutura cristalina, oferece um indicativo da flexibilidade protéica e mudanças conformacionais a cada passo da simulação. Assim, de forma a caracterizar o papel de cada domínio na dinâmica das proteínas, as análises de RMSD foram realizadas para as ambas isoformas de PGHS, tanto não glicosiladas quanto glicosiladas, considerando todos os átomos (RMSD Total) e também seus domínios separadamente (Figuras 13 e 14).



**Figura 13:** Análise de RMSD para o monômero de PGHS-1 nas formas não glicosilada (preto) e glicosilada (vermelho). Os sítios de glicosilação estão localizados nos domínios EGF (Asn68) e catalítico (Asn144 e Asn410).

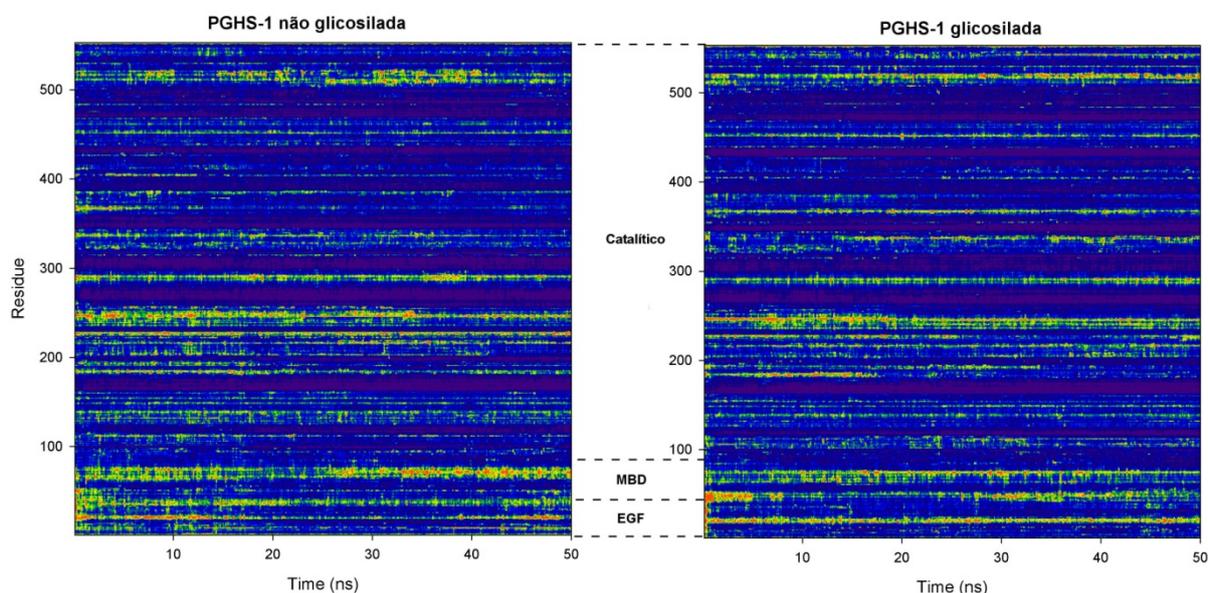


**Figura 14:** Análise de RMSD para o monômero de PGHS-2 nas formas não glicosilada (preto) e glicosilada (vermelho). Os sítios de glicosilação estão localizados nos domínios EGF (Asn68) e catalítico (Asn144 e Asn410).

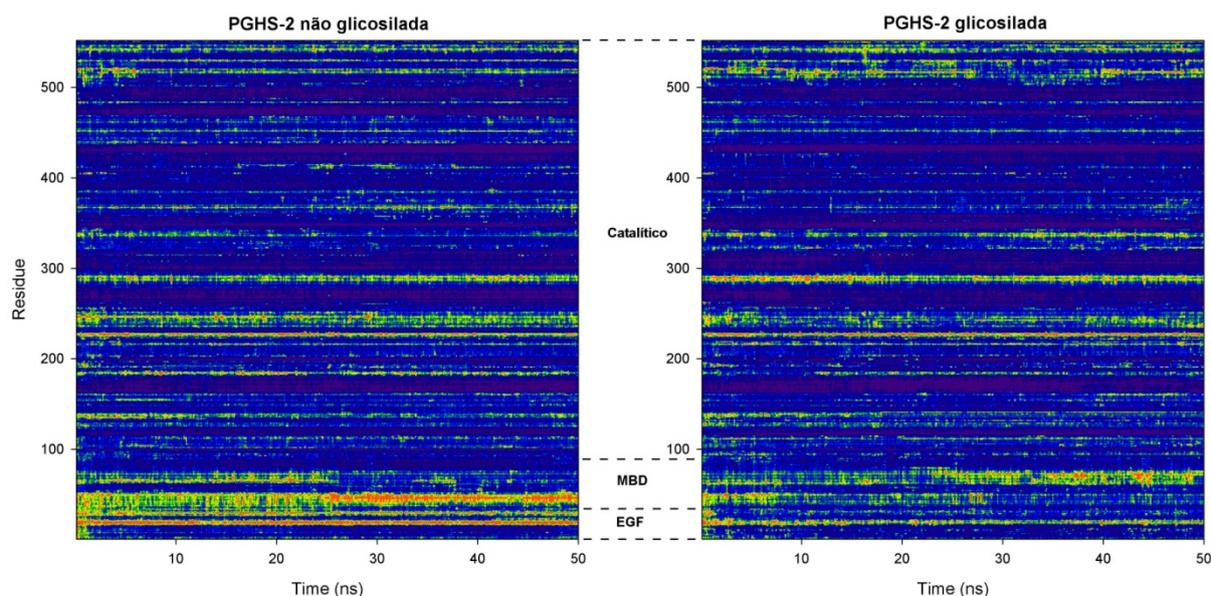
Os dados obtidos sugerem que a presença da glicosilação não altera significativamente a flexibilidade total do sistema, tampouco do domínio catalítico. Há, contudo, algumas interferências locais em domínios específicos, como no domínio EGF, onde foi observada diminuição da flexibilidade quando a enzima está glicosilada e ambas as isoformas da enzima. Este tipo de alteração já havia sido descrito para este domínio em trabalhos prévios do nosso grupo (Pol-Fachin *et al.*, 2009), como no fator VII de coagulação. Adicionalmente, dados obtidos por outros grupos (Perera *et al.*, 2002) têm demonstrado este tipo de efeito de enrijecimento do esqueleto peptídico na presença de glicanas, potencialmente pelo aumento do número de interações polares entre soluto e solvente através dos resíduos de carboidratos adicionados. O aumento de flexibilidade no domínio MBD quando a enzima está glicosilada, em contrapartida, não foi atribuído a nenhuma alteração

específica, e pode não refletir uma situação verdadeira, já que não foram utilizadas restrições de posição neste domínio para mimetizar um ambiente de membrana.

A principal limitação das análises de RMSD se refere à eliminação da informação por resíduo em detrimento da informação temporal. De forma a contornar esta limitação, nosso grupo de pesquisas desenvolveu uma estratégia alternativa, onde a informação de cada resíduo não é perdida durante a análise, oferecendo também indicativos da variação da conformação e flexibilidade protéica como função do tempo, estratégia esta denominada tRMSF. Estas análises foram também realizadas para ambas as isoformas de PGHS-1 a fim de caracterizar as regiões de maior flexibilidade em cada um dos domínios (Figuras 15 e 16).



**Figura 15:** Dados de tRMSF para PGHS-1 nas formas não-glicosilada e glicosilada. Áreas em vermelho/laranja representam as regiões de maior flexibilidade, enquanto áreas azuis representam regiões estáveis. Os três domínios da enzima tem suas regiões delimitadas pelas linhas tracejadas.



**Figura 16:** Dados de tRMSF para PGHS-2 nas formas não-glicosilada e glicosilada. Áreas em vermelho/laranja representam as regiões de maior flexibilidade, enquanto área azuis representam regiões estáveis. Os três domínios da enzima tem suas regiões delimitadas pelas linhas tracejadas.

A partir das análises de RMSF, pode-se observar que o domínio EGF apresenta uma diminuição da flexibilidade geral quando as enzimas estão glicosiladas, principalmente na PGHS-2.

A parte sacarídica de glicoproteínas é, em geral, analisada separadamente da parte protéica. Estas análises encontram-se no trabalho apresentado no item seguinte.

### 4.3 Parte II

## **GROMOS96 43a1 performance in predicting oligosaccharides conformational ensemble within glycoproteins**

Cláudia Lemelle Fernandes, **Liana Guimarães Sachett**, Laercio Pol-Fachin, Hugo Verli

*Carbohydrate Research*, **2010**, 345; 663-671

Considerando a dificuldade da caracterização da estrutura e conformação de glicoproteínas, juntamente com suas cadeias sacarídicas correspondentes, e considerando a adequação dos procedimentos empregados pelo grupo na reprodução do equilíbrio conformacional de glicanas biológicas, o presente trabalho objetivou avaliar a adequação do emprego de estados conformacionais de dissacarídeos, obtidos através de simulações de DM em solução aquosa, como geometrias de referencia para a construção de glicanas compondo glicoproteínas. Tal capacidade preditiva foi analisada a partir da comparação de dados conformacionais obtidos por DM para uma série de dissacarídeos com a geometria dessas ligações glicosídicas, em glicoproteínas, obtidas por RMN.

Os resultados obtidos indicam que a prevalência dos estados conformacionais em solução foi semelhante entre os dissacarídeos isolados e compondo as glicoproteínas empregadas como modelo, sugerindo que o arcabouço protéico, ao menos nos casos estudados, não parece ser capaz de induzir o aparecimento de novos estados conformacionais. Isto sugere que a abordagem proposta pode se constituir em uma promissora estratégia na construção de estruturas completas de glicoproteínas, na ausência de dados experimentais prévios e, assim, contribuir no entendimento do papel destes glicoconjugados em sistemas biológicos. Por fim, esta estratégia foi aplicada na construção de modelos de PGHSs glicosilados.



Contents lists available at ScienceDirect

## Carbohydrate Research

journal homepage: [www.elsevier.com/locate/carres](http://www.elsevier.com/locate/carres)

## GROMOS96 43a1 performance in predicting oligosaccharide conformational ensembles within glycoproteins

C. L. Fernandes<sup>a</sup>, L. G. Sachett<sup>a</sup>, L. Pol-Fachin<sup>a</sup>, H. Verli<sup>a,b,\*</sup><sup>a</sup> Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil<sup>b</sup> Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Av Ipiranga 2752, Porto Alegre 90610-000, RS, Brazil

## ARTICLE INFO

## Article history:

Received 7 August 2009

Received in revised form 14 December 2009

Accepted 18 December 2009

Available online 28 December 2009

## Keywords:

Glycoproteins

Molecular dynamics

Carbohydrates

Glycosidic linkage

GROMOS96

## ABSTRACT

In previous work [Pol-Fachin, L.; Fernandes, C. L.; Verli, H.; *Carbohydr. Res.* **2009**, *344*, 491–500], we had demonstrated that GROMOS96 43a1 force field and Löwdin HF/6-31G<sup>\*\*</sup>-derived atomic charges, adequately represent a glycoprotein's conformational ensemble in aqueous solutions, taking as the starting geometries NMR-determined structures. Based on such data, the present work intends to evaluate the use of the main solution conformations of isolated disaccharides, to build the carbohydrate moiety of glycoproteins, for which no previous experimental information is available. The observed results suggested that the entire glycoprotein scaffold appears unable to promote major modifications in the conformational behavior of glycosidic linkages. Additionally, when compared to energy contour plots, the results support the use of solution ensembles, to refine vacuum conformations of carbohydrate databases in the assembling of glycoproteins 3D structures. Finally, such approach is applied to build a full glycosylated model for COX-1 and COX-2 enzymes.

© 2009 Elsevier Ltd. All rights reserved.

### 1. Introduction

Carbohydrates constitute a group of biological polymers of enormous structural and conformational complexity, presenting several possible monomeric units, adopting different conformational states, linked in diverse ways and presenting a variable degree of branching. However, while the identification of saccharide units being attached to proteins and their connectivity is experimentally accessible, their 3D information is not easily obtained, which increases the difficulties in elucidating carbohydrate functions in glycoproteins.<sup>1–3</sup>

Oligosaccharide 3D structures and conformation are mainly determined by their glycosidic linkage geometries, as a function of composition, linkage type, and solvation.<sup>3</sup> As a consequence, an approach usually employed to describe the conformational preferences of disaccharides consists in the construction of energy contour maps.<sup>4–8</sup> Based on the most stable geometries described by such plots, it is possible to infer the structure of a complete glycoprotein, in which the glycan moiety is built on a previously determined 3D protein structure.<sup>2</sup>

The potential contribution of such theoretic strategy may be estimated, considering the difficulties of some time-honored

methods in describing the structure of glycans and glycoproteins. In X-ray crystallography, the oligosaccharide is usually incomplete, as its high flexibility may impair both the crystallization and acquisition of electron densities.<sup>9–11</sup> Even when these data are obtained, X-ray structures of carbohydrates are usually of poor quality, as demonstrated by a recent survey of PDB entries containing oligosaccharides, which suggests that about one-third of carbohydrate data comprise significant errors in relation to their stereochemistry, nomenclature, and consistency with electron density maps.<sup>12,13</sup> In NMR, one of the main challenges lies on the difficulties in resolving NOE signals, mainly due to the high flexibility of glycans, or to the systems size.<sup>2,9,10</sup> Such flexibility produces a complex conformational ensemble, with the possibility of co-existence of a number of simultaneous conformers in a solution.<sup>14–17</sup>

In this context, the current work intends to evaluate the adequacy of using main solution conformations of the isolated disaccharides, to build the carbohydrate moiety of glycoproteins, in the absence of previous experimental information on a carbohydrate's 3D structure and conformation. Accordingly, the obtained data are applied in building a complete model for COX-1 and COX-2 enzymes, in their fully glycosylated forms. Such strategy is expected to accurately represent glycoproteins in biological solutions, offering a refinement to the currently employed proceedings, and so contributing in the understanding of carbohydrates biological roles at the atomic level.

\* Corresponding author. Tel.: +55 51 3308 7770; fax: +55 51 3308 7309.  
E-mail address: [hverli@cbiot.ufrgs.br](mailto:hverli@cbiot.ufrgs.br) (H. Verli).

## 2. Experimental

### 2.1. Computational methods

#### 2.1.1. Nomenclature and topology

The analyzed disaccharide units represent carbohydrate moieties from NMR-determined glycoproteins: the human complement regulatory protein CD59 (PDB ID 1CDR and 1CDS),<sup>18</sup> the adhesion domain of human CD2 (PDB ID 1GYA)<sup>19</sup> and the  $\alpha$ -subunit of the human chorionic gonadotropin (PDB ID 1HD4)<sup>20</sup> (Fig. 1). Based on the glycoproteins' data, eight different disaccharides, composed of *N*-acetyl- $\beta$ -D-glucosamine ( $\beta$ -D-Glc<sub>1</sub>NAc),  $\beta$ -D-mannose ( $\beta$ -D-Man<sub>1</sub>p), *L*-fucose (*L*-Fuc<sub>1</sub>p), and  $\beta$ -D-galactose ( $\beta$ -D-Gal<sub>1</sub>p), bounded by  $\alpha/\beta(1\rightarrow2)$ ,  $\alpha(1\rightarrow3)$ ,  $\beta(1\rightarrow4)$ , and  $\alpha(1\rightarrow6)$  linkages, were evaluated.

The IUPAC nomenclature recommendations and symbols<sup>21</sup> have been adopted. The relative orientation of a pair of contiguous carbohydrate residues was described, for different types of linkages, by two or three torsional angles at the glycosidic linkage. For a (1→*X*) linkage, where '*X*' is '2', '3', '4', or '6' for the (1→2), (1→3), (1→4), or (1→6) linkages, respectively, the  $\phi$  and  $\psi$  were defined as shown below:

$$\phi = \text{O5-C1-OX-CX'}$$

$$\psi = \text{C1-OX-CX-C(X-1)}$$

For a (1→6) linkage, the  $\omega$  was defined as

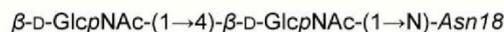
$$\omega = \text{O6-C6-C5-C4}$$

Each of the disaccharide structures was built in Molden<sup>22</sup> and their respective topologies were obtained from the PRODRG server.<sup>23</sup> The HF/6-31\*\*<sup>+</sup>-derived Löwdin charges were employed.<sup>24,25</sup> Improper dihedrals were included to preserve the hexopyranose conformations in accordance with their expected form in aqueous solution: <sup>4</sup>C<sub>1</sub> for  $\beta$ -D-Glc<sub>1</sub>NAc, <sup>1</sup>C<sub>4</sub> for *L*-Fuc<sub>1</sub>p, <sup>4</sup>C<sub>1</sub> for  $\beta$ -D-Man<sub>1</sub>p, and <sup>4</sup>C<sub>1</sub> for  $\beta$ -D-Gal<sub>1</sub>p. These topologies were employed in molecular dynamics (MD) simulations and were analyzed using the GROMACS simulation suite<sup>26</sup> and GROMOS96 43a1 force field,<sup>27</sup> as described previously.<sup>17,24,25,28,29</sup>

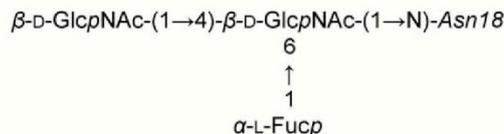
#### 2.1.2. Calculation of energy contours maps

The conformational description of each disaccharide was performed through a calculation of relaxed contour maps around the glycosidic linkages, rotating the torsion angles from  $-180^\circ$  to  $180^\circ$  in steps of  $30^\circ$ , with a total of 144 conformers for each linkage. This was performed using a  $334.8 \text{ kJ mol}^{-1}$  constant force, to make a restriction only to the  $\phi$  and  $\psi$  proper dihedrals during energy minimization, in each of the above-mentioned values, thus allowing the search of the conformational space associated with the disaccharide. Then, using the minimized output conformation,

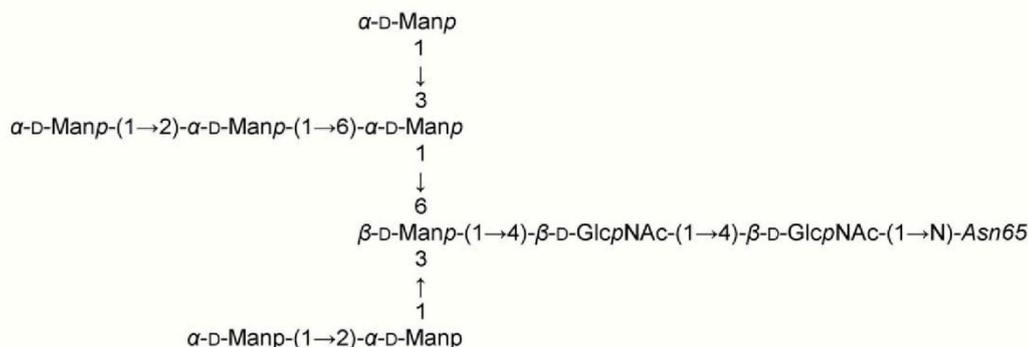
#### (A) Human complement regulatory protein CD59 (PDB ID 1CDS)



#### (B) Human complement regulatory protein CD59 (PDB ID 1CDR)



#### (C) Adhesion domain of human CD2 (PDB ID 1GYA)



#### (D) Human chorionic gonadotropin (PDB ID 1HD4)

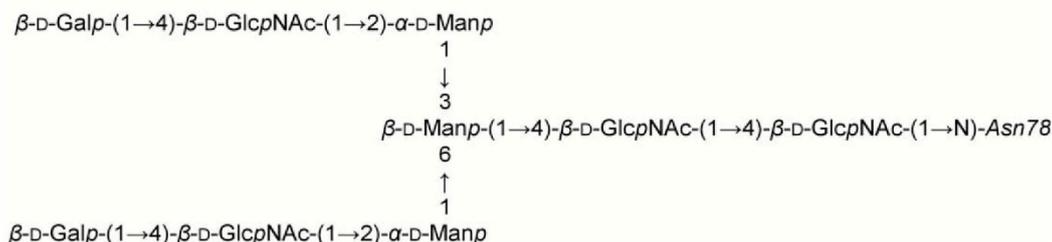


Figure 1. Schematics of the NMR glycan structures used to construct the simulated disaccharides.

a series of MD simulations were performed for 20 ps at 10 K, with an integration step of 0.5 fs to further support the search for minimum-energy conformations. In this process, the rotatable exocyclic groups were allowed to freely search for minimum-energy orientations.<sup>29,30</sup> The relative stabilities of each conformation, obtained from the 10 K MD simulations, based on GROMOS96 force field<sup>27</sup> and HF/6-31G\*\* Löwdin atomic charges, were used to construct relaxed energy contour plots, describing the conformation of each glycosidic linkage, as shown in Figure 2.

