

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

**DESCRIÇÃO CONFORMACIONAL DE CARBOIDRATOS E  
GLICOPROTEÍNAS: VALIDAÇÃO DE PROTOCOLO BASEADO EM  
DINÂMICA MOLECULAR E IMPLICAÇÕES FUNCIONAIS**

Laércio Pol Fachin

Porto Alegre – Brasil  
Outubro de 2009

**DESCRIÇÃO CONFORMACIONAL DE CARBOIDRATOS E GLICOPROTEÍNAS:  
VALIDAÇÃO DE PROTOCOLO BASEADO EM DINÂMICA MOLECULAR E  
IMPLICAÇÕES FUNCIONAIS**

*Laércio Pol Fachin*

Dissertação de mestrado elaborada no Laboratório 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

Outubro de 2009

**DESCRIÇÃO CONFORMACIONAL DE CARBOIDRATOS E GLICOPROTEÍNAS:  
VALIDAÇÃO DE PROTOCOLO BASEADO EM DINÂMICA MOLECULAR E  
IMPLICAÇÕES FUNCIONAIS**

*Laércio Pol Fachin*

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)

---

Arthur Germano Fett Neto (Centro de Biotecnologia - UFRGS)

---

Hubert Karl Stassen (Instituto de Química - UFRGS)

---

Yraima Moura Lopes Cordeiro (Faculdade de Farmácia - UFRJ)

---

Alexandre José Macedo (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**

POL-FACHIN, Laércio.

**Descrição conformacional de carboidratos e glicoproteínas: validação de protocolo baseado em dinâmica molecular e implicações funcionais**

Rio Grande do Sul, UFRGS, Centro de Biotecnologia, 2009.

Dissertação: Mestre em Ciências (Biologia Celular e Molecular)

1. Dinâmica Molecular

2. Glicoproteínas

3. Polissacarídeos

4. Ligação Glicosídica

I. Hugo Verli

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

III. Títulos

## AGRADECIMENTOS

Meus sinceros agradecimentos ao Prof. Hugo Verli, por ter me recebido junto ao Grupo de Bioinformática Estrutural e, desde então, ter se dedicado a me conduzir, por meio de ensinamentos, conselhos e, principalmente, pelo exemplo, a uma admirável conduta profissional e científica.

Aos professores Hubert Karl Stassen e Evelyn Koeche Schroeder, membros da comissão de acompanhamento, pela constante prestatividade e auxílio.

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.

Aos verdadeiros amigos que, mesmo em épocas de tempestade, continuam caminhando lado a lado ao longo da jornada.

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

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

A minha irmã e, principalmente, aos meus pais, pelo inabalável apoio, suporte e incentivo oferecidos durante a realização deste trabalho.

*“O segredo do verdadeiro Mestre é saber quando lutar, assim como o segredo do artista é saber quando representar. O conhecimento de assuntos e métodos técnicos é fundamental, mas não suficiente para garantir o sucesso, em qualquer arte ou ciência: a percepção direta do momento é essencial para a execução do que chamamos de um golpe de Mestre.”*

Adaptado de Sun Bin, discípulo direto de Wang Li.

Dedico esta dissertação a meus pais, Luiz Fachin e Maria Fátima Pol Fachin, e a todos aqueles que contribuíram, de alguma forma, para que o presente trabalho fosse realizado.

## SUMÁRIO

LISTA DE ABREVIATURAS.....	xi
RESUMO .....	xii
ABSTRACT .....	xiii
ÍNDICE DE FIGURAS .....	xiv
ÍNDICE DE TABELAS .....	xvii
1 Introdução .....	1
1.1 Glicosilação de proteínas .....	1
1.2 Glicoproteínas em processos biológicos .....	6
1.2.1 Fator VII da coagulação .....	7
1.2.2 CD2 humana .....	8
1.2.3 CD59 humana .....	8
1.2.4 hCG.....	9
1.3 Carboidratos em processos biológicos.....	11
1.4 Estudos da estrutura e conformação de glicanas e polissacarídeos.....	13
1.5 Caracterização conformacional de biomoléculas através de dinâmica molecular (DM).....	15
2 Objetivos .....	20
3 Metodologia.....	21
3.1 Programas utilizados.....	21
3.2 Cálculos utilizando métodos ab initio .....	21
3.3 Construção de mapas de contorno por mecânica molecular .....	22
3.4 Simulações de DM .....	23
3.4.1 Protocolo de simulação .....	23
3.4.2 Construção de topologias.....	25