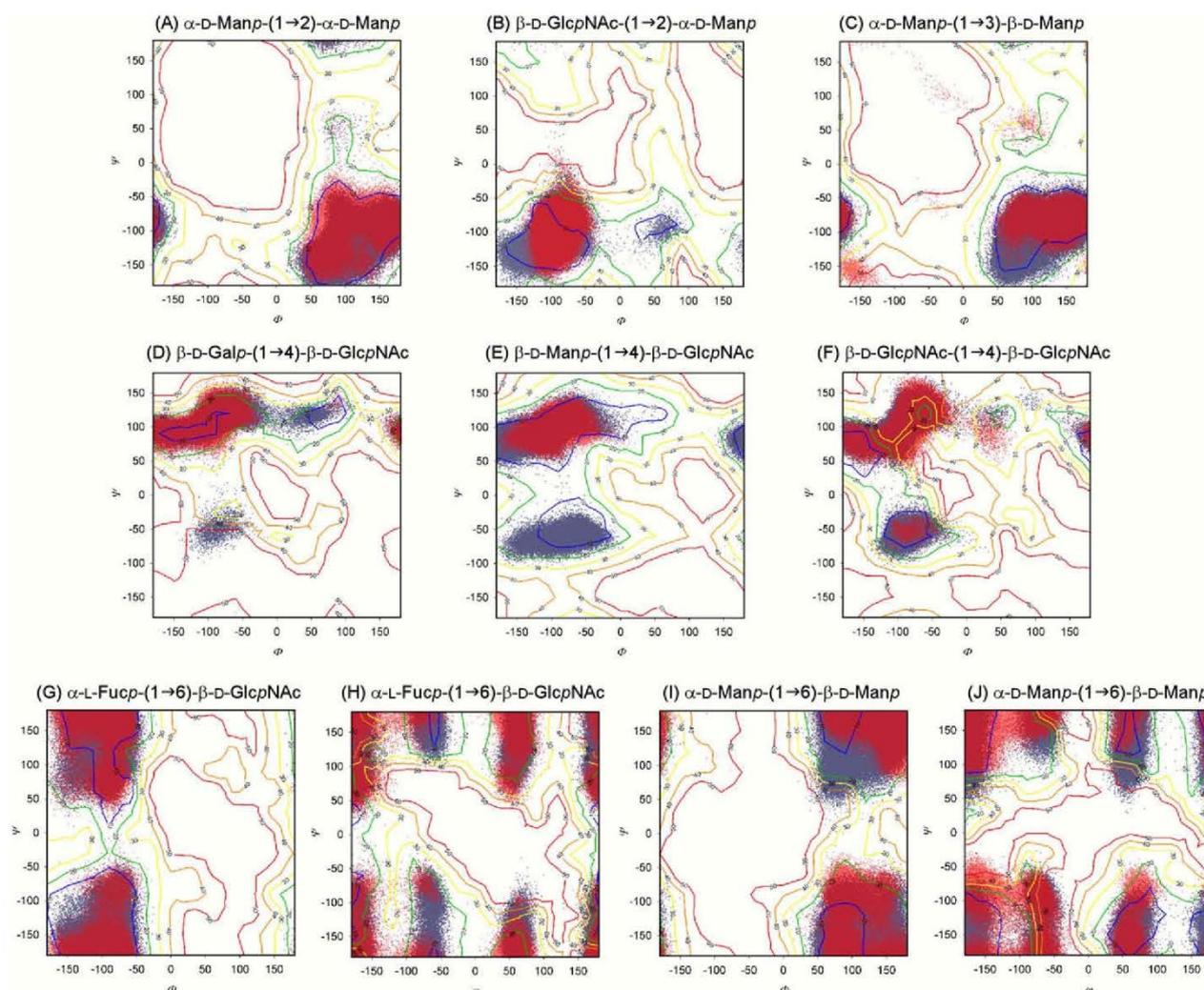
### 2.1.3. MD simulations

The minimum-energy conformations of each of the disaccharides, as obtained from their contour plots, were then used as the starting conformations for MD simulations in an aqueous solution, using the GROMACS simulation suite.<sup>26</sup> Each pair of angles describing a vacuum minimum-energy conformation was used to build the disaccharide conformers that would be further submitted to MD simulations. The disaccharides were solvated in a triclinic box using periodic boundary conditions and SPC water model.<sup>31</sup> The Lincs method<sup>32</sup> was applied to constrain covalent bond lengths, allowing an integration step of 2 fs, after an initial energy minimization using Steepest Descents algorithm. Electrostatic interactions were calculated using the Particle Mesh Ewald meth-

od.<sup>33</sup> Temperature and pressure were kept constant by coupling disaccharides, ions, and solvent to external temperature and pressure baths, with coupling constants of  $\tau = 0.1$  and 0.5 ps,<sup>34</sup> respectively. The dielectric constant was treated as  $\epsilon = 1$ , and the reference temperature was adjusted to 310 K. The systems were slowly heated from 50 to 310 K, in steps of 5 ps, each one increasing the reference temperature by 50 K. Each simulation was extended to 0.1  $\mu$ s.

### 2.1.4. Glycosylated cyclooxygenase (COX) models

The murine COX-2 and ovine COX-1 were retrieved from PDB IDs 1CVU and 1Q4G, respectively. These structures were isolated to their monomeric forms, in which, the N-glycosylation sites (Asn68, Asn144, and Asn410) were filled with a putative model for their glycan moieties, represented in Figure 4, using glycoscience modeling tools.<sup>35–37</sup> This model was proposed based on a consensus between previous biochemical, crystallographic and mass spectrometry data.<sup>38,39</sup> The composition evidenced by spectrometric results was compared to structures deposited in PDB and further refined by biochemical evidence, supporting the identification of a proper anomeric state for each residue and type of linkage for carbohydrate residues. The PDB structures were also submitted to pdb-care software,<sup>40</sup> in order to identify and correct



**Figure 2.** Relaxed contour plots for the studied glycosidic linkages, in which the energy maps are shown, every 10 kJ mol<sup>-1</sup>, from 10 to 50 kJ mol<sup>-1</sup>, superimposed with the results of MD simulations in solution. Blue dots indicate the conformational behavior of disaccharide units, while red dots indicate the conformational behavior of a given disaccharide in a whole glycoprotein environment.

potential structural errors. These so-obtained models for glycosylated COX-1 and COX-2 had their glycosidic linkage geometries adjusted to the main conformations for each linkage, based on their relative abundance in the isolated disaccharides in water, as described in Table 1. Such models were then submitted to 50 ns MD simulations in an aqueous solution, using the GROMACS package<sup>26</sup> and GROMOS96 43a1 force field,<sup>27</sup> following the simulation proceedings described above (Section 2.1.3). As previous results identified that detergent molecules do not appear to be essential to stabilize the membrane-binding domain during simulations<sup>41</sup> and that there is no glycosylation site within this domain, then the proteins were simulated without restraints in order to observe the dynamics of the catalytic and EGF domains, which in turn are glycosylated.

### 3. Results and discussion

#### 3.1. Simulation systems

The strategy employed to evaluate the adequacy of using conformations of disaccharides, derived from MD as guiding geometries, for building the glycan parts of glycoproteins without experimental restraints from NMR, or X-ray crystallography, was based on the following steps: (1) building of disaccharides commonly found in glycoproteins; (2) calculation of their vacuum contour plots in order to obtain minimum-energy conformations,

which were to be further sampled in solution; (3) MD simulations of each disaccharide's minimum-energy conformation in aqueous solutions, under an explicit solvent model; and (4) to compare the obtained conformational samplings, in both vacuum and solution, to previous MD and NMR data of glycoproteins. The studied disaccharides included:  $\alpha$ -D-Manp-(1→2)- $\alpha$ -D-Manp,  $\alpha$ -D-Manp-(1→3)- $\beta$ -D-Manp,  $\alpha$ -D-Manp-(1→6)- $\beta$ -D-Manp,  $\beta$ -D-GlcpNAc-(1→2)- $\alpha$ -D-Manp,  $\beta$ -D-Galp-(1→4)- $\beta$ -D-GlcpNAc,  $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc,  $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc, and  $\alpha$ -L-Fucp-(1→6)- $\beta$ -D-GlcpNAc. The so-obtained data were then applied to describe the complete glycosylated structures of COX isoforms, for which until now, there were no 3D data to illustrate the application of the proposed protocol.

#### 3.2. Vacuum conformational analysis

There is a considerable amount of data on disaccharide conformational preference in carbohydrate databases.<sup>35–37,42</sup> The Glycosciences portal<sup>35–37</sup> has a large amount of data based on MM3<sup>43,44</sup> force field, thus offering a highly detailed representation of carbohydrates conformational features,<sup>45,46</sup> at the least, in gas phase.<sup>47,48</sup> Unfortunately, such a force field has not been as widely used, to study biological systems, as more specifically parameterized force fields, such as CHARMM, AMBER, GROMOS, and OPLS, amongst others.<sup>47,49</sup> In this context, we proceeded to the calculation of relaxed maps for a series of eight disaccharides (Fig. 2), fre-

**Table 1**  
Conformational preferences of disaccharides as observed from MD simulations<sup>a</sup>

Disaccharide	Relative abundances (%)	Angle (°)		
		$\phi$	$\psi$	$\omega$
$\alpha$ -D-Manp-(1→2)-Manp	<b>62</b>	<b>80 ± 25</b>	<b>-135 ± 36</b>	–
	38	160 ± 20	-100 ± 28	–
$\beta$ -D-GlcpNAc-(1→2)-Manp	<b>96</b>	<b>-90 ± 35</b>	<b>-95 ± 32</b>	–
	4	60 ± 25	-100 ± 24	–
$\alpha$ -D-Manp-(1→3)-Manp	33	80 ± 24	-85 ± 28	–
	23	-130 ± 18	–	–
	<b>44</b>	<b>160 ± 21</b>	<b>-80 ± 21</b>	–
$\beta$ -D-Galp-(1→4)-GlcpNAc	1.5	<b>-80 ± 29</b>	-50 ± 23	–
	<b>86</b>	<b>120 ± 21</b>	–	–
	3	50 ± 29	–	–
$\beta$ -D-Manp-(1→4)-GlcpNAc	10	-150 ± 17	90 ± 16	–
	23	-70 ± 30	-55 ± 19	–
	33	110 ± 22	–	–
	<b>44</b>	<b>-160 ± 21</b>	<b>90 ± 18</b>	–
$\beta$ -D-GlcpNAc-(1→4)-GlcpNAc	<b>56</b>	<b>-80 ± 24</b>	<b>115 ± 16</b>	–
	19	-90 ± 28	-60 ± 20	–
	15	-100 ± 18	90 ± 14	–
	11	-150 ± 17	–	–
	14	<b>-90 ± 35</b>	120 ± 27	60 ± 16
$\alpha$ -D-Fucp-(1→6)-GlcpNAc	16	-160 ± 32	–	–
	12	160 ± 34	-60 ± 17	–
	3	-90 ± 24	–	–
	17	90 ± 23	<b>-180 ± 17</b>	–
	<b>38</b>	<b>180 ± 46</b>	–	–
	<b>27</b>	<b>80 ± 23</b>	<b>-160 ± 30</b>	<b>60 ± 17</b>
	6	–	110 ± 29	–
	6	150 ± 18	-160 ± 31	–
$\alpha$ -D-Manp-(1→6)-Manp	7	–	110 ± 24	–
	5	80 ± 22	-90 ± 21	-70 ± 19
	6	–	170 ± 34	–
	1.5	150 ± 19	-90 ± 23	–
	5	–	170 ± 30	–
	11	80 ± 24	-125 ± 21	175 ± 19
	11	–	-180 ± 24	–
	4	–	90 ± 21	–
	3	150 ± 19	-125 ± 23	–
6	–	-180 ± 24	–	
1	–	90 ± 20	–	

<sup>a</sup> The most abundant conformations are marked as bold and were used, as necessary, to construct the fully glycosylated COX enzymes models (see details in the text).

quently found in glycoproteins, under GROMOS96 43a1 force field set of parameters. The eight disaccharides analyzed provided 10 maps, including  $\phi$ - $\psi$  and  $\omega$ - $\psi$  plots.

Similarities between the conformational patterns of  $\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp and  $\alpha$ -D-Manp-(1 $\rightarrow$ 3)- $\beta$ -D-Manp may be observed, as reported,<sup>50–52</sup> with one broad minimum energy region in both  $\phi$  and  $\psi$  dimensions (Fig. 2A and C). Likewise, the  $\beta$ (1 $\rightarrow$ 4) linkages are presented from two to three minimum energy regions, being in the northwest region of the plot, from  $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc,  $\beta$ -D-Manp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc, and  $\beta$ -D-GlcpNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc, and also in the southwest quadrant for the latter two disaccharides. While the inversion in stoichiometry at carbon atoms 2 and 4 in Gal, regarding to Man, appears to increase the rigidity of the linkage, the  $\beta$ (1 $\rightarrow$ 4) disaccharides are globally more flexible than the  $\alpha$ (1 $\rightarrow$ 2) and  $\alpha$ (1 $\rightarrow$ 3) linkages, presenting more than one single energy minima, in agreement with the previous data.<sup>53,54</sup> Additionally, the conformational restriction promoted by the presence of *N*-acetyl group in position C-2, maintaining the location of the minima, was already described.<sup>55</sup>

Concerning the behavior of (1 $\rightarrow$ 6) linkages, the preference of the  $\omega$  dihedrals in the region of  $-60^\circ$ ,  $60^\circ$ , and  $180^\circ$ , usually reported in the literature,<sup>56–59</sup> was observed accordingly for both  $\alpha$ -L-Fucp-(1 $\rightarrow$ 6)- $\beta$ -D-GlcpNAc and  $\alpha$ -D-Manp-(1 $\rightarrow$ 6)- $\beta$ -D-Manp disaccharides (Fig. 2H and J, respectively). Furthermore, at least in vacuum, the steric hindrance promoted by the methyl group, at Fuc carbon 5 and the *N*-acetyl group at GlcNAc carbon 2, reduced the extension and the depth of minimum energy regions when compared to Man-containing disaccharides. As a general feature, all disaccharides presented well-defined minimum energy regions, in agreement with those deposited in the Glycosciences portal,<sup>35–37</sup> indicating a similar conformational description of the main energy minima between MM3 and GROMOS96 43a1 force fields,<sup>27</sup> for the evaluated systems. Therefore, such results indicate that the latter may be used in carbohydrates conformational description with a reasonable level of accuracy.

### 3.3. Solution conformational sampling

Considering that the explicit inclusion of solvent molecules is already described as able to reveal a distinct set of conformers when compared to calculations in its absence,<sup>60</sup> each energy minima from maps (Fig. 2) was submitted to 100 ns MD simulations in order to obtain a description of the solvent influence over each disaccharide (Supplementary data). In fact, such influence may be initially observed on the use of different energy minima, as starting geometries, for distinct sets of simulations. Independent of the starting minima, the obtained conformational ensembles were equivalent, suggesting a major role of solvent on a disaccharide's conformational equilibrium. Also, it indicates the capability of the employed time scale, to effectively describe the conformational transitions, relevant to each system.

As presented in Figure 2 and Table 1, the conformational preference of most of the simulated disaccharides remained located around their vacuum minimum energy regions. However, upon inclusion of a solvent, some less favorable vacuum regions were also occupied, adjacent to minimum energy regions, as well as some high energy regions from the vacuum map, indicating the influence of the solvent in the stabilization of additional regions of the energy maps. For example, the two energy minima shown by the  $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc disaccharide in vacuum, increased to three main regions in the conformational space populated in solution (Fig. 2D and Table 1). Thus, it became more similar to the remaining  $\beta$ (1 $\rightarrow$ 4) simulated disaccharides (Fig. 2E and F), as expected.<sup>36,54,55</sup> Similarly, the  $\alpha$ -D-Manp-(1 $\rightarrow$ 6)- $\beta$ -D-Manp presents an additional stabilization of the  $\omega$  angle at  $-60^\circ$  as a consequence of solvent inclusion (Fig. 2J), acquiring a confor-

mational preference similar to that shown by  $\alpha$ -L-Fucp-(1 $\rightarrow$ 6)- $\beta$ -D-GlcpNAc (Fig. 2H).

On the other hand, some minimum energy regions on the contour plots were not consistently populated in solution, as could be observed for  $\beta$ -D-GlcpNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp (Fig. 2B) and  $\beta$ -D-Manp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc (Fig. 2E). These observations indicate that, while the vacuum conformational maps indeed generated good starting points for the representation of carbohydrates in solution, the conformational space described by the two strategies is not necessarily equivalent. Most importantly, such data reinforce the key importance in using explicit solvent effects when computationally studying biological carbohydrates and glycoconjugates.

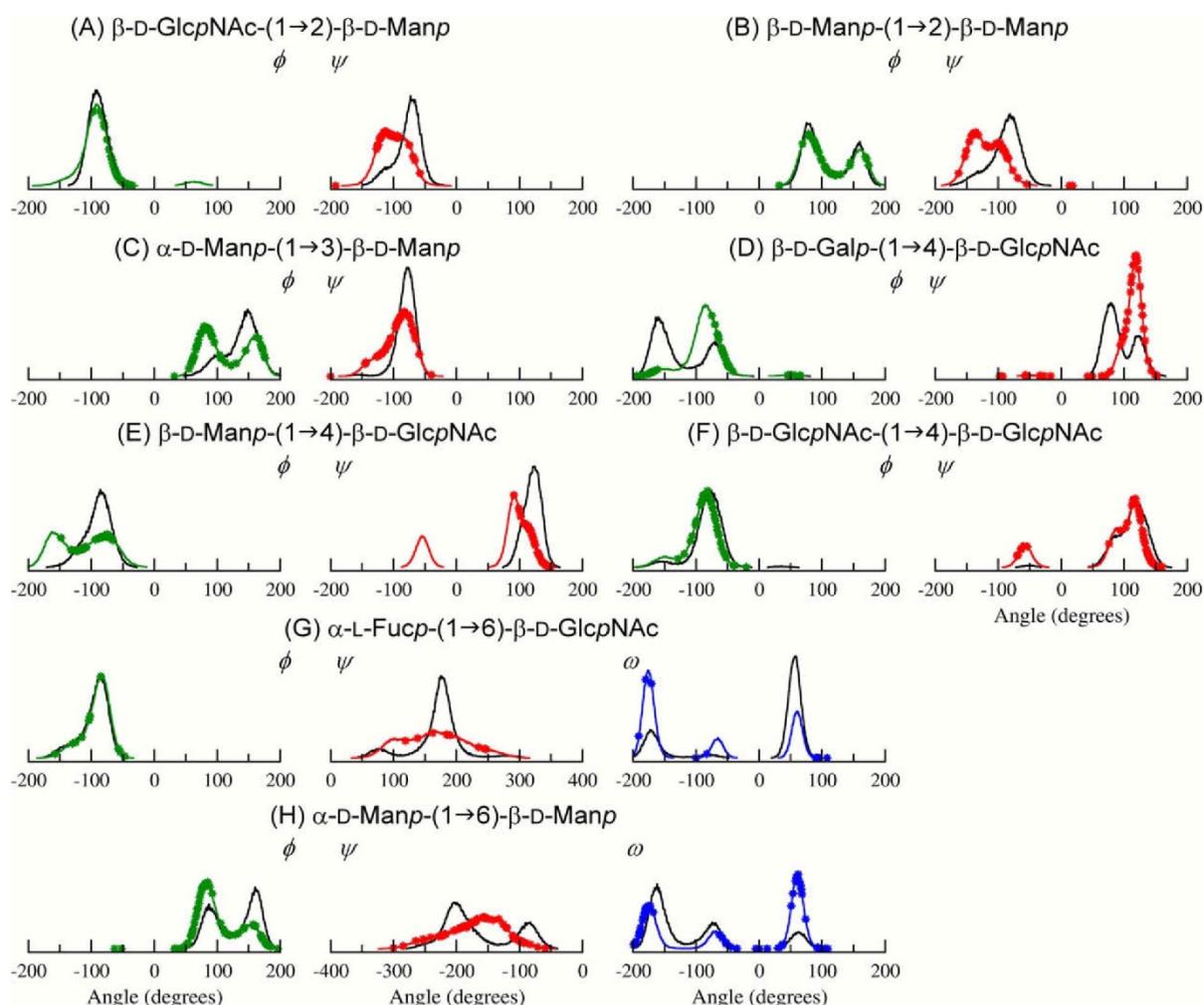
### 3.4. Accuracy in disaccharide use in solution as a conformational reference for building models of glycans in glycoproteins

In spite of the fact that energy contour plots, in the absence of an explicit solvent, are well described in the literature under different force fields and usually having their minima set according to crystallographic data,<sup>35</sup> the direct association of these geometries, to infer solution conformations and ensembles of biological carbohydrates, has some important drawbacks: (1) the high amount of errors in the PDB of carbohydrate containing structures;<sup>12,13</sup> (2) usually, only a few carbohydrate residues are found in glycoproteins crystallographic structures, close to the protein core,<sup>9,10</sup> impairing the acquiring of information about carbohydrates conformations in the extremity of glycans; (3) aqueous solution absent environments tend to reveal a distinct set of conformers when compared to solution data;<sup>60</sup> (4) crystal packing effects promote both glycosidic linkage conformational modifications and hexopyranose ring distortions.<sup>61</sup> In this context, additional strategies capable of supplying carbohydrate and glycoconjugate conformational ensembles of biological significance may strongly contribute to the understanding of the functional roles of such molecules in living organisms.

Employing the GROMACS<sup>26</sup> package as a fast, low cost, simulation package, supporting the achievement of high conformational samplings, we had previously demonstrated that the GROMOS96 43a1 force field,<sup>27</sup> supplemented by Löwdin HF/6-31G<sup>\*\*</sup>-derived atomic charges and PRODRG<sup>23</sup> topologies, for carbohydrate residues, is capable of adequately representing glycoprotein conformational ensemble in aqueous solutions, using the NMR data<sup>17</sup> as a reference ensemble. As this work employed the NMR data from PDB as the starting geometries, its results were compared to those obtained in the current work, where no previous experimental data were used to set the carbohydrate simulations, in order to check the reliability of disaccharide solution ensembles to reproduce the conformational behavior of equivalent units in complex glycans as glycoproteins.

Therefore, the conformational distribution of a set of glycosidic linkages was compared between its isolated disaccharides and also when within protein-bounded glycans, as shown in Figure 3. Besides, each glycosidic linkage conformation, as described in the employed NMR models from PDB, was included in the distribution curves as asterisks, in order to offer an experimental reference of each disaccharide conformational ensemble, in a total of 836 geometries. From these points, at least 90% were populated in each disaccharide, indicating the high reliability in simulating disaccharides to represent biological ensembles of glycoproteins as determined by NMR methods.