3.4.3	Validação das simulações de DM .....	25
4	Resultados .....	27
4.1	Preâmbulo.....	27
4.2	Trabalho I.....	28
4.3	Trabalho II.....	40
4.4	Trabalho III.....	51
5	Discussão Geral.....	61
5.1	Estrutura e conformação de $\alpha$ - e $\beta$ -galactanas .....	61
5.1.1	Descrição da geometria das ligações glicosídicas .....	61
5.2	Reprodução conformacional de glicoproteínas .....	65
5.3	Previsão conformacional de oligossacarídeos .....	67
5.3.1	Predição da conformação de glicanas a partir de seus dissacarídeos constituintes .....	71
5.4	Implicações para a caracterização conformacional de glicoproteínas e polissacarídeos .....	73
6	Conclusões .....	74
7	Perspectivas.....	75
8	Referências Bibliográficas.....	76
9	Anexos .....	98
9.1	Trabalho I.....	99
9.2	Trabalho II.....	100
9.3	Cargas atômicas de Löwdin .....	101
9.4	Topologias de monossacarídeos para GROMOS96 43a1 .....	104
	Curriculum Vitae.....	116

## LISTA DE ABREVIATURAS

Ara – arabinose  
DM – dinâmica molecular  
EGF – fator de crescimento epidermal  
Fuc – fucose  
FucNAc – N-acetil-fucosamina  
GAG – glicosaminoglicano  
Gal – galactose  
GalNAc – N-acetil-galactosamina  
Glc – glicose  
GlcA – ácido glicurônico  
GlcN – glicosamina  
GlcNAc – N-acetil-glicosamina  
GPI – glicofosfatidilinositol  
hCG – gonadotrofina coriônica humana  
Hyl – hidroxilisina  
Hyp – hidroxiprolina  
IdoA – ácido idurônico  
Man – manose  
MD – molecular dynamics  
MM – mecânica molecular  
NeuAc – ácido siálico  
NOE – *nuclear Overhauser effect*  
PDB – Protein Data Bank  
RE – retículo endoplasmático  
Rha – ramnose  
RMN – ressonância magnética nuclear  
Xyl – xilose.

## RESUMO

O desenvolvimento de técnicas complementares a métodos experimentais clássicos, tais como a cristalografia de raios-X e a ressonância magnética nuclear (RMN), capazes de descrever e prever a diversidade estrutural e conformacional de carboidratos e glicoconjugados, consiste em um potencial impacto no entendimento dos processos biológicos em que estas moléculas estão envolvidas. Neste sentido, o presente trabalho tem por objetivo: 1) a validação de um conjunto de parâmetros, desenvolvidos no nosso grupo de pesquisas, na reprodução de dados de RMN descrevendo a conformação de uma série de glicoproteínas em solução aquosa, e 2) a avaliação do emprego de geometrias de dissacarídeos em solução como modelo para a construção de glicanas complexas, na ausência de dados experimentais. Através do emprego de técnicas de modelagem molecular como cálculos *ab initio*, construção de mapas de contorno para dissacarídeos e simulações de dinâmica molecular (DM) em escalas de tempo de até 0.1  $\mu$ s, os resultados obtidos confirmam a adequação dos parâmetros gerados na descrição da conformação de carboidratos e glicoproteínas. Adicionalmente, sugerem que o emprego de populações de confôrmeros de dissacarídeos em solução, obtidos através de simulações de DM, constitui-se em uma estratégia promissora na obtenção de modelos de glicanas condizentes com condições biológicas. Esperamos, a partir destas observações, que os resultados obtidos na presente dissertação possam contribuir no crescimento do número de trabalhos de DM abordando carboidratos e glicoproteínas biológicas e, assim, no crescimento do entendimento de processos biológicos nos quais estas biomoléculas estejam envolvidas.

## ABSTRACT

The development of novel techniques, capable of describing and predicting the structural and conformational diversity of carbohydrates and glycoconjugates, in complement to experimental methods as X-ray crystallography and nuclear magnetic resonance (NMR), might bring a potential impact in the comprehension of the biological processes in which such molecules are involved. In this context, this work aims to: 1) validate a set of parameters, developed by our research group, to reproduce NMR data describing the conformation of a series of glycoproteins in aqueous solutions; and 2) evaluate the use of solution geometries of disaccharides to assemble complex glycans, in the absence of experimental data. Employing molecular modelling techniques as *ab initio* calculations, construction of energy contour plots for disaccharides and molecular dynamics (MD) simulations with time scales up to 0.1  $\mu$ s, the obtained results confirm the adequacy of the studied parameters to describe the conformation of carbohydrates and glycoproteins. Additionally, such data suggest that the use of solution conformations of disaccharides, as obtained through MD simulations, consists in a promising strategy to obtain glycans representation in accordance to biological conditions. We expect, with such observations, that the obtained results contribute to increase the number of works employing MD simulations of biological carbohydrates and glycoproteins and, thus, to raise the comprehension of the biological processes in which such molecules are involved.

## ÍNDICE DE FIGURAS

- Figura 1: Representação esquemática das ligações entre aminoácidos e monossacarídeos na (A) fosfoglicosilação e na (B) C-glicosilação. .... 1
- Figura 2: Representação diagramática de alguns tipos de N- e O-ligações glicosídicas identificados até então. Adaptado de Spiro, 2002. .... 3
- Figura 3: Processo de biosíntese de um oligossacarídeo N-ligado do tipo complexo. O processo tem início com (A) a adição de monossacarídeos a uma molécula de dolicol fosfato ancorada à face citosólica da membrana do RE. Quando sete resíduos (duas GlcNAc e cinco Man) estão ligados, (B) este heptassacarídeo é revertido ao lúmen do RE e (C) mais sete monossacarídeos são adicionados. Em seguida, (D) esse oligossacarídeo é transferido para o polipeptídeo nascente e (E-F) pode sofrer ciclos de adição e remoção de resíduos de glicose. Quando a proteína (G) estabelece seu correto enovelamento e (H) chega ao complexo de Golgi, (I) resíduos de Man são removidos da estrutura glicosídica. Após, (J-K-L) tem início a adição de outros resíduos, tais como GlcNAc, Gal e NeuAc. As formas azuis representam GlcNAc; as rosas, Man; as vermelhas, Glc; as verdes, Gal; as cinzas, NeuAc; e as brancas, Fuc. Adaptado de Helenius & Aebi, 2001. .... 4
- Figura 5: Estruturas da proteína CD59 obtidas a partir de (A) cristalografia de raios-X e (B) RMN, mostrando suas cinco pontes dissulfeto (linhas amarelas), e seu enovelamento, rico em folhas  $\beta$  com uma  $\alpha$ -hélice lateral. .... 9
- Figura 6: Representação das estruturas (A) cristalográfica do hormônio hCG (PDB ID 1HCN), com as subunidades  $\alpha$  (azul) e  $\beta$  (cinza) ligadas, formando a molécula

completa, e da provável estrutura da (B) subunidade  $\alpha$  isolada (vermelho), com reduzido conteúdo de estrutura secundária..... 10

Figura 7: Estrutura de galactanas obtidas de diferentes espécies de ouriço-do-mar:

(A) *G. crenularis*, onde a unidade de repetição consiste em um dissacarídeo contendo resíduos de  $\beta$ -D-galactose unidos através de ligações glicosídicas  $\beta 1 \rightarrow 3$  e sulfatados na posição C-2 intercaladamente; e (B) *E. lucunter*, em que  $\alpha$ -L-galactoses sulfatadas na posição C-2 estão unidas através de ligações glicosídicas  $\alpha 1 \rightarrow 3$ . ..... 12

Figura 8: Representação do pentassacarídeo mínimo necessário para que a heparina tenha atividade potencializadora sobre AT. Nessa estrutura, o oligossacarídeo é composto por resíduos de IdoA 2-sulfatado, glicosamina 2,6 dissulfatada e ácido glicurônico não sulfatado..... 13

Figura 9: Variantes conformacionais e configuracionais associados a monossacarídeos: conformação do anel piranosídico (A e B), estados levógiro (C) ou dextrógiro (D) e estado anomérico  $\alpha$  ou  $\beta$  (D e E). Nessa representação, a glicose é usada como modelo..... 14

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 (interações eletrostáticas/Coulômbicas e interações de van der Waals, também identificadas como potencial de Lennard-Jones). Adaptado de Verli, 2004..... 17

Figura 11: 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)..... 24

Figura 12: Imagens de espécimes de (A) <i>Echinometra lucunter</i> e (B) <i>Glyptocidaris crenularis</i> .....	62
Figura 13: Mapas de contorno para as ligações glicosídicas de $\alpha$ - e $\beta$ -galactanas, sulfatadas e não-sulfatadas, sobrepostos aos perfis conformacionais observados nas simulações de DM em solução (pontos em cinza). Os níveis de energia são mostrados a cada 10 kJ.mol <sup>-1</sup> , de 10 a 50 kJ.mol <sup>-1</sup> .....	64
Figura 14: Distâncias entre hidrogênios vicinais entre monossacarídeos, observados através de técnicas de RMN e reproduzidos pelas simulações de DM em solução aquosa para os dissacarídeos de $\beta$ -galactana.....	65
Figura 15: Cargas atômicas obtidas para as ligações glicosídicas entre aminoácidos e monossacarídeos.....	66
Figura 16: Flutuação do ângulo $\omega$ nos dissacarídeos (A) Fuc-( $\alpha$ 1→6)-GlcNAc e (B) Man-( $\alpha$ 1→6)-Man.....	71

## ÍNDICE DE TABELAS

Tabela 1: Efeitos da glicosilação sobre proteínas.....	6
Tabela 2: Estruturas tridimensionais de glicoproteínas e polissacarídeos determinados por RMN.....	15
Tabela 3: Estudos envolvendo DM de glicoproteínas .....	19
Tabela 4: Dissacarídeos estudados .....	22
Tabela 5: Proteínas e glicoproteínas estudadas .....	24
Tabela 6: Definição dos diedros impróprios utilizados para definir a conformação dos monossacarídeos contidos nos sistemas simulados.....	26
Tabela 7: Definições de RMN e de cristalografia para os diedros componentes de ligações glicosídicas de dissacarídeos. ....	69
Tabela 8: Comparação da geometria dos ângulos $\Phi$ e $\Psi$ , obtidas por DM e RMN, de ligações glicosídicas que compõe a parte sacarídica de glicoproteínas .....	72