From all 18 dihedral angles analyzed, five presented virtually identical behaviors between the disaccharide and the glycan forms of the linkages, as for the  $\phi$  angle from  $\beta$ -D-GlcpNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp, the  $\phi$  angle from  $\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp, the  $\phi$  and  $\psi$  angles from  $\beta$ -D-GlcpNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc, and the  $\phi$  angle



**Figure 3.** Distribution of the  $\phi$  (green),  $\psi$  (red), and  $\omega$  (blue) dihedral angles associated with the studied disaccharides in 0.1  $\mu$ s MD simulation. The behavior of such glycosidic linkages in their protein-bounded glycans is also supplied (black curves). The asterisks (\*) in the isolated disaccharides distribution curves indicate the NMR experimental geometries. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from  $\alpha$ -L-Fucp-(1 $\rightarrow$ 6)- $\beta$ -D-GlcpNAc, which suggests an absence of influence of the entire glycoprotein scaffold on specific disaccharide units, at least on the cases studied here. On the other hand, the six angles presented, as expected, had an increased flexibility in the isolated disaccharides when compared to the same units in the glycoprotein structure, suggesting some degree of conformational susceptibility to the chemical neighborhood, including the  $\psi$  angle from  $\beta$ -D-GlcpNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp, the  $\psi$  angle from  $\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp, the  $\psi$  angle from  $\alpha$ -D-Manp-(1 $\rightarrow$ 3)- $\beta$ -D-Manp, the  $\phi$  and  $\psi$  angles from  $\beta$ -D-Manp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc, and the  $\psi$  angle from  $\alpha$ -L-Fucp-(1 $\rightarrow$ 6)- $\beta$ -D-GlcpNAc. Curiously, both angles from  $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc presented an inverse behavior, with the isolated disaccharides presenting a lower flexibility than any of the glycans, suggesting that the glycoprotein scaffold may be capable of shifting the conformer populations equilibrium for some glycosidic linkages. Nevertheless, the simulation of a glycoprotein, with its  $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc, in the main conformation obtained from the isolated disaccharides, does offer the same conformational pattern as described by the NMR.<sup>17</sup> This indicates the plasticity of these linkages and the capability of MD simulations in explicit solvent, to adequately represent the expected conformational transitions of a given disaccharide unit in two different environments. Furthermore, in most of the cases where the glycosidic linkage geometries were not identical be-

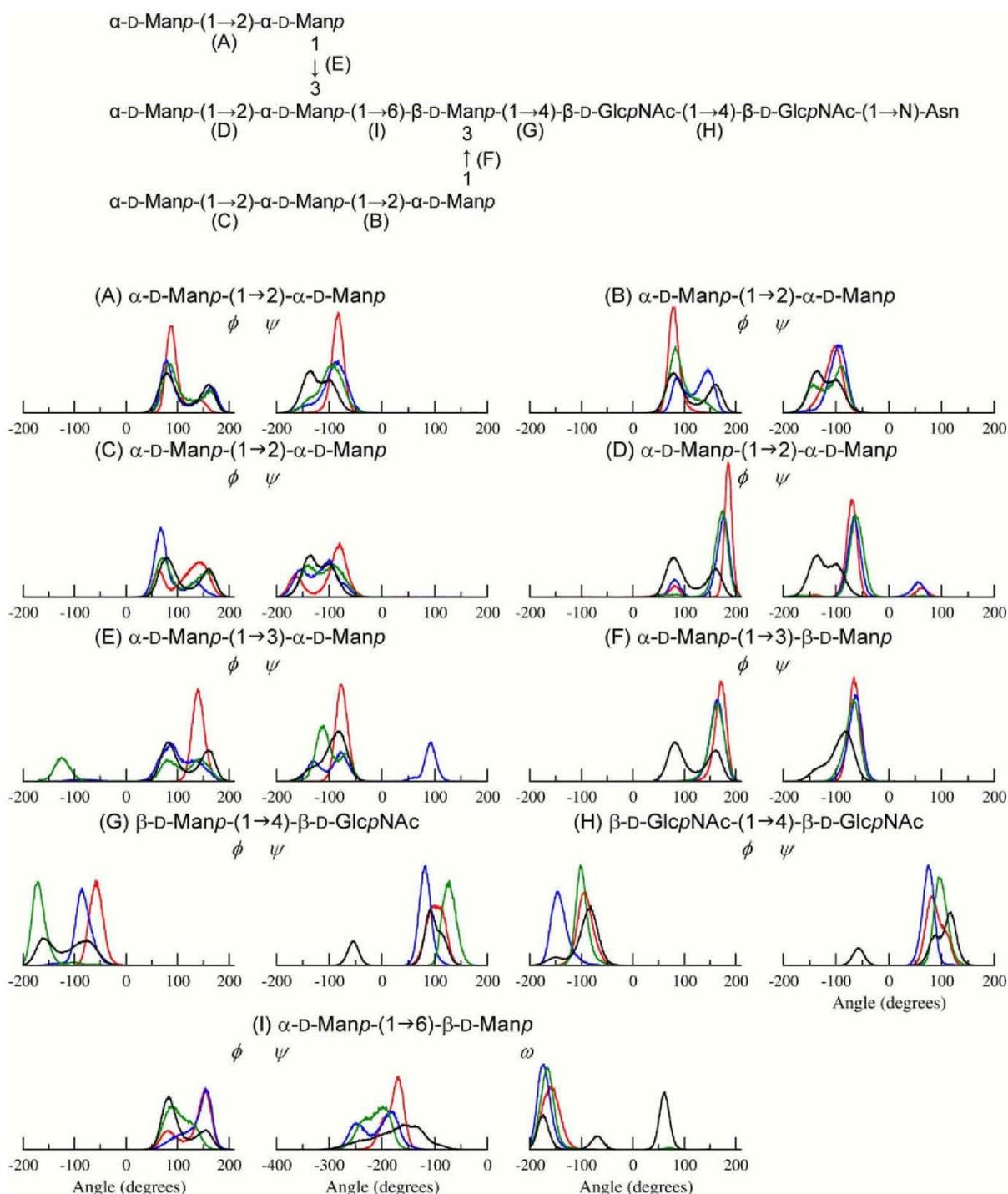
tween the isolated disaccharide and the entire glycoproteins, there were no formations of new conformational peaks, but a modification in the relative abundance of the same conformational state. The major exception was the  $\psi$  angle from  $\alpha$ -D-Manp-(1 $\rightarrow$ 6)- $\beta$ -D-Manp, in which the main conformational state in the isolated disaccharide resembled an average conformation between the two conformational states observed in the same disaccharide when inserted in a glycoprotein molecular context.

In order to illustrate the potential of using disaccharide solution conformations as the starting geometries for building complete glycoproteins, we employed the main conformational states described in Figure 3 and Table 1, to build models for the complete glycosylated COX-1 and COX-2 enzymes. In fact, there are about 30 crystallographic structures deposited in PDB for the two proteins, most of it glycosylated (Supplementary data). Unfortunately, the majority of such structures present a small number of carbohydrate residues, that is, half of all crystal structures has up to two residues in each glycosylation position. Also, the structures presenting higher glycosylation patterns present only three Man residues, and for a single position, mostly Asn144 (PDB IDs 1IGX, 1IGZ, 1FE2, 1DIY, 1Q4G, and 2AYL), while 5–10 Man are expected to occur in all positions.<sup>39,40</sup> Most importantly, several errors have been observed in these structures (Supplementary data), as evidenced by pdb-care<sup>41</sup> and based on the expected glycan structure assem-

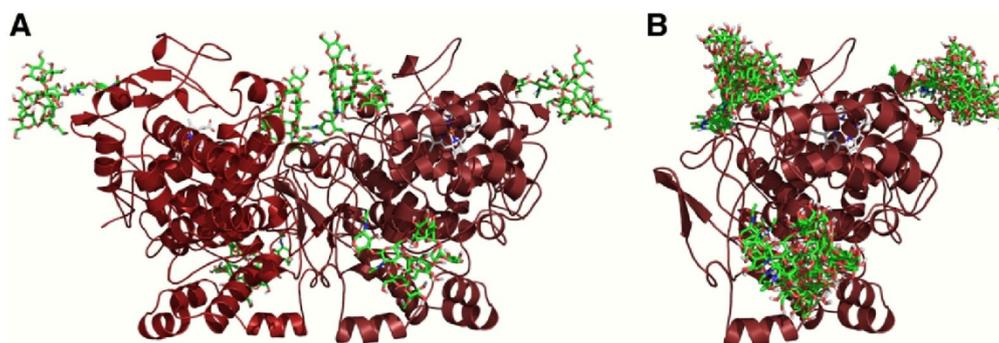
bled in the endoplasmic reticulum.<sup>2,62</sup> In this context, and considering the great therapeutic impact of cyclooxygenase inhibitors, a putative model was built for the glycan moiety of such enzymes (Fig. 4, see Section 2 for details), in agreement with most of the available experimental data on this system.<sup>39,40</sup> Such a glycan was employed to glycosylate both COX-1 and COX-2 in positions Asn68, Asn144, and Asn410.<sup>63</sup> The obtained glycoproteins were further submitted to MD simulations in order to refine the models

and allow for carbohydrate accommodation in the presence of solvent and protein scaffold. A dimeric glycosylated COX-2 is presented in Figure 5, as well as a representation of the glycan ensemble through the simulations.

Each dihedral from glycosidic linkages was analyzed and compared to the conformational pattern observed for the proper disaccharide units (Fig. 4 for COX-2 and in Supplementary data for COX-1). Accordingly, as a general feature and previously pointed in the



**Figure 4.** Distribution of the  $\phi$ ,  $\psi$ , and  $\omega$  dihedral angles, as obtained from MD simulation, from COX-2 glycans at Asn68 (red), Asn144 (blue), and Asn410 (green), together with the conformational patterns associated with the respective disaccharides, isolated in water (black). The A–I letters indicate the relation of a distribution plot to its location on the respective disaccharide unit from the complete glycan. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Figure 5.** Glycosylated structure from murine COX-2 built based on the main conformational states as presented in Figure 3 and Table 1. (A) glycosylated COX-2 dimer and (B) superimposition of snapshots obtained from 50 ns MD of the monomeric fully glycosylated COX-2.

text, most of the linkages do not present new conformer populations, but distinct preferences over the conformations observed in solution for disaccharide units. Nevertheless, in spite of the fact that a given glycosidic linkage also has the same starting conformation, the influence of the distinct positions in the protein sequence (and of the antennae themselves) over the glycans may still be observed, originating distinct relative abundances of conformations, as noted in  $\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp (Fig. 4A–D). This points to the adequacy of the employed protocol in obtaining reliable models for glycoproteins conformational behavior in aqueous solutions.

#### 4. Conclusions

The acquirement of reliable models for the 3D structure of carbohydrates in complex biological systems, such as glycoproteins, represents a challenge for both experimental and theoretical methods. It is a complex picture defined by several possible monomeric units in diverse conformational states, linked in different ways, presenting a variable degree of branching, attached to distinct biological macromolecules and requiring the representation of both spatial and temporal properties in order to be adequately described in solution.

While MD simulations are progressively considered as a potential strategy to obtain a dynamic description of molecules conformational behavior in environments mimicking biological solutions and is successfully applied to the study of proteins, membranes, nucleic acids, and simple carbohydrates, its use to represent glycoproteins and other biological glycoconjugates is much less frequent, mainly due to the lack of parameterization of carbohydrates in its major biological forms. In this context, our group has been dedicated to the conformational representation of glycans of biological significance, in its different levels of complexity, using generally accessible and fast molecular simulation packages.<sup>17,24,25,28–30</sup>

In the current work, we demonstrated that relaxed contour plots obtained from disaccharides employing GROMOS96 43a1 force field<sup>27</sup> added by Löwdin HF/6-31G<sup>\*\*</sup> atomic charges present similar results to those obtained from MM3 force field. What is most important to mention is that the minimum-energy conformations obtained from such plots represent an adequate starting point for further refinement of disaccharides conformational behavior through MD simulations. This combination of vacuum and explicit solution simulation was capable of offering a reliable platform to describe the conformational ensemble of glycosidic linkages, solely based on the carbohydrate sequence and in the absence of the previous experimental information, reproducing more than 90% of the NMR data of complete glycoproteins. In this context, such an approach may constitute a potential strategy to build

accurate models for glycoproteins, as exemplified for COX-1 and COX-2 enzymes, capable of contributing to the understanding of the role of biological glycoconjugates at the atomic level.

#### Acknowledgments

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (#472174/2007-0), MCT, and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), MEC, Brasília, DF, Brazil.

#### Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.carres.2009.12.018.

#### References

- Varki, A. *Glycobiology* **1993**, *3*, 97–130.
- Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
- Imberty, A.; Gerber, S.; Tran, V.; Pérez, S. *Glycoconjugate J.* **1990**, *7*, 27–54.
- Brant, D. A. *Pure Appl. Chem.* **1997**, *69*, 1885–1892.
- Kuttel, M. M.; Naidoo, K. J. *J. Phys. Chem. B* **2005**, *109*, 7468–7474.
- Xia, J.; Daly, R. P.; Chuang, F.-C.; Parker, L.; Jensen, J. H.; Margulis, C. J. *J. Chem. Theory Comput.* **2007**, *3*, 1620–1628.
- Xia, J.; Daly, R. P.; Chuang, F.-C.; Parker, L.; Jensen, J. H.; Margulis, C. J. *J. Chem. Theory Comput.* **2007**, *3*, 1629–1643.
- Xu, Y.; Colletier, J. Ph.; Jiang, H.; Silman, I.; Sussman, J. L.; Weik, M. *Protein Sci.* **2008**, *17*, 601–605.
- Woods, R. J. *Glycoconjugate J.* **1998**, *15*, 209–216.
- Naidoo, K. J.; Denysyk, D.; Brady, J. W. *Protein Eng.* **1997**, *10*, 1249–1261.
- Sureshan, K. M.; Miyasou, T.; Watanabe, Y. *Carbohydr. Res.* **2004**, *339*, 1551–1555.
- Crispin, M.; Stuart, D. I.; Jones, E. Y. *Nat. Struct. Mol. Biol.* **2007**, *14*, 354.
- Berman, H. M.; Henric, K.; Nakamura, H.; Markley, J. *Nat. Struct. Mol. Biol.* **2007**, *14*, 354–355.
- Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664–6676.
- Wormald, M.; Petrescu, A.-J.; Pao, Y.-L.; Glythero, A.; Elliot, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371–387.
- Pérez, S.; Mulloy, B. *Curr. Opin. Struct. Biol.* **2005**, *15*, 517–524.
- Pol-Fachin, L.; Fernandes, C. L.; Verli, H. *Carbohydr. Res.* **2009**, *344*, 491–500.
- Fletcher, C. M.; Harrison, R. A.; Lachmann, P. J.; Neuhaus, D. *Structure* **1994**, *2*, 185–199.
- Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R. N.; Smolyar, A.; Reinherz, E. L.; Wagner, G. *Science* **1995**, *269*, 1273–1278.
- Erbel, P. J. A.; Karimi-Nejad, Y.; van Kuik, J. A.; Boelens, R.; Kamerling, J. P.; Vliegthart, J. F. G. *Biochemistry* **2000**, *39*, 6012–6021.
- I.U.P.A.C.-I.U.B. Commission on Biochemical Nomenclature *Pure Appl. Chem.* **1996**, *68*, 1919–2008.
- Schaftenaar, G.; Noordik, J. H. *J. Comput. Aided Mol. Des.* **2000**, *14*, 123–134.
- Schuettelkopf, A. W.; van Aalten, D. M. F. *Acta Crystallogr., Sect. D* **2004**, *60*, 1355–1363.
- Verli, H.; Guimarães, J. A. *Carbohydr. Res.* **2004**, *339*, 281–290.
- Becker, C. F.; Guimarães, J. A.; Verli, H. *Carbohydr. Res.* **2005**, *340*, 1499–1507.
- van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
- van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Huenenberger, P. H.; Krueger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G.; Simulation, Biomolecular. *The*

- GROMOS96 *Manual and User Guide*; Vdf Hochschulverlag: AG Zurich, Switzerland, 1996.
28. Verli, H.; Guimarães, J. A. J. *Mol. Graph. Model.* **2005**, *24*, 203–212.
  29. Pol-Fachin, L.; Verli, H. *Carbohydr. Res.* **2008**, *343*, 1435–1445.
  30. Castro, M. O.; Pomin, V. H.; Santos, L. L.; Vilela-Silva, A.-C. E. S.; Hirohashi, N.; Pol-Fachin, L.; Verli, H.; Mourão, P. A. S. *J. Biol. Chem.* **2009**, *284*, 18790–18800.
  31. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
  32. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
  33. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
  34. Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
  35. Lütteke, T.; Frank, M.; von der Lieth, C.-W. *Carbohydr. Res.* **2004**, *339*, 1015–1020.
  36. Lütteke, T.; Frank, M.; von der Lieth, C.-W. *Nucleic Acids Res.* **2005**, *33*, D242–D246.
  37. Lütteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C.-W. *Glycobiology* **2006**, *16*, 71R–81R.
  38. Mutsaers, J. H. G. M.; van Halbeek, H.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.* **1985**, *147*, 569–574.
  39. Nemeth, J. F.; Hochesang, G. P.; Marnett, L. J.; Caprioli, R. M. *Biochemistry* **2001**, *40*, 3109–3116.
  40. Lütteke, T.; von der Lieth, C.-W. B. M. C. *Bioinformatics* **2004**, *5*, 69.
  41. Furse, K. E.; Pratt, D. A.; Porter, N. A.; Lybrand, T. P. *Biochemistry* **2006**, *45*, 3189–3205.
  42. Hashimoto, K.; Goto, S.; Kawano, S.; Aoki-Kinoshita, K. F.; Ueda, N.; Hamajima, M.; Kawasaki, T.; Kanehisa, M. *Glycobiology* **2006**, *16*, 63R–70R.
  43. Allinger, N. L.; Yuh, Y. H.; Li, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8567.
  44. Allinger, N. L.; Rahman, M.; Lii, J.-H. *J. Am. Chem. Soc.* **1990**, *112*, 8293–8307.
  45. Pérez, S.; Imbert, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J.-F.; van Eyck, B. P.; Johnson, G.; French, A. D.; Kouwijzer, M. L. C. E.; Grootenuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. *Carbohydr. Res.* **1998**, *314*, 141–155.
  46. Hemmingsen, L.; Madsen, D. E.; Esbensen, A. L.; Olsen, L.; Engelsen, S. B. *Carbohydr. Res.* **2004**, *339*, 937–948.
  47. Mackerell, A. D. *J. Comput. Chem.* **2004**, *25*, 1584–1604.
  48. Mackerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
  49. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
  50. Röss, D. A.; Scott, W. E. *J. Chem. Soc., B* **1971**, 469–479.
  51. Sathyanarayana, B. K.; Rao, V. S. R. *Biopolymers* **1972**, *11*, 1379–1394.
  52. Bluhm, T. L.; Deslandes, Y.; Marchessault, R. H.; Perez, S.; Rinaudo, M. *Carbohydr. Res.* **1982**, *100*, 117–130.
  53. Sathyanarayana, B. K.; Stevens, E. S. *J. Biomol. Struct. Dyn.* **1983**, *1*, 947–959.
  54. Lipkind, G. M.; Verovskiy, V. E.; Kochetkov, N. K. *Carbohydr. Res.* **1984**, *133*, 1–13.
  55. Lipkind, G. M.; Shashkov, A. S.; Kochetkov, N. K. *Carbohydr. Res.* **1985**, *141*, 191–197.
  56. Virudachalam, R.; Rao, V. S. R. *Biopolymers* **1979**, *18*, 571–589.
  57. Shefter, E.; Trueblood, K. N. *Acta Crystallogr.* **1965**, *18*, 1067–1077.
  58. Pope, L. *J. Am. Chem. Soc.* **1993**, *115*, 8421–8426.
  59. Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6676–6683.
  60. Yoneda, J. D.; Albuquerque, M. G.; Leal, K. Z.; Seidl, P. R.; Wheeler, R. A.; Boesch, S. E.; de Alencastro, R. B.; de Souza, M. C. B. V.; Ferreira, V. F. *Theochem* **2006**, *778*, 97–103.
  61. Lütteke, T. *Acta Crystallogr., Sect. D* **2009**, *65*, 156–168.
  62. Rollins, T. E.; Smith, W. L. *J. Biol. Chem.* **1980**, *255*, 4872–4875.
  63. Otto, J. C.; DeWitt, D. L.; Smith, W. L. *J. Biol. Chem.* **1993**, *268*, 18234–18242.