# 1 Introdução

## 1.1 Glicosilação de proteínas

A glicosilação é uma das modificações de proteínas mais freqüentes que, coe/ou pós-traducionalmente, ocorre em praticamente todos os organismos vivos, desde eubactérias até eucariotos, sendo controlada por fatores que mudam significativamente entre tipos celulares e espécies (Spiro, 2002). Adicionalmente, esse processo é a alteração protéica mais variável (Medvedová & Farkas, 2004), podendo abranger uma grande diversidade de aminoácidos e monossacarídeos. Até o momento, encontram-se descritos pelo menos cinco tipos de ligações entre carboidratos e proteínas, envolvendo treze monossacarídeos e oito aminoácidos diferentes (Spiro, 2002): C-glicosilação, fosfoglicosilação, gliptação, O-glicosilação e N-glicosilação.

Tendo sido descritos apenas recentemente, e somente em eucariotos, pouca informação encontra-se disponível para dois dos tipos acima citados: a fosfoglicosilação (Mehta *et al.*, 1996; Haynes, 1998), em que oligossacarídeos estão ligados a resíduos de serina ou treonina via ligações fosfodiéster (Figura 1A); e a C-manosilação ou C-glicosilação (Hofsteenge *et al.*, 1994; de Beer *et al.*, 1995; Krieg *et al.*, 1997; Doucey *et al.*, 1998), na qual um resíduo de  $\alpha$ -manose é adicionado ao grupamento indol de um triptofano (Figura 1B) presente na primeira posição da seqüência consenso Trp-X1-X2-Trp/Cys, sendo que essa ligação é favorecida quando X1 é um aminoácido pequeno ou polar (como serina, alanina, glicina ou treonina) e desfavorecida quando X1 é fenilalanina ou leucina (Krieg *et al.*, 1998; Julenius, 2007).

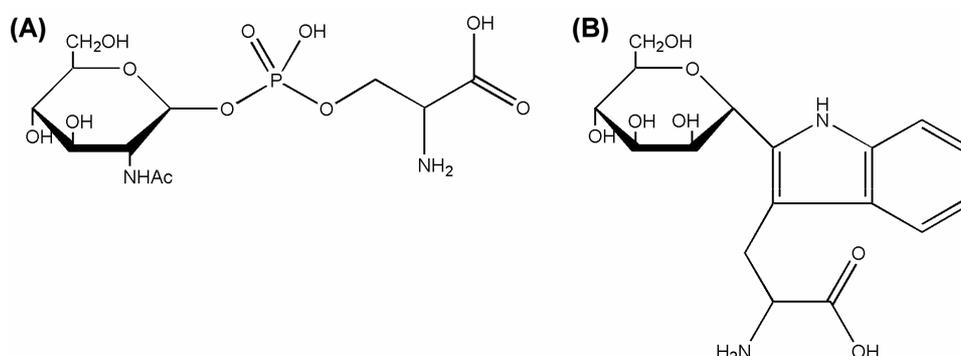


Figura 1: Representação esquemática das ligações entre aminoácidos e monossacarídeos na (A) fosfoglicosilação e na (B) C-glicosilação.

A gliptação, identificada em eucariotos e arqueas (Spiro, 2002), envolve a adição de GPI à região C-terminal de polipeptídeos, que é realizada a fim de ancorar algumas glicoproteínas e proteoglicanas a superfícies celulares (Low, 1989; Paulick & Bertozzi, 2008). Tal modificação tem início quando da clivagem de um peptídeo sinal da região C-terminal da proteína, na qual um grupamento fosfoetanolamina se liga ao grupamento carboxi-terminal protéico e ao esqueleto glicídico [ $\alpha$ -D-Man-(1→2)- $\alpha$ -D-Man-(1→6)- $\alpha$ -D-Man-(1→4)- $\alpha$ -D-GlcN-(1→6)-*myo*-inositol] do GPI, que por sua vez está conectado covalentemente a um fosfolípido. Uma série de modificações (especialmente, ramificações) no esqueleto glicídico do GPI, no grupamento fosfoetanolamina e nos possíveis fosfolípidos envolvidos já foram identificadas (Paulick & Bertozzi, 2008). Nesse sentido, diversas relações entre a estrutura e a atividade associadas a este grupamento tem sido propostas, no entanto a única função biológica realmente confirmada para o GPI é a de conferir às proteínas um ancoramento estável a membranas (Low, 1989; Low & Saltiel, 1988).

A N-glicosilação e a O-glicosilação são os tipos de glicosilação mais comumente encontrados na natureza, identificadas em eucariotos, arqueas e eubactérias (Spiro, 2002; Schmidt *et al.*, 2003; Abu-Qarn *et al.*, 2008). Até hoje, vários tipos de N- e O-ligações glicosídicas foram descritas (Figura 2), embora o estado anomérico de algumas dessas conexões e determinadas seqüências consenso de glicosilação, que indicam o ponto para a adição dos resíduos de carboidratos na cadeia polipeptídica, não estejam identificadas (Spiro, 2002). As conformações dessas ligações glicosídicas e seus efeitos sobre a estrutura protéica consistem em uma ampla área de conhecimento ainda a ser explorada.

A O-ligação glicosídica é a mais diversa e versátil, considerando-se a elevada quantidade de aminoácidos e monossacarídeos envolvidos e que, até o momento, poucas seqüências consenso conservadas foram identificadas (Spiro, 2002). No entanto, o desenvolvimento de bases de dados e algoritmos de predição (Gupta *et al.*, 1999; Li *et al.*, 2006) possibilitaram a identificação de algumas propriedades das regiões de O-glicosilação, tais como a preferência por prolina e a aversão por aminoácidos aromáticos nas proximidades da ligação glicosídica (Thanka Christlet & Veluraja, 2001). Por outro lado, a N-glicosilação possui uma seqüência consenso de reconhecimento determinada, que inclui a presença do tripeptídeo Asn-X-Ser/Thr na superfície da proteína (Hunt & Dayhoff, 1970; Marshall, 1972), onde X pode ser qualquer um dos 20 aminoácidos naturais, exceto prolina (Gavel & von Heijne, 1990;

Marshall, 1974). Em contrapartida à O-glicosilação, enquanto resíduos de prolina impedem a glicosilação (Gavel & von Heijne, 1990; Marshall, 1974), a presença de resíduos aromáticos favorecem a N-glicosilação (Thanka Christlet & Veluraja, 2001).

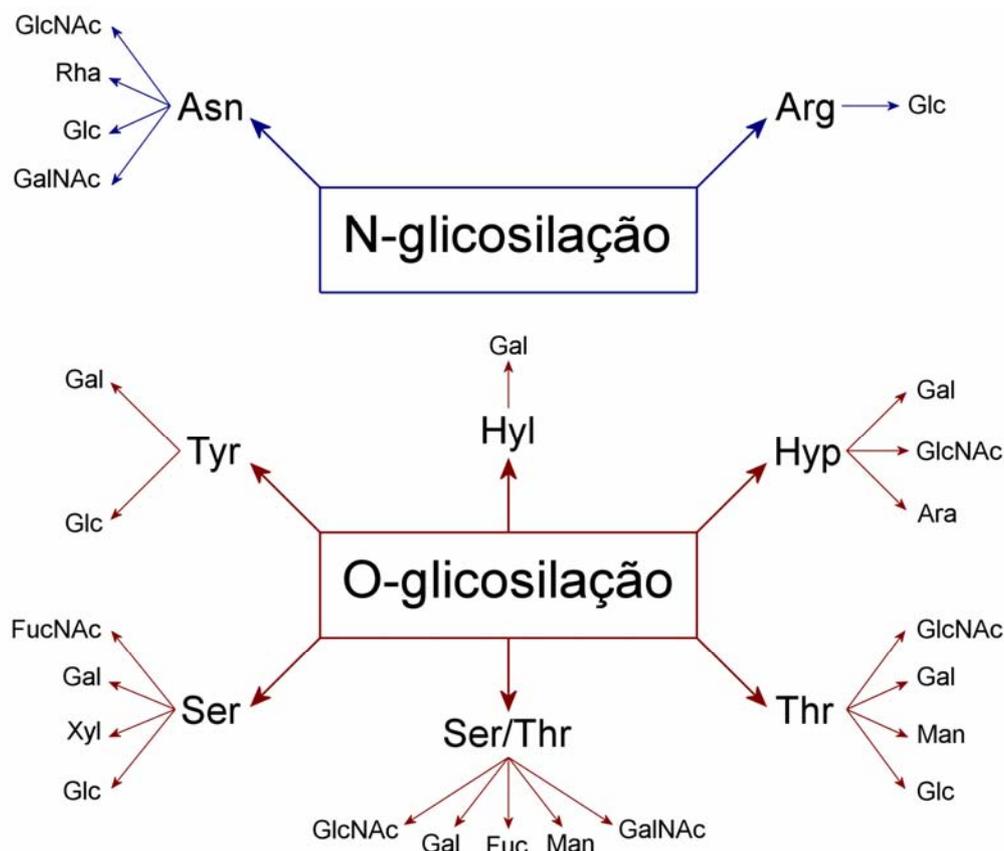


Figura 2: Representação diagramática de alguns tipos de N- e O-ligações glicosídicas identificados até então. Adaptado de Spiro, 2002.