## Supplementary Data

### GROMOS96 43a1 performance on the prediction of oligosaccharides conformational ensemble within glycoproteins

Fernandes, C.L.<sup>a</sup>, Sachett, L.G.<sup>a</sup>, Pol-Fachin, L.<sup>a</sup>, and Verli, H.<sup>a,b,\*</sup>

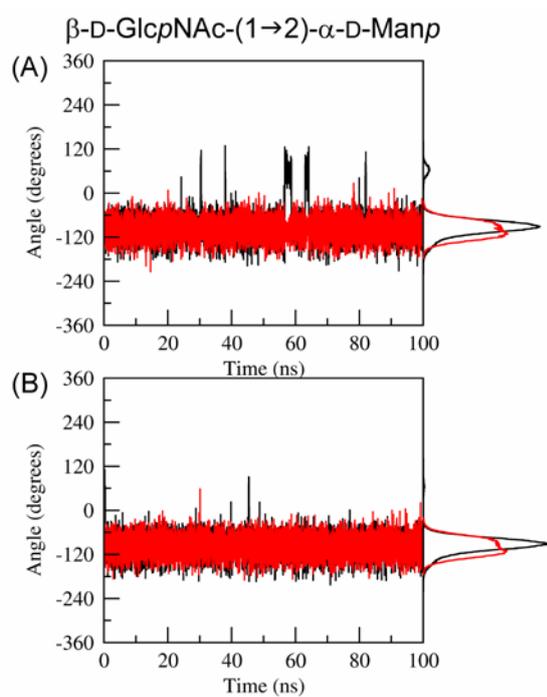
<sup>a</sup>*Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil*

<sup>b</sup>*Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Av Ipiranga 2752, Porto Alegre 90610-000, RS, Brazil*

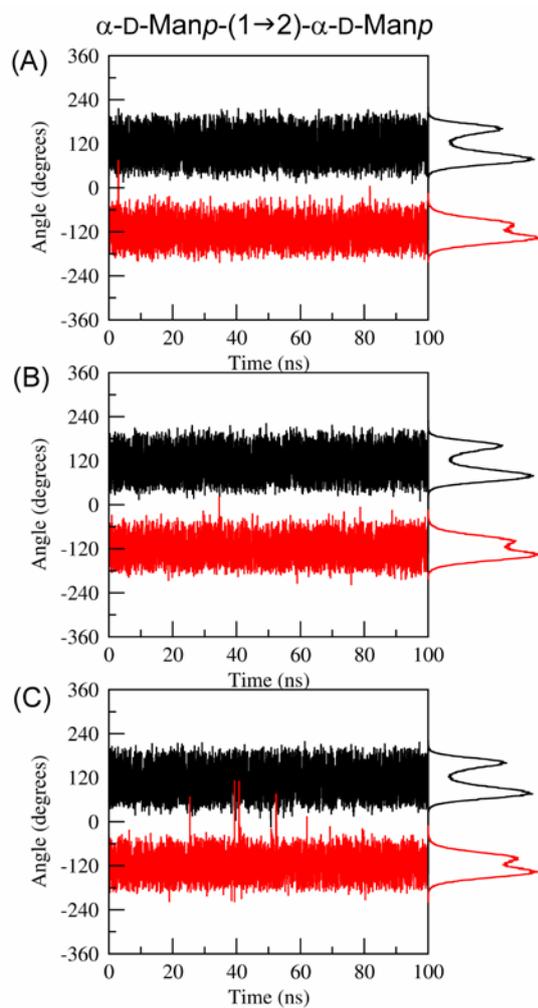
\*Corresponding author. Tel.: +55-51-3308-7770; fax: +55-51-3308-7309; e-mail address: hverli@cbiot.ufrgs.br.

**Table S1:** Minimum energy geometries for dihedral angles  $\phi$ ,  $\psi$  and  $\omega$ , as obtained from relaxed contour plots for a series of disaccharides, refined through 100 ns MD in explicit solvent. Each disaccharide in each pair of dihedrals was submitted to a separate simulation, in a total of 21 simulations (2,100 ns).

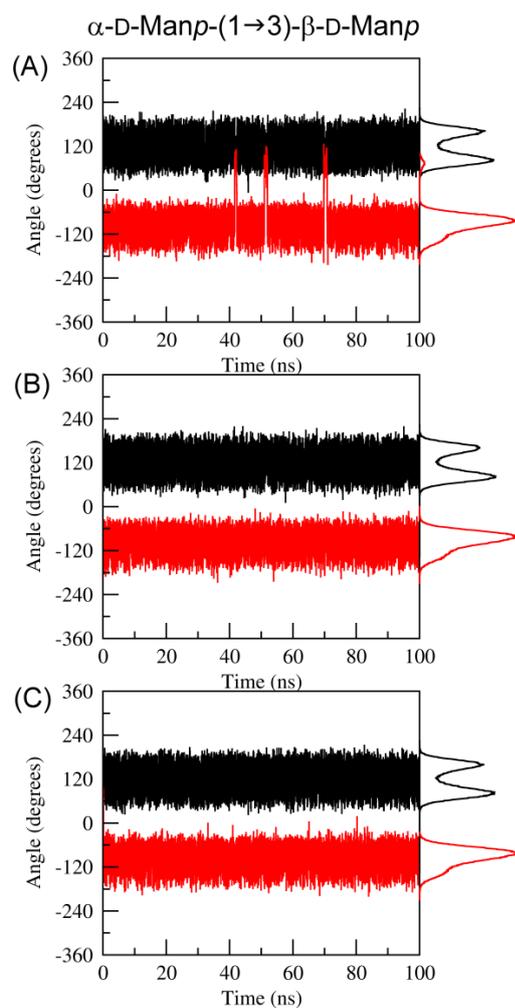
Disaccharides	Simulation	$\phi$	$\psi$	$\omega$
$\beta$ -D-GlcpNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp	A	-60	-150	-
	B	60	90	-
$\alpha$ -D-Manp-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp	A	-180	-90	-
	B	60	-150	-
	C	90	-60	-
$\alpha$ -D-Manp-(1 $\rightarrow$ 3)- $\beta$ -D-Manp	A	-180	-90	-
	B	60	-120	-
	C	120	90	-
$\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc	A	-120	90	-
	B	60	120	-
$\beta$ -D-Manp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc	A	-150	90	-
	B	-60	-60	-
$\beta$ -D-GlcpNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc	A	-150	90	-
	B	-90	-60	-
	C	-60	150	-
	D	60	150	-
$\alpha$ -L-Fucp-(1 $\rightarrow$ 6)- $\beta$ -D-GlcpNAc	A	-150	-150	-60
	B	-90	30	-180
	C	-90	150	60
$\alpha$ -D-Manp-(1 $\rightarrow$ 6)- $\beta$ -D-Manp	A	60	150	60
	B	90	-120	-180



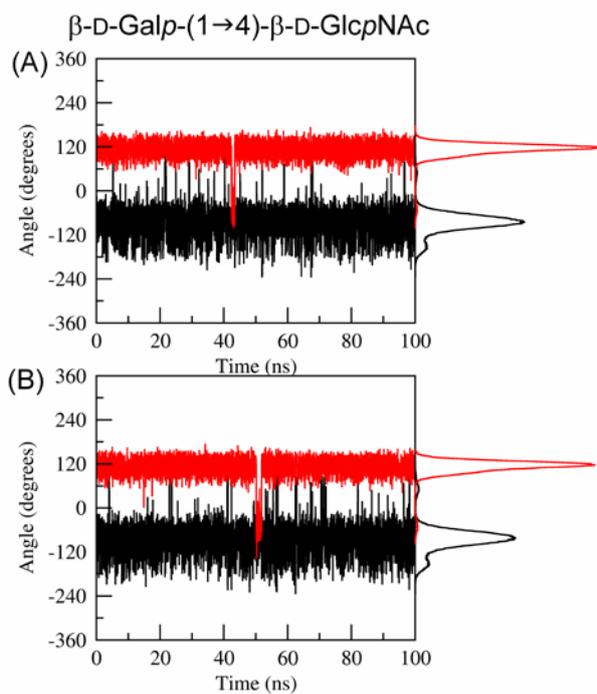
**Figure S1:** Fluctuation and distribution of  $\phi$  (black) and  $\psi$  (red) dihedral angles from disaccharide  $\beta$ -D-GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -D-Manp in the MD simulations. Letters (A) and (B) correspond to initial conformations as shown in Table S1.



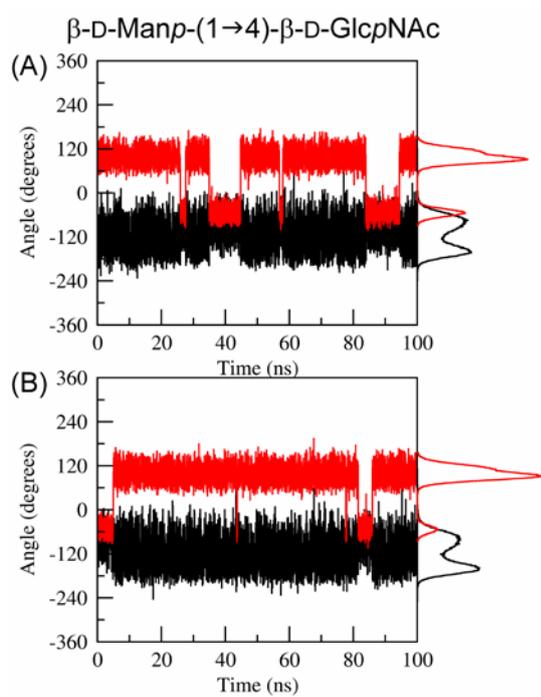
**Figure S2:** Fluctuation and distribution of  $\phi$  (black) and  $\psi$  (red) dihedral angles from in the disaccharide  $\alpha$ -D-Man $\rho$ -(1 $\rightarrow$ 2)- $\alpha$ -D-Man $\rho$  in MD simulations. Letters (A), (B) and (C) correspond to initial conformations as shown in Table S1.



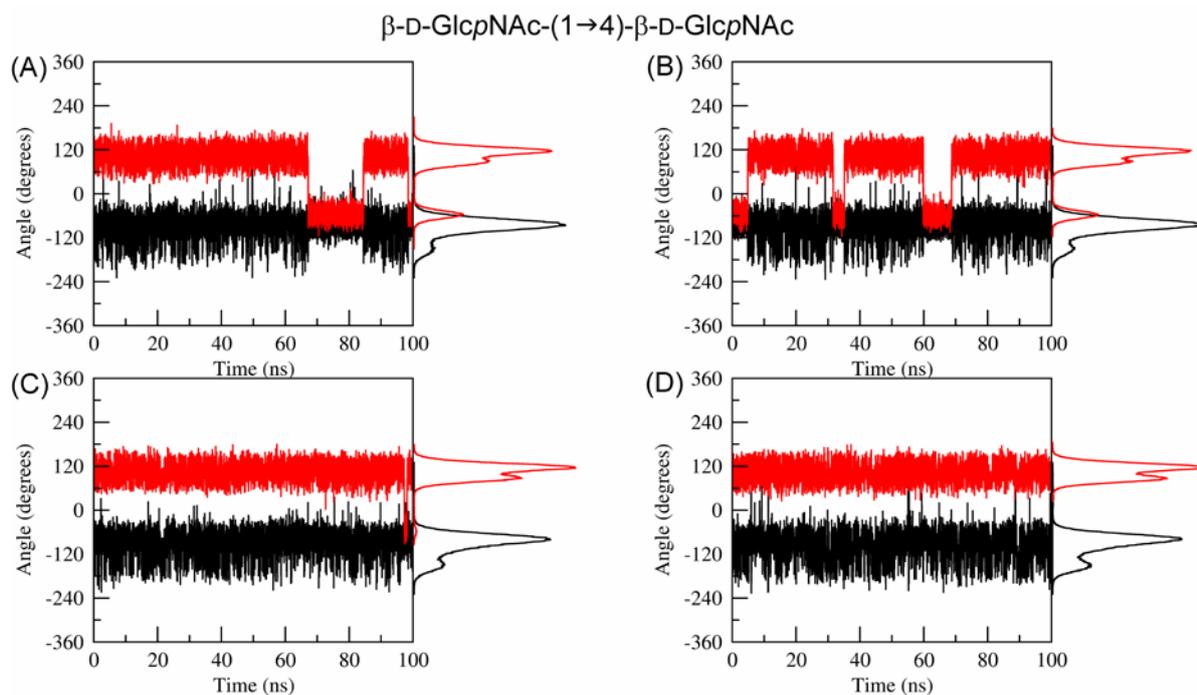
**Figure S3:** Fluctuation and distribution of  $\phi$  (black) and  $\psi$  (red) dihedral angles from in the disaccharide  $\alpha$ -D-Manp-(1 $\rightarrow$ 3)- $\alpha$ -D-Manp in MD simulations. Letters (A), (B) and (C) correspond to initial conformations as shown in Table S1.



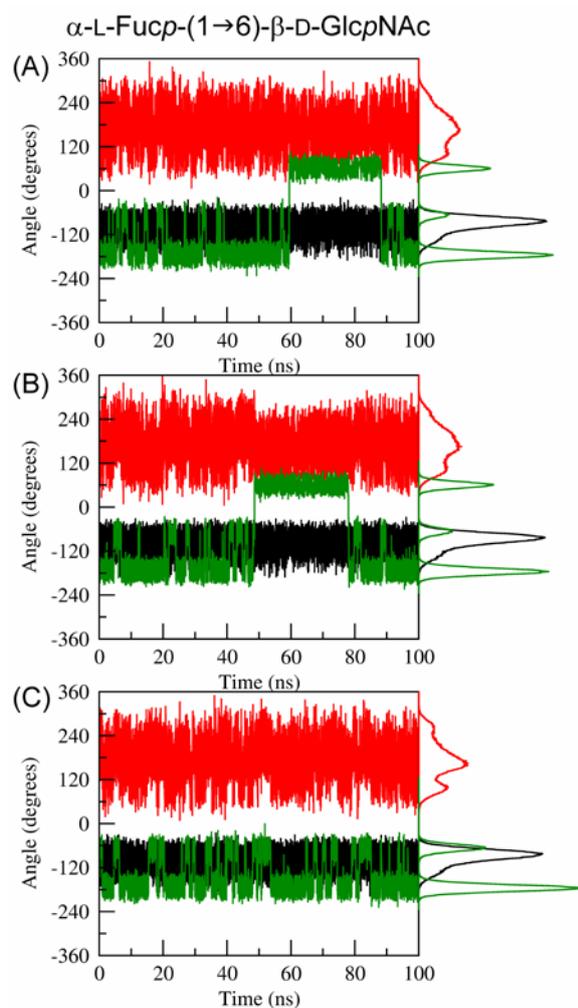
**Figure S4:** Fluctuation and distribution of  $\phi$  (black) and  $\psi$  (red) dihedral angles from in the disaccharide  $\beta$ -D-Galp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc in MD simulations. Letters (A) and (B) correspond to initial conformations as shown in Table S1.



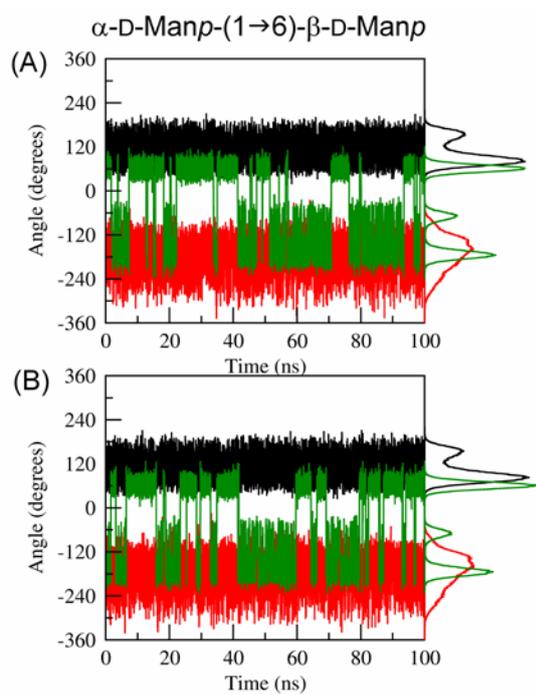
**Figure S5:** Fluctuation and distribution of  $\phi$  (black) and  $\psi$  (red) dihedral angles from in the disaccharide  $\beta$ -D-Manp-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc in MD simulations. Letters (A) and (B) correspond to initial conformations as shown in Table S1.



**Figure S6:** Fluctuation and distribution of  $\phi$  (black) and  $\psi$  (red) dihedral angles from in the disaccharide  $\beta$ -D-GlcpNAc-(1 $\rightarrow$ 4)- $\beta$ -D-GlcpNAc in MD simulations. Letters (A), (B), (C) and (D) correspond to initial conformations as shown in Table S1.



**Figure S7:** Fluctuation and distribution of  $\phi$  (black),  $\psi$  (red) and  $\omega$  (green) dihedral angles from in the disaccharide  $\alpha$ -L-Fucp-(1 $\rightarrow$ 6)- $\beta$ -D-GlcpNAc in MD simulations. Letters (A), (B) and (C) correspond to initial conformations as shown in Table S1.



**Figure S8:** Fluctuation and distribution of  $\phi$  (black),  $\psi$  (red) and  $\omega$  (green) dihedral angles from in the disaccharide  $\alpha$ -D-Manp-(1 $\rightarrow$ 6)- $\beta$ -D-Manp in MD simulations. Letters (A) and (B) correspond to initial conformations as shown in Table S1

**Table S2.** Glycosylation structures obtained from crystallographic structures available for COX-1. Pdb-care errors found in carbohydrates are shown as red.

PDB	Glycosylation
1PRH 3.5 Å	No glycosylation structure
1EQG 2.6 Å	No glycosylation structure
1EQH 2.7 Å	No glycosylation structure
1HT5 2.7 Å	No glycosylation structure
1HT8 2.7 Å	No glycosylation structure
1PGG 4.5 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→N)-Asn410
1PGF 4.5 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→N)-Asn410
1PGE 3.5 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→N)-Asn410
1PTH 3.4 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→N)-Asn410
2OYE 2.8 Å	$\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-Manp-(1→3)- $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410
2OYU 2.7 Å	$\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410
1IGX 3.1 Å	$\alpha$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\alpha$ -D-Manp-(1→3)- $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410
1IGZ 2.9 Å	$\alpha$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\alpha$ -D-Manp-(1→3)- $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410
1FE2 3.0 Å	$\alpha$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-Manp-(1→3)- $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410
1DIY 3.0 Å	$\alpha$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\alpha$ -D-Manp-(1→3)- $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410
1CQE 3.1 Å	$\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410
1Q4G 2.0 Å	$\alpha$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\alpha$ -D-Manp-(1→6)- $\alpha$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\alpha$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410 $\alpha$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn68 $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\alpha$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn144 $\beta$ -D-Manp-(1→6)- $\beta$ -D-Manp-(1→4)- $\beta$ -D-GlcpNAc-(1→4)- $\beta$ -D-GlcpNAc-(1→N)-Asn410

2AYL 2.0 Å	$\alpha$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-Manp-(1→6)-β-D-Manp-(1→6)-β-D-Manp-(1→4)- $\alpha$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-Manp-(1→6)-β-D-Manp-(1→4)-β-D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn410
1U67 3.1 Å	$\alpha$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-Manp-(1→6)-β-D-Manp-(1→6)-β-D-Manp-(1→4)- $\alpha$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-Manp-(1→6)-β-D-Manp-(1→4)-β-D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn410
1EBV 3.2 Å	$\alpha$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-Manp-(1→6)-β-D-Manp-(1→4)-β-D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn410
	$\beta$ -D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410

**Table S3.** Glycosylation structures obtained from crystallographic structures available for COX-2.

PDB	Glycosylation
1PXX 2.9 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410
1CVU 2.4 Å	$\beta$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn68
	$\alpha$ -D-Manp-(1→6)- $\alpha$ -D-Manp-(1→4)-β-D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410
1DDX 3.0 Å	$\beta$ -D-GlcpNAc-(1→4)-β-D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410
1CX2 3.0 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410
3PGH 2.5 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410
4COX 2.9 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410
5COX 3.0 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410
6COX 2.8 Å	$\beta$ -D-GlcpNAc-(1→N)-Asn68
	$\beta$ -D-GlcpNAc-(1→N)-Asn144
	$\beta$ -D-GlcpNAc-(1→N)-Asn410



#### 4.4 Parte III

### Dynamics of Different Arachidonic Acid Orientations bound to Prostaglandin Endoperoxide Synthases

Liana Guimarães Sachett, Hugo Verli

*Eur. J. Med Chem.* em preparação.

As enzimas PGHS catalisam o primeiro passo na biossíntese de prostaglandinas e tromboxanos ao converterem o seu substrato AA a  $PGH_2$ , em uma reação que ocorre em duas etapas distintas. Para esta reação é necessário que o AA esteja em uma conformação catalítica, onde o radical tirosil em Tyr385 seja capaz de abstrair o átomo de hidrogênio do carbono C13 deste substrato. Além desta orientação catalítica, existem estruturas cristalográficas em que este substrato está orientado de maneira alternativa, impossibilitando a reação catalítica. Esta orientação vem sendo discutida de maneira não conclusiva e sua caracterização estrutural em solução é ausente. A orientação alternativa para o AA é muito semelhante à orientação de alguns AINES, como o diclofenaco, por exemplo. Ainda, outro inibidor da enzima, o composto SC-558, foi recentemente caracterizado por estudos conformacionais, onde foi demonstrado que este inibidor possui duas orientações possíveis no sítio ativo desta enzima. Desta maneira, neste trabalho é apresentado um estudo do comportamento dinâmico das diferentes orientações do substrato AA complexadas a PGHSs. Os resultados obtidos indicam que a orientação alternativa não é estável, possivelmente sendo uma orientação minoritária em solução.

# Dynamics of Different Arachidonic Acid Orientations bound to Prostaglandin Endoperoxide Synthases

Liana Guimarães Sachett<sup>a</sup> and Hugo Verli<sup>b,\*</sup>

<sup>a</sup>*Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil.*

<sup>b</sup>*Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Av. Ipiranga 2752, Porto Alegre 90610-000, RS, Brazil.*

## ABSTRACT

Prostaglandin endoperoxide synthases (PGHSs) catalyze the conversion of arachidonic acid (AA) into prostaglandin endoperoxide H<sub>2</sub>. This reaction requires a specific orientation of AA within the active site, but an alternative crystallographic binding orientation for AA also exists. The origin of this alternative complex, and its potential relevance, have been neglected so far. Therefore, we have characterized the dynamic behaviors of both catalytic and alternative orientations of AA, bound to PGHS-1 and -2, using molecular dynamics simulations. Our results indicate that AA in the alternative orientation seems to be less stable, moving towards the active site exit. This potentially minor orientation of AA can be related to crystallographic complexes of anti-inflammatory agents (e.g. diclofenac), providing insights for designing new PGHSs inhibitors.

## *Keywords:*

Arachidonic acid, prostaglandin endoperoxide synthase, cyclooxygenase, molecular dynamics

## *Abbreviations:*

PGHS: prostaglandin endoperoxide synthase; AA: arachidonic acid; NSAID: non-steroidal anti-inflammatory drugs; NO-NSAID: nitric-oxide-releasing anti-inflammatory drugs; COX: cyclooxygenase; POX: peroxidase; MD: molecular dynamics; EDS: electron density server

## 1. Introduction

Prostaglandin endoperoxide synthases (PGHSs), also known as cyclooxygenases, are homodimeric integral membrane enzymes, that are located in the endoplasmic reticulum lumen and nuclear envelope [1], where they catalyze the committed step in prostanoid biosynthesis, converting arachidonic acid (AA) and O<sub>2</sub> to prostaglandin endoperoxide H<sub>2</sub> (PGH<sub>2</sub>) [2]. The reaction occurs in two steps, processed in two distinctive active sites. There are two well characterized isoforms of the enzyme. The first, PGHS-1, is constitutively expressed in most tissue types and performs a “housekeeping” function, generating low levels of prostanoid biosynthesis for maintaining physiological homeostasis [3]. The second isoform, PGHS-2, is known as the inducible form, since it is present in mammalian tissues at very low levels. PGHS-2 expression increase is triggered by specific stimuli, such as cytokines, growth factors and inflammation [4].