A síntese de O-ligações glicosídicas é atribuída a várias enzimas diferentes, mesmo em eucariotos ou em eubactérias separadamente, de maneira que a ligação de carboidratos a proteínas ocorreria através da transferência pós-traducional de monossacarídeos isolados – GalNAc, Man, Xyl, Fuc, a aminoácidos que possuam um grupamento hidroxila livre – Thr e Ser, principalmente – em polipeptídeos já enovelados (Imperiali & Hendrickson, 1995).

Por outro lado, a síntese e modificação de oligossacarídeos durante a N-glicosilação são amplamente estudadas e compreendidas, nas quais, inclusive, determinadas enzimas estão relacionadas a etapas específicas do processo (Helenius & Aebi, 2001). Essencialmente, essa cascata de eventos começa na face citosólica da membrana do RE, onde são ligados alguns dos monossacarídeos

constituintes de um tetradecassacarídeo inicial (Figura 3A) que, posteriormente, já no lúmen dessa organela, é transferido para a cadeia polipeptídica nascente (Figura 3D). Em seguida, após o completo enovelamento da glicoproteína, inicia-se o processamento desse oligossacarídeo, no complexo de Golgi (Figura 3H), onde tem origem a imensa diversidade observada nos glicoconjugados que atingem as superfícies celulares (Helenius & Aebi, 2001).

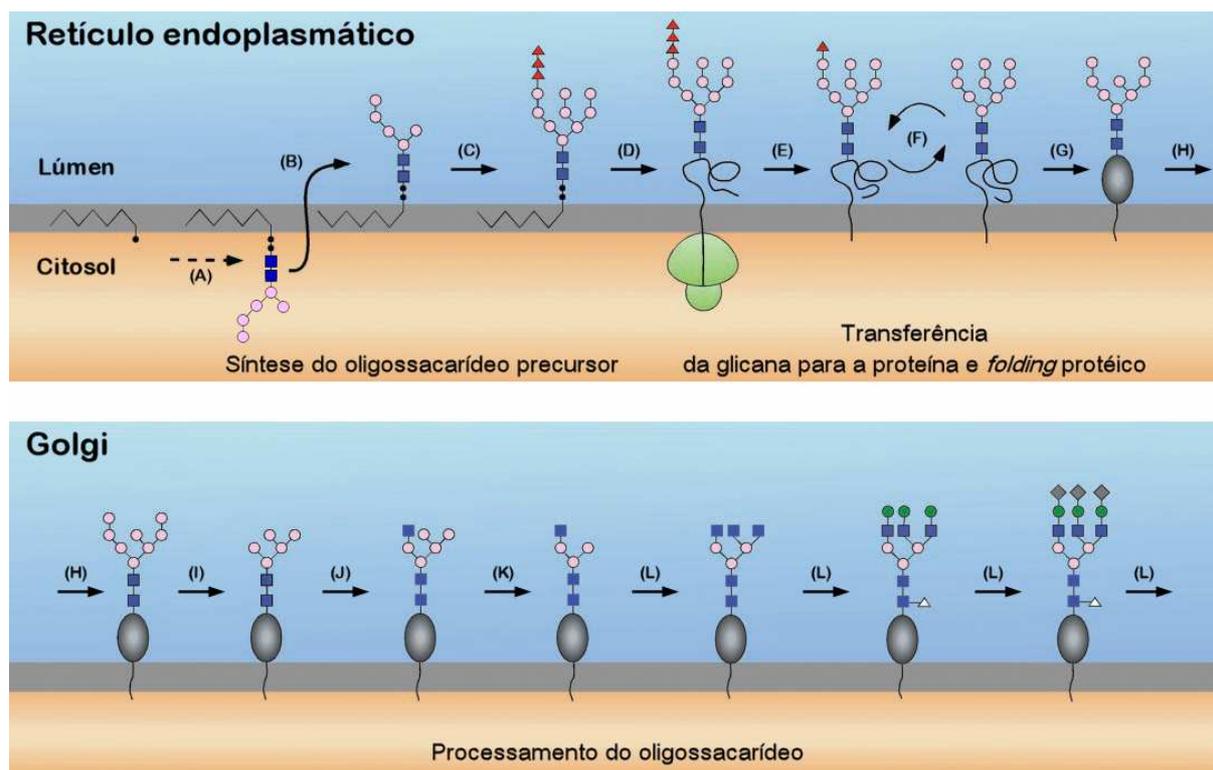


Figura 3: Processo de biosíntese de um oligossacarídeo N-ligado do tipo complexo. O processo tem início com (A) a adição de monossacarídeos a uma molécula de dolicol fosfato ancorada à face citosólica da membrana do RE. Quando sete resíduos (duas GlcNAc e cinco Man) estão ligados, (B) este heptassacarídeo é revertido ao lúmen do RE e (C) mais sete monossacarídeos são adicionados. Em seguida, (D) esse oligossacarídeo é transferido para o polipeptídeo nascente e (E-F) pode sofrer ciclos de adição e remoção de resíduos de glicose. Quando a proteína (G) estabelece seu correto enovelamento e (H) chega ao complexo de Golgi, (I) resíduos de Man são removidos da estrutura glicosídica. Após, (J-K-L) tem início a adição de outros resíduos, tais como GlcNAc, Gal e NeuAc. As formas azuis representam GlcNAc; as rosas, Man; as vermelhas, Glc; as verdes, Gal; as cinzas, NeuAc; e as brancas, Fuc. Adaptado de Helenius & Aebi, 2001.