The PGHSs structure consists of three distinct domains: an N-terminal epidermal growth factor domain, followed by a membrane binding domain (MBD), and a C-terminal catalytic domain containing the cyclooxygenase (COX) and the peroxidase (POX) active sites [5]. Both isoforms of PGHS are N-glycosylated in the residues Asn68, Asn144 and Asn410 [6-8]. The glycosyl moieties in such structures are of a high-mannose type and are essential for the catalytic activity of PGHSs, probably by promoting proper protein folding [9].

The COX reaction consists of oxygenating AA, to produce prostaglandin G<sub>2</sub> (PGG<sub>2</sub>). The mechanism for this reaction requires the enzyme to be activated at Tyr385, producing a tyrosyl radical which abstracts the hydrogen from the carbon 13 of AA. This generates an activated arachidonyl radical that undergoes cyclization and an oxygenation reaction, which renders it PGG<sub>2</sub>. This intermediate moves to the POX active site to be reduced to PGH<sub>2</sub>. A catalytic orientation of AA in the COX active site is required for the reaction to occur. This is maintained by a salt bridge between AA carboxyl and Arg120 guanidine groups. The Ser530 residue is also important as it coordinates the geometry of attack in the reaction [10,11]. In contradiction to this observation, an alternative orientation of the substrate is found in crystallographic data (PDB IDs: 1CVU and 3HS5). This alternative orientation appears not to be viable for catalysis, as the carboxylic end of AA is bound to the side chain of Tyr385 and Ser530 [12,13], impairing its reaction with the carbon 13. Nevertheless, the biological relevance of the observed orientation has not been discussed so far and conformational studies are only available for the catalytic orientation [12-14].

The crystallographic data concerning PGHSs inhibitors demonstrates that some of these compounds are oriented to interact with the enzyme at Tyr385 and Ser530. This is the case for diclofenac (PDB ID: 1PXX) [15] and SC-558 (PDB ID: 6COX) [16]. Thus, understanding the interaction dynamics of the alternative orientation of AA within the COX active site of the enzyme may help in the interpretation of this binding region and, so forth, contribute to the design of new PGHSs inhibitors. Selective PGHS-2 inhibitors, called coxibs, were developed for treating pain and inflammation, in an attempt to reduce the gastrointestinal side effects caused by non-selective non-steroidal anti-inflammatory drugs (NSAIDs) [17-19]. Although these compounds are considered an important strategy in treating and preventing some types of cancer [20-22], some coxibs were linked to increased risk of blood clotting, heart attack and stroke [23-25]. In order to prevent these cardiovascular side effects, new NSAIDs are being studied, specifically, nitric-oxide-releasing NSAIDs (NO-NSAIDs), an emerging class of PGHSs inhibitors, capable of improving cardiovascular safety as well as reducing gastrointestinal toxicity [26,27]. PGHSs inhibition is still focus of attention in treatment of pain and inflammation, and understanding the different modes of interaction that could occur within the active site of this enzyme, is essential for the rational drug-design of new compounds. In this context, the present study characterizes, for the first time, the conformational dynamics of the alternative orientation of AA, in solution, in both PGHSs, in their fully glycosylated states.

## 2. Results and Discussion

### a. *Simulation systems*

The presence of an alternative orientation for the PGHSs substrate AA has been interpreted as a non-conclusive mode, and the study of such orientation has not yet been characterized in solution. In order to conformationally characterize, in solution, the diverging orientations of AA, within the catalytic site of PGHSs, the following systems were simulated by molecular dynamics (MD): 1) PGHS-1, complexed to AA in the catalytic orientation; 2) PGHS-1, complexed to AA in the alternative orientation; 3) PGHS-2, complexed to AA in the catalytic orientation; and 4) PGHS-2, complexed to AA in the alternative orientation. In all systems the PGHSs were considered in its proper fully glycosylated forms.

### b. *Substrate and inhibitors orientations*

In order to unequivocally characterize the orientations of AA binding to PGHSs in residues Tyr385 and Ser530, the electron density maps for the crystallographic structures found in PDB for both alternative and catalytic orientations were regenerated using EDS [28] (Figure 1). Accordingly, the data clearly shows the hydrogen binding of AA carboxylate group to Tyr355 (3.1 Å) and Arg120 (2.4 Å), for the catalytic orientation, and Tyr385 (2.7 Å) and Ser530 (2.7 Å), for the alternative orientation. While such an alternate complex for AA is not related in literature to biological phenomena, even though it is observed in two different crystallographic structures (PDB IDs: 1CVU and 3HS5), it is not only related to AA. Also, crystallographic data for the PGHSs inhibitors demonstrated that some of them present interactions with Tyr385 and Ser530 in a similar way to the alternative orientation of AA (Figure 1). Diclofenac binds to PGHS-2 with its carboxylic group hydrogen-bound to Tyr385 and Ser530, in the first experimental demonstration that the carboxylic group of an NSAID can bind to a PGHS enzyme in an orientation that precludes the formation of a salt bridge with Arg120 [15]. Additional mutagenesis data also suggest the close relation of Ser530 to the binding of other NSAID, such as piroxicam and nimesulid [15].

The presence of potential multiple orientations, is not only observed for AA, but for other compounds complexed to PGHSs. For example, interaction of SC-558 and celecoxib with PGHS appears to be unrelated to the bioisosteric relationship between carboxylate and sulfonamide group. Instead, the orientation of SC-558, found in crystallographic data, demonstrated that the trifluoromethyl group is interacting with Arg120 [16] (Figure 1). Also, celecoxib had recently been resolved by x-ray crystallography bound to PGHS-1 [29], demonstrating that it interacts similarly to SC-558 in the crystallographic pose (Figure 1). Still, conformational analysis of SC-558 bound to PGHSs suggested an alternative orientation for this inhibitor, where the sulfonamide group interacts with Arg120 [30], which has only been found when the inhibitor was complexed to PGHS-2 and agrees with the bioisosteric relation between carboxylate and sulfonamide groups.

### *c. Catalytic and non-catalytic orientation stability*

During the performed MD simulations, the dynamics of AA within the PGHSs' active site was dependent on its orientation. In the catalytic orientation, the AA remains hydrogen-bonded to Arg120 (Figure 2, A and B), suggesting a stable complex with both PGHSs. Accordingly, the distance between AA carboxylic group and Arg120 guanidine group remains at  $3.2\pm 0.7$  Å during 75% of the simulation when bound to PGHS-2, and  $3.3\pm 0.7$  Å during 40% of the simulation time when bound to PGHS-1 (Figure 3, E, F, G and H), while the

distance between the phenolic oxygen atom from Tyr385 and C13 from AA (Figure 4) is  $7.0 \pm 2.6$  Å for PGHS-2 and  $5.7 \text{Å} \pm 2.0$  for PGHS-1, demonstrating that the orientation reported as catalytic, maintains the interactions reported as necessary for the catalytic conversion of AA to PGG<sub>2</sub> in the performed simulations.

On the other hand, the alternative orientation of AA bound to both PGHSs presents a distinct behavior. The initial binding of the AA carboxylic group to Tyr385 and Ser530, as determined in crystallographic data, is progressively lost in detriment of a new hydrogen bond with Arg120 (Figure 2, C and D), for both PGHS-1 and PGHS-2. Such behavior may be observed in the distances between AA carboxylic group and Arg120 guanidine group, which changes from  $\sim 15$  Å to the proper hydrogen bond distance of  $3.5 \pm 0.7$  Å during 40% of the simulation time, for both PGHSs (Figure 3, A, B, C and D). These results suggest that the alternative orientation of AA is less stable, in the active site, than the catalytic orientation.

#### d. *Implications in drug-design*

The alternative orientation of AA may be related to previous data on NSAIDs interaction to PGHS. In fact, the crystallographic structure of diclofenac complexed to PGHS-2 does present its carboxylic group hydrogen binding to Tyr385 and Ser530 (Figure 1). Simultaneously, its dichlorinated ring is close to Arg120, in a similar way to the orientation of the double bond between C14 and C15 of the AA hydrophobic chain. Such similarity supports the idea that the alternative orientation of AA corresponds to a minor complex able to be observed in solution, and then selected during crystallization. As a possible orientation, already adopted by diclofenac, it points to an additional strategy to the design of new inhibitors of PGHS-2 enzyme, offering a distinct point of view for structure-based development of such compounds.

### 3. Conclusions

The PGHSs non-selective inhibitors have been associated with undesirable side effects, including gastrointestinal toxicity [17-19]. Moreover, PGHS-2 is the isoform rapidly induced by inflammatory cytokines and mitogens and is largely speculated to account for prostaglandin formation in inflammation and cancer [20-22] and these selective PGHS-2 inhibitors were developed for treating pain and inflammation in an attempt to reduce the gastrointestinal side effects. Nevertheless, some coxibs were linked to increased risk of blood clotting, heart attack and stroke [23-25], leading to the study of new NSAIDs like NO-

NSAIDs, an emerging class of PGHS inhibitors capable of improving cardiovascular safety as well as reducing gastrointestinal toxicity [26,27].

Our results demonstrated that the alternative orientation of the PGHSs substrate AA found in crystallographic data does not seem to correspond to a stable orientation when the enzyme is placed in aqueous solution, suggesting that this orientation corresponds to a minor orientation of AA in solution. Although the alternative bound AA does not seem to be stable in the catalytic site, previous MD studies demonstrated that inhibitors like SC-558 can interact and inhibit the enzyme in more than one orientation [30]. Crystallographic orientation for diclofenac demonstrates that the interactions with residues Tyr385 and Ser530 are also a target for PGHS inhibition, in addition to the most common interaction with Arg120, indicating that there is more than one important region at the COX active site where the inhibitors interact. Since new NSAIDs are being studied, aiming for compounds that present lesser side effects, the results found in this study may provide new insights to the search of new PGHS inhibitors.

#### 4. Experimental section

##### *a. Glycosylated PGHSs models*

Crystallographic structures for PGHS-1 and PGHS-2 have been obtained from PDB under IDs 1Q4G [31] and 1CVU [12], respectively. These structures were isolated to their monomeric forms, in which the N-glycosylation sites (Asn68, Asn144, and Asn410) were filled with a putative model (proposed to be based on a consensus between previous biochemical, crystallographic and mass spectrometry data) [6,7] for their glycan moieties and using the glycoscience modeling tools [32], as previously described [8].

##### *b. Electron density maps regeneration*

Electron density maps have been regenerated using the EDS [28] for crystallographic structures PDB ID: 1DIY [10], which contained the AA in the alternative orientation, and PDB ID: 1CVU [12], containing the AA in the catalytic orientation.

##### *c. MD simulations*

All calculations had been performed using GROMACS simulation suite [33] and GROMOS96 force field [34]. Briefly, glycosylated PGHS-1 and -2 monomers complexed to AA, in its two possible orientations, were solvated with the SPC water model and submitted

to an energy minimization using the Steepest Descents algorithm. Counter ions were added to neutralize the systems and the Lincs method [35] was applied to constrain covalent bond lengths, allowing an integration step of 2 fs after an initial energy minimization. The systems were then submitted to MD simulations, where position restraints were used on the MBD, in order to mimic the membrane environment. This was achieved by employing backbone constraints of  $1,000 \text{ kJ.mol}^{-1}$  constant force on the MBD. The simulations were carried out to 50 ns, employing the Particle-Mesh Ewald method [36]. Temperature and pressure were kept constant by coupling protein and solvent to external temperature and pressure baths with coupling constants of  $\tau=0.1$  and 0.5 ps, respectively [37]. The reference temperature of the systems was adjusted to 310 K after a slowly heating from 50 to 310 K, in steps of 5 ps, each one increasing the reference temperature by 50 K. A reference value of 3.5 Å between heavy atoms was considered for hydrogen-bonds and a cut-off angle of 30° was used between hydrogen-donor-acceptor [33].

## 5. References

- [1] D. L. DeWitt, T. E. Rollins, J. S. Day, J. A. Gauger, W. L. Smith, *J. Biol. Chem.* 256 (1981) 10375-10382.
- [2] W. L. Smith, D. L. DeWitt, Prostaglandin endoperoxide synthases-1 and -2, in: *Advances in Immunology*, ed. FJ Dixon, Academic Press, Orlando, FL, 1996, pp. 167–215.
- [3] D. L. DeWitt, E. A. Meade, *Arch. Bioche. Biophys.* 306 (1993) 94-102.
- [4] W. Xie, J. G. Chipman, D. L. Robertson, R. L. Erikson, D. L. Simmons, *Proc. Natl. Acad. Sci.* 88 (1991) 2692-2696.
- [5] D. Picot, P. J. Loll, R. M. Garavito, *Nature* 367 (1994) 243-249.
- [6] J. H. G. M. Mutsaers, H. van Halbeek, J. P. Kamerling, J. F. G. Vliegthart, *Eur. J. Biochem.* 147 (1985) 569-574.
- [7] J. F. Nemeth, G. P. Hochesang, L. J. Marnett, R. M. Caprioli, *Biochemistry* 40 (2001) 3109-3116.

- [8] C. L. Fernandes, L. G. Sachett, L. Pol-Fachin, H. Verli, *Carbohydr. Res.* 345 (2010) 663–671.
- [9] J. C. Otto, D. L. DeWitt, W. L. J. Smith, *J. Biol. Chem.* 268 (1993) 18234-18242.
- [10] M. G. Malkowski, S. Ginell, W. L. Smith, R. M. Garavito, *Science* 289 (2000) 1933–1937.
- [11] C. A. Rouzer, L. J. Marnett, *Chem. Rev.* 103 (2003) 2239-2304.
- [12] J. R. Kiefer, J. L. Pawlitz, K. T. Moreland, R. A. Stegeman, W. F. Hood, J. K. Gierse, A. M. Stevens, D. C. Goodwin, S. W. Rowlinson, L. J. Marnett, W. C. Stallings, R. G. Kurumbail, *Nature* 405 (2000) 97-101.
- [13] A. J. Vecchio, D. M. Simmons, M. G. Malkowski, *J. Biol. Chem.* 285 (2010) 22152-22163.
- [14] K. E. Furse, D. A. Pratt, N. A. Porter, T. P. Lybrand, *Biochemistry* 45 (2006) 3189-3205.
- [15] S. W. Rowlinson, J. R. Kiefer, J. J. Prusakiewicz, J. L. Pawlitz, K. R. Kozak, A. S. Kalgutkar, W. C. Stallings, R. G. Kurumbail, L. J. Marnett, *J. Biol. Chem.* 278 (2003) 45763-45769.
- [16] R. G. Kurumbail, A. M. Stevens, J. K. Gierse, J. J. McDonald, R. A. Stegeman, J. Y. Pak, D. Gildehaus, J. M. Miyashiro, T. D. Penning, K. Seibert, P. C. Isakson, W. C. Stallings, *Nature* 384 (1996) 644-648.
- [17] E. A. Meade, W. L. Smith, D. L. DeWitt, *J. Biol. Chem.* 268 (1993) 6610-6614.
- [18] J. R. Vane, R. M. Botting, *Inflamm. Res.* 44 (1995) 1-10.
- [19] T. D. Warner, F. Giuliano, I. Vojnovic, A. Bukasa, J. A. Mitchell, J. R. Vane, *Proc. Natl. Acad. Sci. U.S.A.* 96 (1999) 7563-7568.
- [20] B. Agarwal, C. V. Rao, S. Bhendwal, W. R. Ramey, H. Shirin, B. S. Reddy, P. R. Holt, *Gastroenterology* 117 (1999) 838–847.

- [21] G. Steinbach, P. M. Lynch, R. K. Phillips, M. H. Wallace, E. Hawk, G. B. Gordon, N. Wakabayashi, B. Saunders, Y. Shen, T. Fujimura, L. K. Su, B. Levin, *N. Engl. J. Med.* 342 (2000) 1946–1952.
- [22] M. J. Thun, S. J. Henley, C. Patrono, *J. Natl. Cancer Inst.* 94 (2002) 252-266.
- [23] C. Bombardier, L. Laine, A. Reicin, D. Shapiro, R. Burgos-Vargas, B. Davis, R. Day, M. B. Ferraz, C. J. Hawkey, M. C. Hochberg, T. K. Kvien, T. J. Schnitzer, *N. Engl. J. Med.* 343 (2000) 1520-1528.
- [24] T. Grosser, S. Fries, G. A. FitzGerald, *J. Clin. Invest.* 116 (2006) 4-15.
- [25] L. J. Marnett, *The Ann. Rev. Pharmacol. Toxicol.* 49 (2009) 265-290.
- [26] J. E. Keeble, P. K. Moore, *Br. J. Pharmacol.*, 137 (2002) 295-310.
- [27] E. Koç, S. G. Küçükgülzel, *Mini Rev. Med. Chem.* 9 (2009) 611-619.
- [28] G. J. Kleywegt, M. R. Harris, J. Y. Zou, T. C. Taylor, A. Wählby, T. A. Jones, *Acta Cryst.* 60 (2004) 2240-2249.
- [29] G. Rimon, R. S. Sidhu, D. A. Lauver, J. Y. Lee, N. P. Sharma, C. Yuan, R. A. Frierler, R. C. Trievel, B. R. Lucchesi, W. L. Smith, *Proc. Natl. Acad. Sci. U.S.A.* 107 (2010) 28-33.
- [30] V. Limongelli, M. Bonomi, L. Marinelli, F. L. Gervasio, A. Cavalli, E. Novellino, M. Parrinello, *Proc. Natl. Acad. Sci. U.S.A.* 107 (2010) 5411-5416.
- [31] K. Gupta, B. S. Selinsky, C. J. Kaub, A. K. Katz, P. J. Loll, , *J. Mol. Biol.* 335 (2004) 503-518.
- [32] T. Lütteke, M. Frank, C.-W. von der Lieth, *Carbohydr. Res.* 339 (2004) 1015-1020.
- [33] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* 26 (2005) 1701–1718.
- [34] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krueger, A. E. Mark, W. R. P. Scott, I. G. Tironi, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; University of Groningen, Groningen, and ETH, Zurich, 1996.

- [35] B. Hess, H. Bekker, H. J. Berendsen, J. G. Fraaije, *J. Comput. Chem.* 18 (1997) 1463–1472.
- [36] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* 98 (1993) 10089–10092.
- [37] H. J. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.* 91 (1987) 6269–6271.

## FIGURE CAPTION

**Figure 1.** Orientations of PGHSs substrate AA and PGHSs inhibitors as described in crystallographic data. (A) Catalytic orientation of AA from PDB ID: 1DIY. (B) Alternative orientation of AA from PDB ID: 1CVU. (C) Crystallographic orientations of PGHS inhibitors celecoxib (PDB ID: 3KK6), diclofenac (PDB ID: 1PXX) and SC-558 (PDB ID: 6COX). Electron density maps were regenerated around carboxylic groups of AA and its interacting residues, by using the EDS for (A) and (B).

**Figure 2.** Variation of AA orientation during MD simulations. Conformation of AA in catalytic orientation bound to (A) PGHS-1 and to (B) PGHS-2, and conformation of AA in alternative orientation bound to (C) PGHS-1 and to (D) PGHS-2. (A) and (B) demonstrate the stability of the interaction between the carboxylic group of AA and the guanidine group of Arg120. (C) and (D) show the dislocation of AA towards the guanidine group of Arg120. Images were obtained from the initial (0ns) and the final (50ns) structures of MD.

**Figure 3.** Analysis of the distances between the oxygen atoms from the carboxylic group of AA and the nitrogen atoms from the guanidine group of Arg120. For the alternative orientation of AA, the distances between NH1-O1 (black) and NH1-O2 (red) of AA, bound to (A) PGHS-1 and to (B) PGHS-2 have been determined, as well as the distances between NH2-O1 (black) and NH2-O2 (red) of AA bound to (C) PGHS-1 and to (D) PGHS-2. This analysis demonstrates that there is a dislocation of AA towards the guanidine group of Arg120 during MD. For the catalytic orientation of AA, were measured the distances between NH1-O1 (black) and NH1-O2 (red) of AA bound to (E) PGHS-1 and to (F) PGHS-2, and the distances between NH2-O1 (black) and NH2-O2 (red) of AA, bound to (G) PGHS-1 and to (H) PGHS-2. This data supports the stability of the catalytic orientation of AA.

**Figure 4.** Distance analysis for the catalytic reaction. The stability of the distance between the hydroxyl group of Tyr385 and the carbon 13 (C13) of AA in the catalytic orientation, bound to PGHS-1 (A) and PGHS-2 (B), demonstrates that the catalysis is favored along the MD simulation.