No caso da N-glicosilação, o oligossacarídeo precursor pode ser convertido em diferentes tipos de glicanas (Figura 4), dependendo da proteína a que esse carboidrato está ligado e de fatores teciduais ou celulares (Spiro, 2002), podendo influenciar de maneira variada diversas propriedades de glicoproteínas. Nesse contexto, estudos prévios em modelos animais indicaram que os tipos desses oligossacarídeos podem mudar durante o desenvolvimento, da mesma forma que se identificou que tipos específicos dessas glicanas podem ser expressos em diferentes estágios da diferenciação (Cipollo *et al.*, 2005), possivelmente alterando os alvos moleculares dessas proteínas e/ou suas atividades (Fukuda, 1991; Saéz *et al.*, 2001). Modificações similares nos carboidratos presentes em proteínas de superfície celular estão associadas com várias patologias, tais como câncer (Couldrey & Green, 2000; Takahashi *et al.*, 2008) e um grupo de patologias derivadas de defeitos no processo de N-glicosilação, ao que se denomina síndrome de glicoproteínas com deficiência de carboidratos (Jaeken & Carchon, 1993).

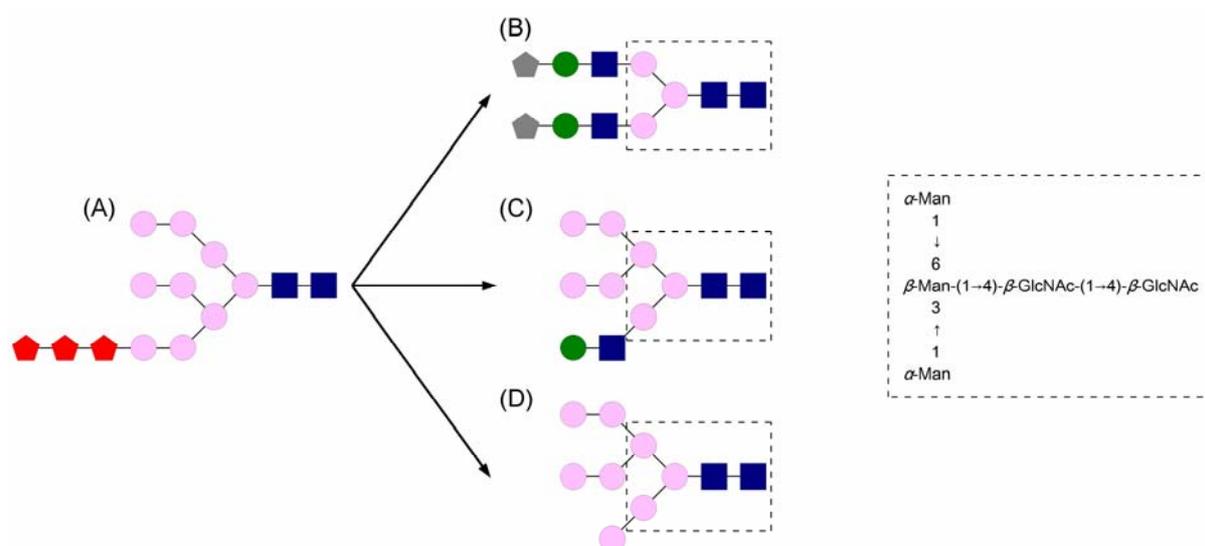


Figura 4: Representação esquemática do (A) tetradecassacarídeo precursor e dos três tipos principais de glicanas a que esse oligossacarídeo pode ser convertido através de processamento no RE e no complexo de Golgi: (B) tipo complexo, (C) tipo híbrido e (D) tipo *high-mannose*, também chamado de oligomanose. As formas azuis representam GlcNAc; as rosas, Man; as vermelhas, Glc; as verdes, Gal; e as cinzas, NeuAc. Em linhas tracejadas, é apresentada a estrutura do pentassacarídeo central dos oligossacarídeos N-ligados.

## 1.2 Glicoproteínas em processos biológicos

A forma mais comum na qual diversidade funcional é gerada para proteínas consiste na alteração da ordem e composição de seus aminoácidos (Mitra *et al.*, 2006). No entanto, mesmo com a manutenção dessa organização, um termo adicional de variabilidade pode ser adicionado através da glicosilação, de maneira a promover efeitos não somente sobre a função, mas sobre diversas outras propriedades protéicas (Tabela 1).

Tabela 1: Efeitos da glicosilação sobre proteínas.