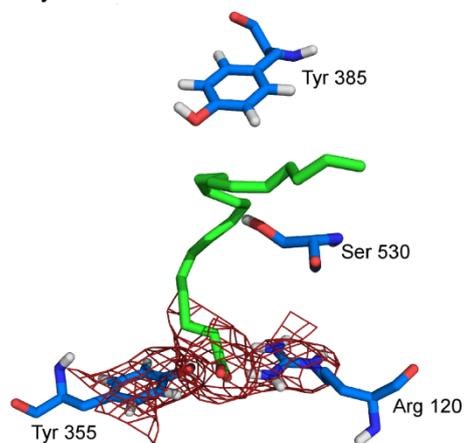
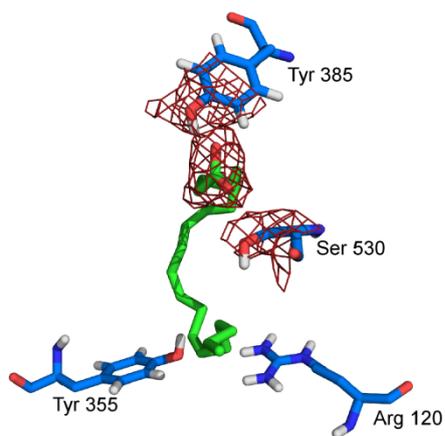
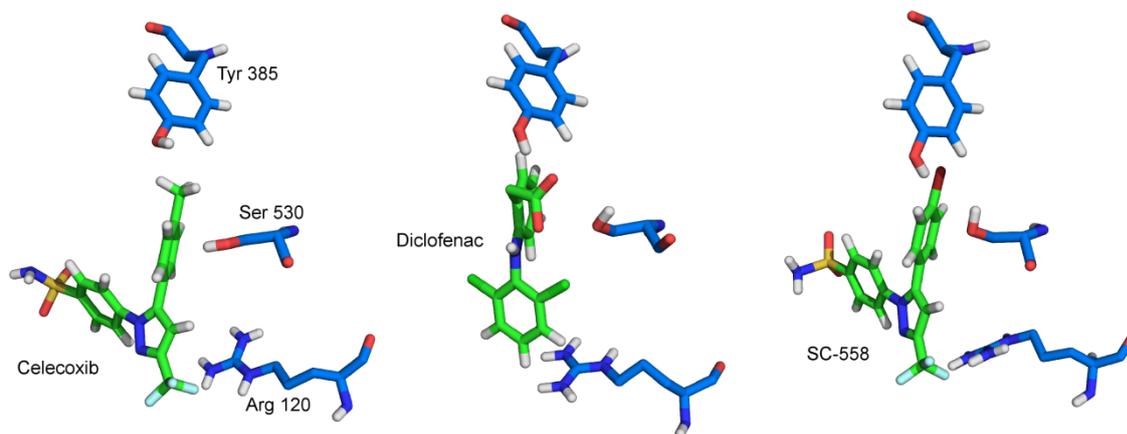
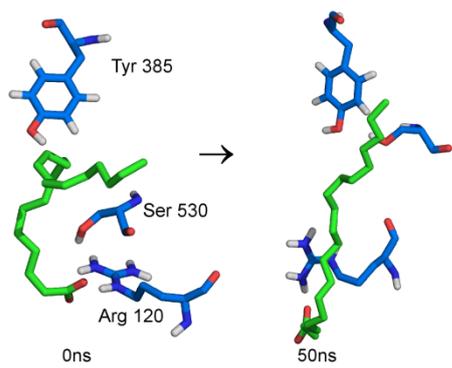
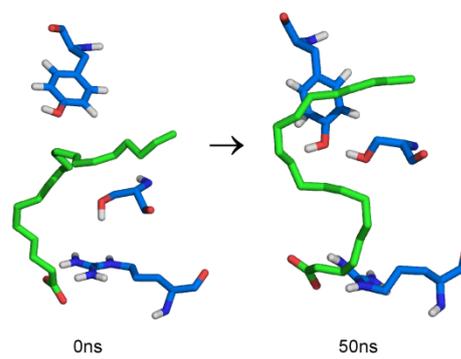
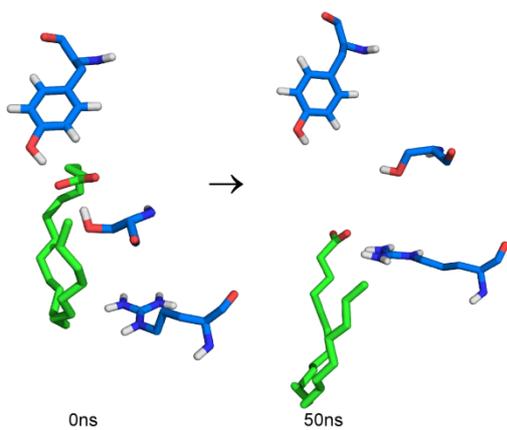
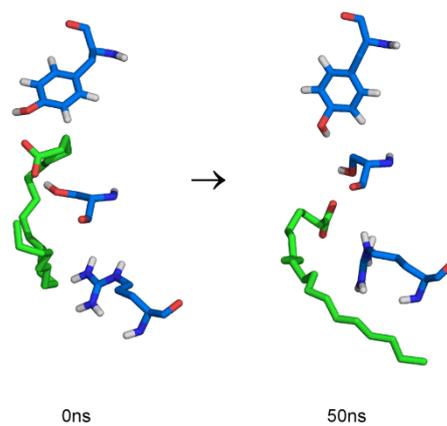
**FIGURES:****A. Catalytic orientation of AA****B. Alternative orientation of AA****C. Orientations of PGHS inhibitors**

FIGURE 1.

**A. PGHS-1: catalytic orientation of AA****B. PGHS-2: catalytic orientation of AA****C. PGHS-1: alternative orientation of AA****D. PGHS-2: alternative orientation of AA****FIGURE 2.**

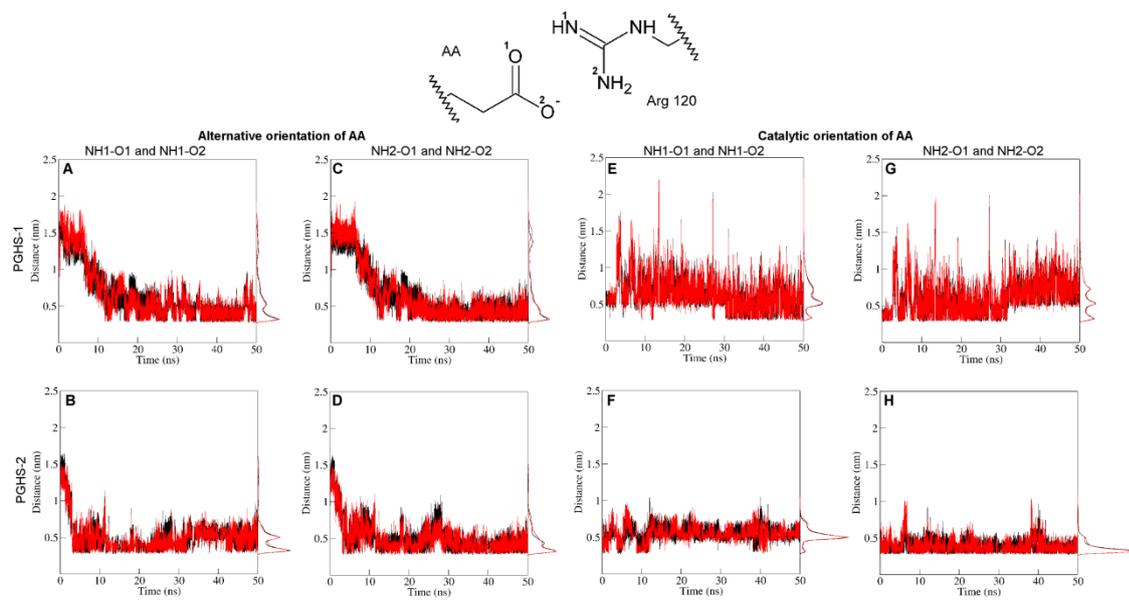


FIGURE 3.

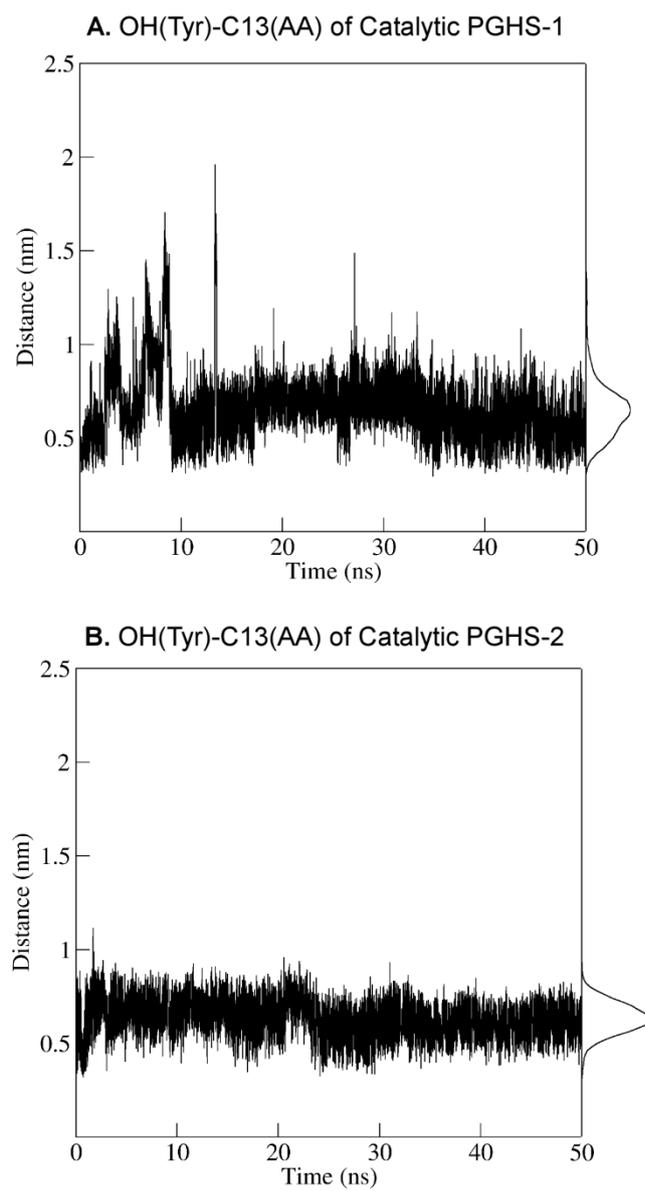


FIGURE 4.

## 5 Discussão Geral

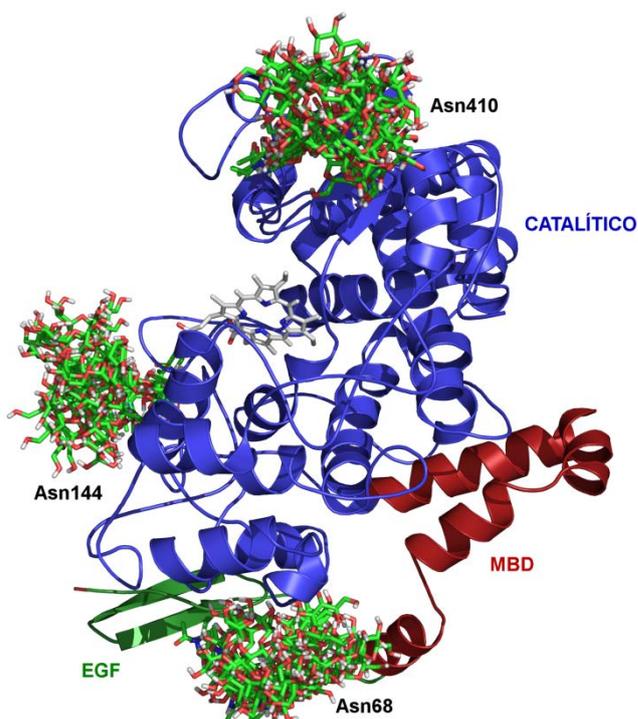
### 5.1 Influência da glicosilação na dinâmica de PGHSs

Existem diversas dificuldades para a obtenção da estrutura de carboidratos por métodos de cristalografia de raios-X e RMN. No contexto do uso de cristalografia de raios-X, a flexibilidade conformacional dos oligossacarídeos, estejam eles livres ou ligados à superfície da proteína, geralmente dificulta a formação de cristais (Bohne-Lang & von der Lieth, 2002; Petrescu *et al.*, 2006). Adicionalmente, no caso de glicoproteínas, quando os cristais são obtidos, o mapa de densidade eletrônica é prejudicado pela alta mobilidade dessas glicanas, comprometendo sua correta determinação conformacional e configuracional (Petrescu *et al.*, 1999). Dessa forma, não são incomuns erros na determinação do estado anomérico, conformações distorcidas de monossacarídeos (Lütteke *et al.*, 2004), ou na orientação relativa de dois resíduos unidos através de uma ligação glicosídica (Lütteke, 2009). O emprego de técnicas de RMN, por sua vez, obtém estruturas que representam as conformações observadas em solução para as moléculas em estudo. No entanto, uma determinação conformacional precisa de carboidratos, utilizando essa metodologia, só é possível se forem obtidos três ou mais sinais de NOE inter-resíduos em uma ligação glicosídica (Wormald *et al.*, 2002), o que, na maioria das vezes, não permite a determinação conformacional de oligossacarídeos maiores que uma dezena de resíduos. Adicionalmente, é importante considerar que os sinais de NOE são medidos em uma escala de tempo que varia entre 50 ms e 1 s, o que não afeta a determinação conformacional de moléculas rígidas, mas, em sistemas flexíveis, ou que coexistem em dois ou mais subestados conformacionais, acarreta em resultados que representando propriedades conformacionais médias dessas moléculas (Woods, 1998; Wormald *et al.*, 2002), o que não necessariamente corresponde aos subestados conformacionais existentes em solução. Nesse contexto, as geometrias mais prevalentes, extraídas do perfil conformacional em solução de dissacarídeos, podem ser utilizadas na construção de partes sacarídicas de glicoproteínas, servindo como estruturas de entrada para o estudo destas moléculas por técnicas de modelagem molecular (Pol-Fachin *et al.*, 2009).

A parte sacarídica de uma glicoproteína pode influenciar diversas propriedades da cadeia polipeptídica à qual está ligada, de modo que a correta descrição

conformacional desses oligossacarídeos é essencial no entendimento dos efeitos da glicosilação sobre o arcabouço protéico. A porção sacarídica pode participar diretamente em eventos de reconhecimento, bem como modificar propriedades das proteínas, principalmente devido ao grande tamanho e à alta flexibilidade dos carboidratos, que desta maneira podem encobrir uma grande parte da superfície protéica (Dwek, 1996). Dessa forma, o perfil conformacional da parte protéica das isoformas de PGHSs estudadas foi comparado ao destas isoformas não-glicosiladas a fim de avaliar o efeito da glicosilação sobre a parte protéica destas glicoproteínas.

A DM de monômeros de PGHSs glicosilados demonstrou que a porção sacarídica de fato encobre uma grande parte da superfície protéica (Figura 17). As análises das simulações de DM permitiram observar que as maiores alterações conformacionais nas enzimas ocorreram no domínio EGF, onde a presença da glicosilação diminuiu a flexibilidade deste domínio, estabilizando-o. Este mesmo domínio já havia sido estudado no nosso grupo onde, através das simulações por DM, foi possível observar que oligossacarídeos ligados ao domínio EGF do fator de coagulação fVII promoveram uma estabilização conformacional em regiões de alça próximas ao sítio de glicosilação ou em regiões espacialmente adjacentes da glicana (Pol-Fachin *et al.*, 2009). Perfil semelhante foi também observado para a proteína regulatória do sistema complemento e para a subunidade  $\alpha$  da gonadotrofina coriônica (Pol-Fachin *et al.*, 2009).



**Figura 17:** Sobreposição dos passos da DM (a cada 10 ns) para as oligomanoses nos três sítios de N-glicosilação das PGHSs, demonstrando a alta flexibilidade desta porção sacarídica.

Este efeito de estabilização, contudo, não foi uniforme para toda a proteína. A região da proteína que contém os dois sítios catalíticos se manteve praticamente estável durante as simulações realizadas. Esta pouca variação observada no domínio catalítico das PGHSs quando a enzima estava ou não glicosilada parece estar de acordo com o papel destas estruturas sacarídicas na enzima pois, mesmo sendo este domínio o que apresenta dois dos três sítios de glicosilação considerados, sabe-se que a presença ou ausência de glicosilação não afeta as propriedades catalíticas desta enzima (Otto *et al.*, 1993).

Dessa forma a metodologia utilizada, baseada na utilização de estados conformacionais majoritários de dissacarídeos simulados livres em solução como ponto de partida para estudos conformacionais de glicoproteínas, mostrou-se útil na caracterização da influência da glicosilação sobre a parte protéica das PGHSs. Esta metodologia de DM empregada, que utiliza programas gratuitos e implementados com algoritmos de simulação rápidos, pode contribuir no entendimento dos diversos processos biológicos em que glicoproteínas e polissacarídeos estejam envolvidos.

## **5.2 Caracterização da dinâmica do substrato das enzimas PGHSs em suas diferentes orientações**

A presença de uma orientação alternativa para o substrato das PGHSs, AA, é sugerida como uma orientação não-catalítica, pois o radical tirosil na Tyr385 não está a uma distância capaz de abstrair o hidrogênio do carbono C13 deste substrato (Kiefer *et al.*, 2000; Vecchio *et al.*, 2010). Entretanto, esta orientação tem sido interpretada de modo não-conclusivo, e não existem estudos que caracterizem esta orientação quando a enzima está em solução (Kiefer *et al.*, 2000; Furse *et al.*, 2006; Vecchio *et al.*, 2010).

Os resultados obtidos a partir dos estudos de DM das isoformas da enzima, complexadas ao seu substrato nas orientações catalítica e alternativa, permitiram observar um comportamento diferente entre elas. Quando o substrato encontra-se na orientação catalítica no sítio ativo das PGHSs, pode-se observar que o AA permaneceu ligado à Arg120, bem como manteve uma distância de seu carbono C13 com o resíduo de Tyr385 compatível com o processo catalítico, sugerindo um complexo estável do substrato com ambas isoformas da enzima.

Diferentemente, a orientação alternativa do AA, quando complexado a ambas PGHSs, apresentou um comportamento distinto daquele observado na orientação catalítica. A ligação inicial do grupamento carboxílico aos resíduos Tyr385 e Ser530 foi progressivamente perdida em detrimento de uma nova ligação de hidrogênio com o resíduo Arg120 em ambas isoformas da enzima. Estes resultados sugerem que a orientação alternativa é menos estável no sítio ativo do que a orientação catalítica.

Ainda, a orientação alternativa do substrato AA no sítio COX da enzima pode ser comparada à orientação do AINE diclofenaco, que também apresenta uma estrutura cristalográfica onde seu grupamento carboxílico está interagindo com a Tyr385 e Ser530 (Rowlinson *et al.*, 2003). Esta similaridade corrobora a idéia de que a orientação alternativa do AA corresponde a um complexo minoritário em solução que, mesmo sendo menos abundante, pode indicar novas estratégias de planejamento de inibidores ao indicar novas formas de complexação e interações.

Além do diclofenaco, estudos conformacionais demonstraram que outro inibidor da enzima, SC-558, apresenta mais de uma orientação capaz de inibir a enzima no sítio COX (Limongelli *et al.*, 2010). Assim, há claros indícios de que existe mais de uma região importante para interação de fármacos no sítio COX da enzima. Desta maneira, considerando a constante busca por novos AINES capazes de inibirem as

PGHSs que apresentem menores efeitos colaterais, os resultados obtidos no estudo podem fornecer novos *insights* para a busca destes compostos.

## 6 Conclusões

A partir dos objetivos traçados e dos resultados obtidos pode-se concluir que:

- Os dados obtidos indicam uma diminuição da flexibilidade, na presença da glicosilação, no domínio EGF da enzima em ambas as isoformas;
- O emprego de estados conformacionais de dissacarídeos em solução como blocos para a construção de glicanas complexas ligadas a glicoproteínas se mostrou como um protocolo capaz de reproduzir adequadamente geometrias de carboidratos determinadas por RMN;
- A análise das diferentes orientações do ácido araquidônico demonstrou que a orientação alternativa (não catalítica) parece não ser estável no sítio de ligação, havendo uma tendência de o substrato se deslocar para interagir com a Arg120;
- A interconversão da forma não catalítica para uma forma semelhante à catalítica, durante a DM, pode indicar que a orientação cristalográfica alternativa seja uma orientação minoritária em solução.

Globalmente, os resultados obtidos permitiram caracterizar as duas isoformas de PGHS e contribuir para a validação de um protocolo de DM para o estudo de glicoproteínas, bem como para o entendimento da interação da enzima com as diferentes orientações do substrato, de maneira a auxiliar a busca por novos inibidores desta enzima.

## 7 Perspectivas

Considerando-se os resultados obtidos, as seguintes perspectivas podem ser traçadas:

- Caracterização da interação de inibidores de PGHSs que possuam mais de uma possível orientação de interação no sítio ativo;
- Caracterização do efeito da dimerização da enzima;
- Suporte estrutural para o planejamento de novos agentes anti-inflamatórios.

## 8 Referências Bibliográficas

### Referências citadas no texto e nos artigos publicados/submetidos

- Agarwal, B.; Rao, C. V.; Bhendwal, S.; Ramey, W. R.; Shirin, H.; Reddy, B. S.; Holt, P. R.: Lovastatin augments sulindac-induced apoptosis in colon cancer cells and potentiates chemopreventive effect of sulindac. *Gastroenterology*, **1999**, *117*, 838–847.
- Allinger, N. L.; Yuh, Y. H.; Lii, J. -H.: Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.*, **1989**, *111*, 8551-8567.
- Allinger, N. L.; Rahman, M.; Lii, J. -H.: A molecular mechanics force field (MM3) for alcohols and ethers. *J. Am. Chem. Soc.*, **1990**, *112*, 8293-8307.
- Becker, C. F.; Guimarães, J. A.; Verli, H.: Molecular dynamics and atomic charge calculations in the study of heparin conformation in aqueous solution. *Carbohydr. Res.*, **2005**, *340*, 1499-1507.
- Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P.: The missing term in effective pair potentials. *J. Phys. Chem.*, **1987**, *91*, 6269-6271.
- Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R.: Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.*, **1984**, *81*, 3684-3690.
- Berman, H. M.; Henric, K.; Nakamura, H.; Markley, J.: Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.*, **2007**, *14*, 354-355.
- Bluhm, T. L.; Deslandes, Y.; Marchessault, R. H.; Perez, S.; Rinaudo, M.: Solid-state and solution conformation of scleroglucan. *Carbohydr. Res.*, **1982**, *100*, 117-130.
- Bohne-Lang, A.; von der Lieth, C. W.: Glycosylation of proteins: a computer based method for the rapid exploration of conformational space of N-glycans. *Pac. Symp. Biocomput.*, **2002**, *7*, 285-296.
- Bombardier, C.; Laine, L.; Reicin, A.; Shapiro, D.; Burgos-Vargas, R.; Davis, B.; Day, R.; Ferraz, M. B.; Hawkey, C. J.; Hochberg, M. C.; Kvien, T. K.; Schnitzer, T. J.: Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. *N. Engl. J. Med.*, **2000**, *343*, 1520-1528.
- Brant, D. A.: Shapes and motions of polysaccharide chains. *Pure Appl. Chem.*, **1997**, *69*, 1885-1892.
- Brater, D. C.; Harris, C.; Redfern, J.S.; Gertz, B.J.: Renal effects of COX-2-selective inhibitors. *Am. J. Nephrol.*, **2001**, *21*, 1-15.

- Case, D. A.; Cheatham, T. E. 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M. Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J.: The Amber Biomolecular Simulation Programs. *J. Comput. Chem.*, **2005**, *26*, 1668-1688.
- Castro, M. O., Pomin, V. H., Santos, L. L.; Vilela-Silva, A. C. E. S., Hirohashi, N., Pol-Fachin, L., Verli, H. & Mourão, P. A. S.: A unique 2-sulfated  $\beta$ -galactan from the egg jelly of the sea urchin *Glyptocidaris crenularis*: Conformational flexibility versus induction of the sperm acrosome reaction *J. Biol. Chem.*, **2009**, *284*, 18790-18800.
- Chandrasekharan, N. V.; Dai, H., Roos, K. L. T., Evanson, N. K., Tomsik, J., Elton, T. S.; Simmons, D. L: From the cover: COX-3, a cyclooxygenase-1 variant inhibited by acetaminophen and other analgesic/antipyretic drugs: cloning, structure and expression. *Proc Natl Acad Sci USA*, **2002**, *99*, 13926–13931.
- Clark, M.; Cramer, R. D.; van Opdenbosch, N.: Validation of the general-purpose Tripos 5.2 force-field. *J. Comput. Chem.*, **1989**, *10*, 982-1012.
- Crispin, M.; Stuart, D. I.; Jones, E. Y.: Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.*, **2007**, *14*, 354.
- Cumming, D. A.; Carver, J. P.: Virtual and solution conformations of oligosaccharides. *Biochemistry*, **1987**, *26*, 6664-6676.
- Cumming, D. A.; Carver, J. P.: Reevaluation of rotamer population for 1,6 linkages: reconciliation with potential energy calculations. *Biochemistry*, **1987**, *26*, 6676-6683.
- Darden, T.; York, D.; Pedersen, L.: Particle Mesh Ewald – an  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.*, **1993**, *98*, 10089-10092.
- Dayasu, H.; Toh, H.: Molecular Evolution of the Myeloperoxidase Family. *J. Mol. Evol.*, **2000**, *51*, 433-445.
- de Sant'Anna, C. M. R.: Glossário de termos usados no planejamento de fármacos (recomendações IUPAC 1997). *Quim. Nova*, **2002**, *25*, 505-512.
- DeLano, W. L.: *The PyMOL Molecular Graphics System*, DeLano Scientific, Sao Carlos, CA, USA, **2002**. <http://www.pymol.org>
- DeWitt, D. L.; Meade, E. A.: Serum and glucocorticoid regulation of gene transcription and expression of the prostaglandin H synthase-1 and prostaglandin H synthase-2 isozymes. *Arch. Bioche. Biophys.*, **1993**, *306*, 94-102.

- DeWitt, D. L.; Rollins, T. E.; Day, J. S.; Gauger, J. A.; Smith, W. L.: Orientation of the active site and antigenic determinants of prostaglandin endoperoxide synthase in the endoplasmic reticulum. *J. Biol. Chem.*, **1981**, *256*, 10375-10382.
- Dinchuk, J. E.; Liu, R. Q.; Trzaskos, J. M.: COX-3: in the wrong frame in mind. *Immunol. Lett.*, **2003**, *86*, 121.
- Dou, W.; Jiao, Y.; Goorha, S.; Raghov, R.; Ballou, L.R.: Nociception and the differential expression of cyclooxygenase-1 (COX-1), the COX-1 variant retaining intron-1 (COX-1v), and COX-2 in mouse dorsal root ganglia (DRG). *Prostaglandins Other Lipid Mediat.*, **2004**, *74*, 29–43.
- Dwek, R. A.: Glycobiology: Toward Understanding the Function of Sugars. *Chem. Rev.*, **1996**, *96*, 683-720.
- Erbel, P. J. A.; Karimi-Nejad, Y.; van Kuik, J. A.; Boelens, R.; Kamerling, J. P.; Vliegthart, J. F. G.: Effects of the N-Linked Glycans on the 3D Structure of the Free  $\alpha$ -Subunit of Human Chorionic Gonadotropin. *Biochemistry*, **2000**, *39*, 6012-6021.
- Fauth, E. V. F.: *Caracterização estrutural e conformacional de toxinas da família das actinoporinas*. Dissertação (Mestrado em Biologia Celular e Molecular), Centro de Biotecnologia do Estado do Rio Grande do Sul, Universidade Federal do Rio Grande do Sul, Porto Alegre, **2010**.
- Fletcher, C. M.; Harrison, R. A.; Lachmann, P. J.; Neuhaus, D.: Structure of a soluble, glycosylated form of the human complement regulatory protein CD59. *Structure*, **1994**, *2*, 185-199.
- Funk, C. D.; Funk, L. B.; Kennedy, M. E.; Pong, A. E.; FitzGerald, G. A.: Human platelet/erythroleukemia cell prostaglandin G/H synthase: cDNA cloning, expression and gene chromosomal assignment. *FASEB J.*, **1991**, *5*, 2304-2312.
- Furse, K. E.; Pratt, D. A.; Porter, N. A.; Lybrand, T. P.: Molecular Dynamics Simulations of Arachidonic Acid Complexes with COX-1 and COX-2: Insights into Equilibrium Behavior. *Biochemistry*, **2006**, *45*, 3189-3205.
- Garavito, R. M.; Picot, D.; Loll, P. J.: Prostaglandin H synthase. *Curr. Opin. Struct. Biol.*, **1994**, *4*, 529-535.
- Hashimoto, K.; Goto, S.; Kawano, S.; Aoki-Kinoshita, K. F., Ueda, N.; Hamajima, M.; Kawasaki, T.; Kanehisa, M.: KEGG as a glycome informatics resource. *Glycobiology*, **2006**, *16*, 63R-70R.

- Hemmingsen, L.; Madsen, D. E.; Esbensen, A. L.; Olsen, L.; Engelsen, S. B.: Evaluation of carbohydrate molecular mechanical force fields by quantum mechanical calculations. *Carbohydr. Res.*, **2004**, *339*, 937-948.
- Hemler, M.; Lands, W. E.: Purification of the cyclooxygenase that forms prostaglandins: demonstration of two forms of iron in the holoenzyme. *J. Biol. Chem.*, **1976**, *251*, 5575-5579.
- 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.
- Humphrey, W.; Dalke, A.; Schulten, K.: VMD – Visual Molecular Dynamics. *J. Mol. Graph.*, **1996**, *14*, 33-38. <http://www.ks.uiuc.edu/Research/vmd/>.
- Imberty, A.; Gerber, S.; Tran, V.; Pérez, S.: Data bank of three-dimensional structures of disaccharides, a tool to build 3-D structures of oligosaccharides. *Glycoconj. J.*, **1990**, *7*, 27-54.
- IUPAC-IUB Commission on Biochemical Nomenclature: Nomenclature for carbohydrates. *Pure Appl. Chem.*, **1996**, *68*, 1919-2008.
- Jones, M. K.; Wang, H.; Peskar, B. M.; Levin, E.; Itani, R. M.; Sarfeh, I. J.; Tarnawski, A. S.: Inhibition of angiogenesis by nonsteroidal anti-inflammatory drugs: insight into mechanisms and implications for cancer growth and ulcer healing. *Nat. Med.*, **1999**, *5*, 1418–1423.
- Jorgensen, W. L.; Tirado-Rives, J.: The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, **1988**, *110*, 1657-1666.
- Kalgutkar, A. S.; Zhao, Z.: Discovery and design of selective cyclooxygenase-2 inhibitors as non-ulcerogenic, anti-inflammatory drugs with potential utility as anti-cancer agents. *Curr. Drug Targets*, **2001**, *2*, 79–106.
- Keeble, J. E.; Moore, P. K.: Pharmacology and potential therapeutic applications of nitric oxide-releasing non-steroidal anti-inflammatory and related nitric oxide-donating drugs. *Br. J. Pharmacol.*, **2002**, *137*, 295-310.
- Kiefer, J. R.; Pawlitz, J. L.; Moreland, K. T.; Stegeman, R. A.; Hood, W. F.; Gierse, J. K.; Stevens, A. M.; Goodwin, D. C.; Rowlinson, S. W.; Marnett, L. J.; Stallings, W. C.; Kurumbail, R. G.: Structural insights into the stereochemistry of the cyclooxygenase reaction. *Nature*, **2000**, *405*, 97-101.

- Kitson, D. H.; Hagler, A. T.: Theoretical-studies of the structure and molecular-dynamics of a peptide crystal. *Biochemistry*, **1988**, *27*, 5246-5257.
- Koç, E.; Küçükgülzel, S. G.: Medicinal chemistry and anti-inflammatory activity of nitric oxide-releasing NSAID drugs. *Mini Rev. Med. Chem.*, **2009**, *9*, 611-619.
- Koshkin, V.; Dunford, H.B.: Coupling of the peroxidase and cyclooxygenase reactions of prostaglandin H synthase. *Biochim. Biophys. Acta*, **1999**, *1430*, 341-348.
- Kurumbail, R.; Stevens, A.; Gierse, J.; Mc-Donald, J.; Stegeman, R. A.; Pak, J. Y.; Gildehaus, D.; Miyashiro, J. M.; Penning, T. D.; Seibert, K.; Isakson, P. C.; Stallings, W. C.: Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature*, **1996**, *384*, 644–648.
- Kuttel, M. M.; Naidoo, K. J.: Free Energy Surfaces for the  $\alpha(1-4)$ -Glycosidic Linkage: Implications for Polysaccharide Solution Structure and Dynamics. *J. Phys. Chem. B*, **2005**, *109*, 7468-7474.
- Leach, A. R.: *Molecular Modelling Principles and Applications*, 2<sup>nd</sup> Ed., **2001**, Longman, Cingapura.
- Limongelli, V.; Bonomi, M.; Marinelli, L.; Gervasio, F. L.; Cavalli, A.; Novellino, E.; Parrinello, M.: Molecular basis of cyclooxygenase enzymes (COXs) selective inhibition. *Proc. Natl. Acad. Sci. U.S.A.*, **2010**, *107*, 5411-5416.
- Lipkind, G. M.; Verovsky, V. E.; Kochetkov, N. K.: Conformational states of cellobiose and maltose in solution: A comparison of calculated and experimental data. *Carbohydr. Res.*, **1984**, *133*, 1-13.
- Lipkind, G. M.; Shashkov, A. S.; Kochetkov, N. K.: Nuclear overhauser effect and conformational states of cellobiose in aqueous solution. *Carbohydr. Res.*, **1985**, *141*, 191-197.
- Lütteke, T.; Frank, M.; von der Lieth, C. -W.: Data mining the protein data bank: automatic detection and assignment of carbohydrate structures. *Carbohydr. Res.*, **2004**, *339*, 1015-1020.
- Lütteke, T.; von der Lieth, C. -W.: pdb-care (PDB CARbohydrate RESidue check): a program to support annotation of complex carbohydrate structures in PDB files. *Bioinformatics*, **2004**, *5*, 69.
- Lütteke, T., Frank, M.; von der Lieth, C. -W.: Carbohydrate Structure Suite (CSS): analysis of carbohydrate 3D structures derived from the PDB. *Nucleic Acids Res.*, **2005**, *33*, D242-D246.

- Lütteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C. -W.: GLYCOSCIENCES.de: an Internet portal to support glycomics and glycobiology research. *Glycobiology*, **2006**, *16*, 71R-81R.
- Lütteke, T.: Analysis and validation of carbohydrate three-dimensional structures. *Acta Crystallogr. D Biol. Crystallogr.*, **2009**, *65*, 156-168.
- MackKerell, A. D. Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L. Jr.; 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. 3<sup>rd</sup>; 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.
- MackKerell, A. D. Jr.: Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.*, **2004**, *25*, 1584-1604.
- Malkowski, M. G.; Ginell, S.; Smith, W. L.; Garavito, R. M.: The x-ray structure of prostaglandin endoperoxide H synthase-1 complexed with arachidonic acid. *Science*, **2000**, *289*, 1933-1937.
- Marnett, L. J.: The COXIB Experience: A Look in the Rearview Mirror. *Annu. Rev. Pharmacol. Toxicol.*, **2009**, *49*, 265-290.
- Mbonye, U. R.; Yuan, C.; Harris, C. E.; Sidhu, R. S.; Song, I.; Arakawa, T.; Smith, W. L.: Two distinct pathways for cyclooxygenase-2 protein degradation. *J. Biol. Chem.*, **2008**, *283*, 8611-8623.
- McCammon, J. A.; Gelin, B. R.; Karplus, M.: Dynamics of folded proteins. *Nature*, **1977**, *267*, 585-590.
- McGeer, P. L.; McGeer, E. G.: Inflammation of the brain in Alzheimer's disease: implications for therapy. *J. Leukoc. Biol.*, **1999**, *65*, 409-415.
- Meade, E. A.; Smith, W. L.; DeWitt, D. L.: Differential inhibition of prostaglandin endoperoxide synthase (cyclooxygenase) isozymes by aspirin and other non-steroidal anti-inflammatory drugs. *J. Biol. Chem.*, **1993**, *268*, 6610-6614.
- Mitchell, J. A.; Akarasereenont, P.; Thiemermann, C.; Flower, R. J.; Vane, J. R.: Selectivity of nonsteroidal anti-inflammatory drugs as inhibitors of constitutive and inducible cyclooxygenase. *Proc. Natl. Acad. Sci. USA*, **1993**, *90*, 11693-11697.

- Mutsaers, J. H. G. M.; van Halbeek, H.; Kamerling, J. P.; Vliegthart, J. F. G.: Determination of the structure of the carbohydrate chains of prostaglandin endoperoxide synthase from sheep. *Eur. J. Biochem.*, **1985**, *147*, 569-574.
- Naidoo, K. J.; Denysyk, D.; Brady, J. W.: Molecular dynamics simulations of the N-linked oligosaccharide of the lectin from *Erythrina corallodendron*. *Protein Eng.*, **1997**, *10*, 1249-1261.
- Nemeth, J. F.; Hochensang, G. P.; Marnett, L. J.; Caprioli, R. M.: Characterization of the glycosylation sites in cyclooxygenase-2 using mass spectrometry. *Biochemistry*, **2001**, *40*, 3109-3116.
- Otto, J. C.; DeWitt, D. L.; Smith, W. L.: N-glycosylation of prostaglandin endoperoxide synthases-1 and -2 and their orientations in the endoplasmic reticulum. *J. Biol. Chem.*, **1993**, *268*, 18234-18242.
- Paganini-Hill, A.: Aspirin and colorectal cancer: the Leisure World cohort revisited. *Prev. Med.*, **1995**, *24*, 113-115.
- Patrignani, P.; Panara, M. R.; Greco, A.; Fusco, O.; Natoli, C.; Iacobelli, S.; Cipollone, F.; Granci, A.; Créminon, C.; Maclouf, J.: Biochemical and pharmacological characterization of the cyclooxygenase activity of human blood prostaglandin endoperoxide synthases. *J. Pharmacol. Exp. Ther.*, **1994**, *271*, 1705-1712.
- Perera L.; Darden, T. A.; Pedersen, L. G.: Predicted Solution Structure of Zymogen Human Coagulation FVII. *J. Comput. Chem.*, **2002**, *23*, 35-47.
- Pérez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J.-F.; van Eyck, B. P.; Johnson, G.; French, A. D.; Kouwijzer, M. L. C. E.; Grootenuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K.: A comparison and chemometric analysis of several molecular mechanics force fields and parameter sets applied to carbohydrates. *Carbohydr. Res.*, **1998**, *314*, 141-155.
- Pérez, S.; Mulloy, B.: Prospects for glycoinformatics. *Curr. Opin. Struc. Biol.*, **2005**, *15*, 517-524.
- Petrescu, A. J.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R.: A statistical analysis of N- and O-glycan linkage conformations from crystallographic data. *Glycobiology*, **1999**, *9*, 343-352.

- Petrescu, A. J.; Wormald, M. R.; Dwek, R. A.: Structural aspects of glycomes with a focus on N-glycosylation and glycoprotein folding. *Curr. Opin. Struct. Biol.*, **2006**, *16*, 600-607.
- Picot, D.; Loll, P. J.; Garavito, R. M.: The X-ray crystal structures of the membrane protein prostaglandin H<sub>2</sub> synthase-1. *Nature*, **1994**, *367*, 243-249.
- Pol-Fachin, L.: *Descrição conformacional de carboidratos e glicoproteínas: validação de protocolo baseado em dinâmica molecular e implicações funcionais*. Dissertação (Mestrado em Biologia Celular e Molecular), Centro de Biotecnologia do Estado do Rio Grande do Sul, Universidade Federal do Rio Grande do Sul, Porto Alegre, **2009**.
- Pol-Fachin, L.; Fernandes, C. L.; Verly, H.: GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations. *Carbohydr. Res.*, **2009**, *344*, 491-500.
- Pol-Fachin, L.; Verli, H.: Depiction of the forces participating in the 2-O-sulfo- $\alpha$ -L-iduronic acid conformational preference in heparin sequences in aqueous solutions. *Carbohydr. Res.*, **2008**, *343*, 1435-1445.
- Ponder, J. W.; Case, D. A.: Force fields for protein simulations. Em *Advances in protein chemistry*, volume 66: protein simulations. Richards, F. M.; Eisenberg, D. S.; Kuriyan, J. Eds., Elsevier Academic Press, San Diego, **2003**, pp 27-85.
- Pope, L.: Modeling carbohydrate conformations from NMR data: maximum entropy rotameric distribution about the C5-C6 bond in gentiobiose. *J. Am. Chem. Soc.*, **1993**, *115*, 8421-8426.
- Qin, N.; Zhang, S.-P.; Reitz, T.L.; Mei, J.M.; Flores, C.M.: Cloning, expression and functional characterization of human COX-1 splicing variants: evidence for intron 1 retention. *J. Pharmacol. Exp. Ther.*, **2005**, *315*, 1298-1305.
- Rao, V. S. R.; Qasba, P. K.; Balaji, P. V.; Chandrasekaran, R.: *Conformation of Carbohydrates*, 1.ed. Amsterdam: Harwood Academic Publishers, **1998**.
- Ress, D. A.; Scott, W. E.: Polysaccharide conformation. Part VI. *J. Chem. Soc. B*, **1971**, 469-479.
- Riendeau, D.; Percival, M. D.; Brideau, C.; Charleson, S.; Dube', D.; Ethier, D.; Falguyret, J.-P.; Friesen, R. W.; Gordon, R.; Greig, G.; Guay, J.; Mancini, J.; Ouellet, M.; Wong, E.; Xu, L.; Boyce, S.; Visco, D.; Girard, Y.; Prasit, P.; Zamboni, R.; Rodger, I. W.; Gresser, M.; Ford-Hutchinson, A. W.; Young, R. N.; Chan, C.-C.: Etoricoxib (MK-0663): preclinical profile and comparison with

- other agents that selectively inhibit cyclooxygenase-2. *J. Pharmacol. Exp. Ther.*, **2001**, 296, 558-566.
- Roche, V. F.: A Receptor-Grouped Approach to Teaching Nonsteroidal Antiinflammatory Drug Chemistry and Structure-Activity Relationships. *Am. J. Pharm. Educ.*, **2009**, 73, article 143.
- Rollins, T. E.; Smith, W. L.: Subcellular localization of prostaglandin-forming cyclooxygenase in Swiss mouse 3T3 fibroblasts by electron microscopic immunocytochemistry. *J. Biol. Chem.*, **1980**, 255, 4872-4875.
- Rouzer, C. A.; Marnett, L. J.: Mechanism of free radical oxygenation of polyunsaturated fatty acids by cyclooxygenases. *Chem. Rev.*, **2003**, 103, 2239-2304.
- Rowlinson, S. W.; Kiefer, J. R.; Prusakiewicz, J. J.; Pawlitz, J. L.; Kozak, K. R.; Kalgutkar, A. S.; Stallings, W. C.; Kurumbail, R. G.; Marnett, L. J.: A novel mechanism of cyclooxygenase-2 inhibition involving interactions with Ser-530 and Tyr-385. *J. Biol. Chem.*, **2003**, 278, 45763-45769.
- Sathyanarayana, B. K.; Rao, V. S. R.: Conformational studies of  $\beta$ -glucans. *Biopolymers*, **1972**, 11, 1379-1394.
- Sathyanarayana, B. K.; Stevens, E. S.: Theoretical study of the conformations of pustulan [(1 $\rightarrow$ 6)- $\beta$ -D-glucan]. *J. Biomol. Struct. Dyn.*, **1983**, 1, 947-959.
- Schaftenaar, G.; Noordik, J. H.: MOLDEN: a pre- and post-processing program for molecular and electronic structures. *J. Comput. Aided Mol. Des.*, **2000**, 14, 123-134.
- Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A.: General atomic and molecular electronic structure system. *J. Comp. Chem.*, **1993**, 14, 1347-1363.
- Schüttelkopf, A. W.; van Aalten, D. M. F.: PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr.*, **2004**, D60, 1355-1363.
- Seibold, S. A.; Cerda, J. F.; Mulichak, A. M.; Song, I.; Garavito, R. M.; Arakawa, T.; Smith, W. L.; Babcock, G. T.: Peroxidase activity in prostaglandin endoperoxide H synthase-1 occurs with a neutral histidine proximal heme ligand. *Biochemistry*, **2000**, 39, 6616-6624.

- Sevigny, M. B.; Li, C. F.; Alas, M.; Hughes-Fulford, M.: Glycosylation regulates turnover of cyclooxygenase-2. *FEBS Lett.*, **2006**, *580*, 6533-6536.
- Shefter, E.; Trueblood, K. N.: The crystal and molecular structure of D(+)-barium uridine-5'-phosphate. *Acta Crystallogr.*, **1965**, *18*, 1067-1077.
- Silverstein, F. E.; Faich, G.; Goldstein, J. L.; Simon, L. S.; Pincus, T.; Whelton, A.; Makuch, R.; Eisen, G.; Agrawal, N. M.; Stenson, W. F.; Burr, A. M.; Zhao, W. W.; Kent, J. D.; Lefkowitz, J. B.; Verburg, K. M.; Geis, G. S.: Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomised controlled trial. Celecoxib Long-term Arthritis Safety Study. *J. Am. Med. Assoc.*, **2000**, *284*, 1247-1255.
- Smith, W. L.; DeWitt, D. L.: Prostaglandin endoperoxide synthases-1 and -2. In: *Adv. Immunol.*, Academic Press, Orlando, FL, pp167-215, **1996**.
- Smith, W. L.; Garavito, R. M.; DeWitt, D. L.: Prostaglandin Endoperoxide H Synthases (Cyclooxygenases)-1 and -2. *J. Biol. Chem.*, **1996**, *271*, 33157-33160.
- Smith, W. L.; DeWitt, D. L.; Garavito, R. M.: Cyclooxygenases: Structural, Cellular, and Molecular Biology. *Annu. Rev. Biochem.*, **2000**, *69*, 145–182.
- Smythies, J.: On the function of neuromelanin. *Proc. R. Soc. London Ser. B*, **1996**, *263*, 487-489.
- Song, X.; Lin, H. P.; Johnson, A. J.; Tseng, P. H.; Yang, Y. T.; Kulp, S. K.; Chen, C. S.: Cyclooxygenase-2, player or spectator in cyclooxygenase-2 inhibitor-induced apoptosis in prostate cancer cells. *J. Natl. Cancer Inst.*, **2002**, *94*, 585-591.
- Steinbach, G.; Lynch, P. M.; Phillips, R. K.; Wallace, M. H.; Hawk, E.; Gordon, G. B.; Wakabayashi, N.; Saunders, B.; Shen, Y.; Fujimura, T.; Su, L. K.; Levin, B.: The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis. *N. Engl. J. Med.*, **2000**, *342*, 1946–1952.
- Sumitani, K.; Kamijo, R.; Toyoshima, T.; Nakanishi, Y.; Takizawa, K.; Hatori, M.; Naguno, M.: Specific inhibition of cyclooxygenase-2 results in inhibition of proliferation of oral cancer cell lines via suppression of prostaglandin E2 production. *J. Oral Pathol. Med.*, **2001**, *30*, 41–47.

- Sureshan, K. M.; Miyasou, T.; Watanabe, Y.: Crystal structure of 1L-1,2:4,5-di-O-isopropylidene-*allo*-inositol; A comparison of its conformation in solid and solution states. *Carbohydr. Res.*, **2004**, 339, 1551-1555.
- Thun, M. J.; Henley, S. J.; Patrono, C.: Nonsteroidal anti-inflammatory drugs as anticancer agents: mechanistic, pharmacologic, and clinical issues, *J. Natl. Cancer Inst.*, **2002**, 94, 252-266.
- van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C.: GROMACS: fast, flexible and free. *J. Comput. Chem.*, **2005**, 26, 1701-1718.
- van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krueger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G.: *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. Vdf Hochschulverlag AG, Zurich, Suíça, **1996**.
- van Gunsteren, W. F.; Bakowies, D.; Baron, R.; Chandrasekhar, I.; Christen, M.; Daura, X.; Gee, P.; Geerke, D. P.; Glättli, A.; Hünenberger, P. H.; Kastenholz, M. A.; Oostenbrink, C.; Schenk, M.; Trzesniak, D.; van der Vegt, N. F.; Yu, H. B.: Biomolecular modeling: Goals, problems, perspectives. *Angew. Chem. Int. Ed. Engl.*, **2006**, 45, 4064-4092.
- Van der Ouderaa, F. J.; Buytenhek, M.; Nugteren, D. H.; Van Dorp, D. A.: Acetylation of prostaglandin endoperoxide synthetase with acetylsalicylic acid. *Eur. J. Biochem.*, **1980**, 109, 1-8.
- Vane, J. R.; Botting, R. M.: New insights into the mode of action of anti-inflammatory drugs. *Inflamm. Res.*, **1995**, 44, 1-10.
- Vane, J. R.; Bakhle, Y. S.; Botting, R. M.: Cyclooxygenases 1 and 2. *Annu. Rev. Pharmacol. Toxicol.* **1998**, 38, 97-120.
- Varki, A.: Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, **1993**, 3, 97-130.
- Vecchio, A. J.; Simmons, D. M.; Malkowski, M. G.: Structural basis of fatty acid substrate binding to cyclooxygenase-2. *J. Biol. Chem.* **2010**, 285, 22152-22163.
- Verli, H.; Guimarães, J. A.: Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. *Carbohydr. Res.*, **2004**, 339, 281-290.

- Verli, H.; Guimarães, J. A.: Insights into the induced fit mechanism in antithrombin–heparin interaction using molecular dynamics simulations. *J. Mol. Graph. Model.*, **2005**, *24*, 203-212.
- Verli, H.: *Tese de Doutorado*. Universidade Federal do Rio Grande do Sul, Brasil, **2005**.
- Virudachalam, R.; Rao, V. S. R.: Theoretical Studies on Peptidoglycans. II. Conformations of the Disaccharide-Peptide Subunit and the Three-Dimensional Structure of Ptidoglycan. *Biopolymers*, **1979**, *18*, 571-589.
- Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A.: Development and testing of a general amber force field.. *J. Comput. Chem.*, **2004**, *25*, 1157-1174.
- Whittle, B. J. R.; Boughton-Smith, N. K.; Moncada, S.; Vane, J. R.: Actions of prostacyclin (PGI<sub>2</sub>) and its product 6-oxo-PGF<sub>1α</sub> on the rat gastric mucosa *in vivo* and *in vitro*. *Prostaglandins*, **1978**, *15*, 955-968.
- Woods, R. J.: Computational carbohydrate chemistry: what theoretical methods can tell us. *Glycoconj. J.*, **1998**, *15*, 209-216.
- Wormald, M. R.; Petrescu, A. J.; Pao, Y. -L.; Glithero, A.; Elliott, T.; Dwek, R. A.: Conformational Studies of Oligosaccharides and Glycopeptides: Complementarity of NMR, X-ray Crystallography, and Molecular Modelling. *Chem. Rev.*, **2002**, *102*, 371-386.
- Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R. N.; Smolyar, A.; Reinherz, E. L.; Wagner, G.: Conformation and Function of the N-linked Glycan in the Adhesion Domain of Human CD2. *Science*, **1995**, *269*, 1273-1278.
- Wyss, D. F.; Wagner, G.: The structural role of sugars in glycoproteins. *Curr. Opin. Biotechnol.*, **1996**, *7*, 409-416.
- Xia, J.; Daly, R. P.; Chuang, F. -C.; Parker, L.; Jensen, J. H.; Margulis, C. J.: Sugar Folding: A Novel Structural Prediction Tool for Oligosaccharides and Polysaccharides 1. *J. Chem. Theory Comput.*, **2007**, *3*, 1620-1628.
- Xia, J.; Daly, R. P.; Chuang, F. -C.; Parker, L.; Jensen, J. H.; Margulis, C. J.: Sugar Folding: A Novel Structural Prediction Tool for Oligosaccharides and Polysaccharides 2. *J. Chem. Theory Comput.*, **2007**, *3*, 1629-1643.

- Xie, W.; Chipman, J. G.; Robertson, D. L.; Erikson, R. L.; Simmons, D. L.: Expression of a mitogen-responsive gene encoding prostaglandin synthase is regulated by mRNA splicing. *Proc. Natl. Acad. Sci.* **1991**, *88*, 2692-2696.
- Xu, Y.; Colletier, J. P.; Jiang, H.; Silman, I.; Sussman J. L.; Weik, M.: Induced-fit or pre-existing equilibrium dynamics? Lessons from protein crystallography and MD simulations on acetylcholinesterase. *Protein Sci.*, **2008**, *17*, 601-605.
- Yoneda, J. D.; Albuquerque, M. G.; Leal, K. Z.; Seidl, P. R.; Wheeler, R. A.; Boesch, S. E.; de Alencastro, R. B.; de Souza, M. C. B. V.; Ferreira, V. F.: Molecular dynamics simulations of a nucleoside analogue of 1,4-dihydro-4-oxoquinoline-3-carboxylic acid synthesized as a potential antiviral agent: Conformational studies in vacuum and in water. *Theochem*, **2006**, *778*, 97-103.

## 9 Anexos

### 9.1 Topologia utilizada para o AA

```
[ moleculetype ]
; Name nrexcl
DRG      3
```

```
[ atoms ]
```

```
;  nr      type  resnr  resid  atom  cgnr  charge  mass
   1         OM    1  DRG   OAU    1  -0.635  15.9994
   2          C    1  DRG   CAT    1   0.270  12.0110
   3         OM    1  DRG   OAV    1  -0.635  15.9994
   4        CH2    1  DRG   CAS    1   0.000  14.0270
   5        CH2    1  DRG   CAR    1   0.000  14.0270
   6        CH2    1  DRG   CAQ    2   0.000  14.0270
   7          C    1  DRG   CAP    2   0.000  13.0190
   8          C    1  DRG   CAO    2   0.000  13.0190
   9        CH2    1  DRG   CAN    2   0.000  14.0270
  10          C    1  DRG   CAM    2   0.000  13.0190
  11          C    1  DRG   CAL    2   0.000  13.0190
  12        CH2    1  DRG   CAK    3   0.000  14.0270
  13          C    1  DRG   CAJ    3   0.000  13.0190
  14          C    1  DRG   CAI    3   0.000  13.0190
  15        CH2    1  DRG   CAH    3   0.000  14.0270
  16          C    1  DRG   CAG    4   0.000  13.0190
  17          C    1  DRG   CAF    4   0.000  13.0190
  18        CH2    1  DRG   CAE    4   0.000  14.0270
  19        CH2    1  DRG   CAD    4   0.000  14.0270
  20        CH2    1  DRG   CAC    5   0.000  14.0270
  21        CH2    1  DRG   CAB    5   0.000  14.0270
  22        CH3    1  DRG   CAA    5   0.000  15.0350
```

## 9.2 Arquivos .mdp

Os arquivos “.mdp” definem parâmetros para as simulações de dinâmica molecular. A seguir são apresentados, à maneira de exemplos, arquivos “.mdp” empregados em uma das simulações de PGHS.

➤ Etapa de minimização de energia empregando o algoritmo Steepest Descent:

```
Input file
title           = Yo
cpp             = /lib/cpp
define          = -DFLEX_SPC
constraints     = none
integrator      = steep
emtol          = 1
emstep         = 0.001
nsteps         = 100000
nstcomm        = 1
nstxout        = 100
nstvout        = 100
nstfout        = 0
nstlog         = 100
nstenergy      = 100
nstlist        = 10
ns_type        = grid
coulombtype    = PME
rlist          = 0.9
rcoulomb       = 0.9
rvdw           = 0.9
pme_order      = 4
Tcoupl         = no
Pcoupl         = no
```

➤ Etapa 1 da termalização:

```

Input file
title           = Yo
cpp             = /lib/cpp
define         = -DPOSRES
constraints    = all-bonds
integrator     = md
tinit         = 0.0
dt            = 0.002 ; ps !
nsteps        = 500 ; total 1 ps.
nstcomm       = 1
nstxout       = 100
nstvout       = 100
nstfout       = 0
nstlog        = 100
nstenergy     = 100
nstlist       = 10
ns_type       = grid
coulombtype   = PME
rlist         = 0.9
rcoulomb      = 0.9
rvdw          = 0.9
fourierspacing = 0.12
optimize_fft  = yes
pme_order     = 4
ewald_rtol    = 1e-5
Tcoupl        = berendsen
tc-grps       = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
tau_t         = 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t         = 50 50 50 50 50 50 50 50 50 50 50 50 50
energygrps    = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
Pcoupl        = berendsen
Pcoupltype    = isotropic
tau_p         = 0.5
compressibility = 4.5e-5
ref_p         = 1.0
gen_vel       = yes
gen_temp      = 50.0
gen_seed      = 173529

```

➤ Etapa 2 da termalização:

```

Input file
title           = Yo
cpp             = /lib/cpp
define         = -DPOSRES
constraints    = all-bonds
integrator     = md
tinit         = 1.0
dt            = 0.002 ; ps !
nsteps        = 4500 ; total 9 ps.
nstcomm       = 1
nstxout       = 100
nstvout       = 100
nstfout       = 0
nstlog        = 100
nstenergy     = 100
nstlist       = 10
ns_type       = grid
coulombtype   = PME
rlist         = 0.9
rcoulomb      = 0.9
rvdw          = 0.9
fourierspacing = 0.12
optimize_fft  = yes
pme_order     = 4
ewald_rtol    = 1e-5
Tcoupl        = berendsen
tc-grps       = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
tau_t         = 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t         = 50 50 50 50 50 50 50 50 50 50 50 50 50
energygrps    = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
Pcoupl        = berendsen
Pcoupltype    = isotropic
tau_p         = 0.5
compressibility = 4.5e-5
ref_p         = 1.0
gen_vel       = no
gen_temp      = 50.0
gen_seed      = 173529

```

➤ Etapa 3 da termalização:

```

Input file
title           = Yo
cpp             = /lib/cpp
define          = -DPOSRES
constraints     = all-bonds
integrator      = md
tinit          = 10.0
dt             = 0.002 ; ps !
nsteps         = 2500 ; total 5 ps.
nstcomm        = 1
nstxout        = 100
nstvout        = 100
nstfout        = 0
nstlog         = 100
nstenergy      = 100
nstlist        = 10
ns_type        = grid
coulombtype    = PME
rlist          = 0.9
rcoulomb       = 0.9
rvdw           = 0.9
fourierspacing = 0.12
optimize_fft   = yes
pme_order      = 4
ewald_rtol     = 1e-5
Tcoupl         = berendsen
tc-grps        = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
tau_t          = 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t          = 100 100 100 100 100 100 100 100 100 100 100 100 100
energygrps     = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
Pcoupl         = berendsen
Pcoupltype     = isotropic
tau_p          = 0.5
compressibility = 4.5e-5
ref_p          = 1.0
gen_vel        = no
gen_temp       = 100.0
gen_seed       = 173529

```

➤ Etapa 4 da termalização:

```

Input file
title           = Yo
cpp             = /lib/cpp
define          = -DPOSRES
constraints     = all-bonds
integrator      = md
tinit          = 15.0
dt             = 0.002 ; ps !
nsteps         = 2500 ; total 5 ps.
nstcomm        = 1
nstxout        = 100
nstvout        = 100
nstfout        = 0
nstlog         = 100
nstenergy      = 100
nstlist        = 10
ns_type        = grid
coulombtype    = PME
rlist          = 0.9
rcoulomb       = 0.9
rvdw           = 0.9
fourierspacing = 0.12
optimize_fft   = yes
pme_order      = 4
ewald_rtol     = 1e-5
Tcoupl         = berendsen
tc-grps        = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
tau_t          = 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t          = 150 150 150 150 150 150 150 150 150 150 150 150 150
energygrps     = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
Pcoupl         = berendsen
Pcoupltype    = isotropic
tau_p          = 0.5
compressibility = 4.5e-5
ref_p          = 1.0
gen_vel        = no
gen_temp       = 150.0
gen_seed       = 173129

```

➤ Etapa 5 da termalização:

```

;      Input file
title          = Yo
cpp            = /lib/cpp
define         = -DPOSRES
constraints    = all-bonds
integrator     = md
tinit         = 20.0
dt            = 0.002 ; ps !
nsteps        = 2500 ; total 5 ps.
nstcomm       = 1
nstxout       = 100
nstvout       = 100
nstfout       = 0
nstlog        = 100
nstenergy     = 100
nstlist       = 10
ns_type       = grid
coulombtype   = PME
rlist         = 0.9
rcoulomb      = 0.9
rvdw          = 0.9
fourierspacing = 0.12
optimize_fft  = yes
pme_order     = 4
ewald_rtol    = 1e-5
Tcoupl        = berendsen
tc-grps       = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
tau_t         = 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t         = 200 200 200 200 200 200 200 200 200 200 200 200 200
energygrps    = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
Pcoupl        = berendsen
Pcoupltype    = isotropic
tau_p         = 0.5
compressibility = 4.5e-5
ref_p         = 1.0
gen_vel       = no
gen_temp      = 200.0
gen_seed      = 173529

```

➤ Etapa 6 da termalização:

```

Input file
title           = Yo
cpp             = /lib/cpp
define         = -DPOSRES
constraints    = all-bonds
integrator     = md
tinit         = 25.0
dt            = 0.002 ; ps !
nsteps        = 2500 ; total 5 ps.
nstcomm       = 1
nstxout       = 100
nstvout       = 100
nstfout       = 0
nstlog        = 100
nstenergy     = 100
nstlist       = 10
ns_type       = grid
coulombtype   = PME
rlist         = 0.9
rcoulomb      = 0.9
rvdw          = 0.9
fourierspacing = 0.12
optimize_fft  = yes
pme_order     = 4
ewald_rtol    = 1e-5
Tcoupl        = berendsen
tc-grps       = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
tau_t         = 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t         = 250 250 250 250 250 250 250 250 250 250 250 250 250
energygrps    = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
Pcoupl        = berendsen
Pcoupltype    = isotropic
tau_p         = 0.5
compressibility = 4.5e-5
ref_p         = 1.0
gen_vel       = no
gen_temp      = 250.0
gen_seed      = 173529

```

➤ Etapa 7 da termalização:

```

Input file
title           = Yo
cpp             = /lib/cpp
define         = -DPOSRES
constraints    = all-bonds
integrator     = md
tinit         = 30.0
dt            = 0.002 ; ps !
nsteps        = 2500 ; total 5 ps.
nstcomm       = 1
nstxout       = 100
nstvout       = 100
nstfout       = 0
nstlog        = 100
nstenergy     = 100
nstlist       = 10
ns_type       = grid
coulombtype   = PME
rlist         = 0.9
rcoulomb      = 0.9
rvdw          = 0.9
fourierspacing = 0.12
optimize_fft  = yes
pme_order     = 4
ewald_rtol    = 1e-5
Tcoupl        = berendsen
tc-grps       = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
tau_t         = 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t         = 300 300 300 300 300 300 300 300 300 300 300 300 300
energygrps    = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
Pcoupl        = berendsen
Pcoupltype    = isotropic
tau_p         = 0.5
compressibility = 4.5e-5
ref_p         = 1.0
gen_vel       = no
gen_temp      = 300.0
gen_seed      = 173529

```

➤ Exemplo de etapa de dinâmica sob temperatura constante:

```

Input file
title           = Yo
cpp             = /lib/cpp
define         = -DPOSRES
constraints    = all-bonds
integrator     = md
tinit         = 500.0
dt            = 0.002 ; ps !
nsteps        = 250000 ; total 500-1000 ps.
nstcomm       = 1
nstxout       = 250
nstvout       = 1000
nstfout       = 0
nstlog        = 100
nstenergy     = 100
nstlist       = 10
ns_type       = grid
coulombtype   = PME
rlist         = 0.9
rcoulomb      = 0.9
rvdw          = 0.9
fourierspacing = 0.12
optimize_fft  = yes
pme_order     = 4
ewald_rtol    = 1e-5
Tcoupl        = berendsen
tc-grps       = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
tau_t         = 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
ref_t         = 310 310 310 310 310 310 310 310 310 310 310 310 310
energygrps    = Protein SOL NA+ NAG1 NAG2 BMA1 MAN9 MANZ MAN6 MAN5
MAN0 HEME DRG
Pcoupl        = berendsen
Pcoupltype    = isotropic
tau_p         = 0.5
compressibility = 4.5e-5
ref_p         = 1.0
gen_vel       = no
gen_temp      = 310.0
gen_seed      = 173529

```

## Curriculum Vitae

Os itens II, III e IV restringem-se a atividades realizadas durante o período abrangido pela presente dissertação

### I. Formação acadêmica:

Graduação em Farmácia pela Universidade Federal do Rio Grande do Sul, de 2004/1 a 2008/1. *Glicosilação de proteínas*. Orientador: Prof. Hugo Verli.

### II. Trabalhos científicos apresentados em congressos

#### a. Nacionais:

**Sachett, L. G.;** Brizolara, T.; Verli, H. **2009**. Putative model for glycosylated prostaglandin endoperoxide synthases. *Programas e Resumos da XXXVIII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, N-32. XXXVIII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 16 a 19 de maio de 2009, Águas de Lindóia, SP, Brasil.

**Sachett, L. G.;** Verli, H. **2010**. Structural and conformational characterization of COXs. *Programas e Resumos da XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, N-22. XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 18 a 21 de maio de 2010, Foz do Iguaçu, PR, Brasil.

### III. Publicações em periódicos especializados:

#### a. Internacionais:

Fernandes, C. L.; **Sachett, L. G.;** Pol-Fachin, L.; Verli, H. **2010** GROMOS96 43a1 performance in predicting oligosaccharides conformational ensemble within glycoproteins *Carbohydr. Res.*, 345, 663-671.

### IV. Bolsa recebida:

Bolsista CAPES do programa de mestrado em biologia molecular e celular pelo Centro de Biotecnologia da UFRGS a partir de março de 2009.