Propriedades protéicas	Efeitos*
Físicas	Resistência à proteólise
	Resistência à desnaturação
	Alteração do pH ótimo para atividade
	Aumento de viscosidade da solução
	Diminuição do ponto de congelamento
Enovelamento	Aumento de solubilidade da proteína
	Prevenção de agregação protéica
	Promoção de interação com chaperonas
Atividades biológicas	Promoção de enovelamento em folhas $\beta$
	Mudança de reconhecimento proteína-proteína
	Mudança de reconhecimento proteína-carboidrato
	Alteração na taxa de transporte e secreção
	Alteração na oligomerização protéica

\* Dados adaptados de Varki, 1993 e Sears & Wong, 1998.

Dentre os possíveis efeitos de carboidratos sobre propriedades protéicas, merece destaque o papel de glicanas sobre a dinâmica conformacional de cadeias polipeptídicas (Wyss *et al.*, 1996). Tendo em vista a elevada hidrossolubilidade das estruturas sacarídicas, o número de glicanas ligadas a proteínas, mais do que a composição dessas estruturas, está associada a este efeito (Shental-Bechor & Levy, 2008). Adicionalmente, o aumento de estabilidade conformacional conferida por oligossacarídeos N-ligados a proteínas é atribuído a porções da unidade pentassacarídica (Figura 4) mais próxima da ligação à cadeia polipeptídica (Hanson

*et al.*, 2009), que é a única parte desses oligossacarídeos que não sofre modificações durante seu processamento (Helenius & Aebi, 2001). Da mesma forma, a origem desse processo foi relacionada à desestabilização dos estados não-enovelados da proteína, ao invés da estabilização dos estados protéicos enovelados (Shental-Bechor & Levy, 2008).

Nesse sentido, a glicosilação pode desempenhar papel chave sobre os eventos e funções biológicas em que glicoproteínas estejam envolvidas, tais como as de enzimas, hormônios, citosinas, fatores de crescimento e fatores de coagulação (Varki, 1993). Devido a essa ampla variedade de funções, as glicoproteínas podem estar envolvidas em uma série de processos biológicos, tais como adesão celular, reconhecimento intercelular, inflamação e resposta imune, da mesma forma que podem estar envolvidas com eventos mais específicos, tais como atividade anticongelamento (Varki, 1993; Dwek, 1996; Imperiali, 1997; Sears & Wong, 1998). A seguir, algumas glicoproteínas serão apresentadas como exemplos da influência da glicosilação sobre suas conformações e funções.

### **1.2.1 Fator VII da coagulação**

O fator VII da coagulação é um zimogênio que circula no plasma e é ativado através da ação catalítica de um outro fator da coagulação, a trombina, de maneira a formar o fator VII ativado, ou fVIIa (Hagen *et al.*, 1986). A atividade catalítica de sua forma ativa é muito pequena, tornando-se pronunciada somente após sua interação com o fator tecidual, uma proteína integral de membrana exposta como consequência de danos vasculares agudos (Nakagaki *et al.*, 1991), evento considerado como a primeira etapa da cascata de coagulação (Butenas e Mann, 1996).

A estrutura do fVIIa é formada por uma cadeia pesada, composta por 254 resíduos de aminoácidos, e uma cadeia leve, de 152 resíduos (Hagen *et al.*, 1986). Adicionalmente, há, ainda, quatro sítios de glicosilação: três em sua cadeia leve, sendo dois de O-glicosilação (Ser52 e Ser60) e outro de N-glicosilação (Asn145), e o último, na cadeia pesada, uma N-glicosilação no resíduo Asn170 (Hansson & Stenflo, 2005). Sua estrutura tridimensional já foi estudada através de cristalografia de raios-X, principalmente quando em complexo com o fator tecidual (Banner *et al.*, 1996). Contudo, pouco se sabe sobre sua conformação quando livre em solução, tampouco sobre a influência da glicosilação sobre sua estrutura. Nesse contexto,

técnicas de RMN foram utilizadas no estudo de uma parte dessa proteína, o primeiro de dois domínios EGF (Kao *et al.*, 1999), localizado na região N-terminal da proteína, no qual se observou que a presença de O-glicosilação no resíduo Ser60 não causa mudanças conformacionais no domínio, mas parece aumentar sua afinidade por íons  $\text{Ca}^{2+}$ , envolvidos na interação do fVIIa completo com o fator tecidual.

### 1.2.2 CD2 humana

A proteína CD2 humana, também conhecida como LFA-2, é uma glicoproteína transmembrana de aproximadamente 55 kDa, presente em células de defesa, tais como células T e células *natural killer*. Essa molécula é um importante mediador de adesão celular e transdução de sinal, função exercida por meio de sua ligação com seu receptor, CD58, também conhecido como LFA-3 (Moingeon *et al.*, 1989). Embora sua estrutura tridimensional não esteja totalmente elucidada, sua porção extracelular, referente a dois domínios ricos em folhas  $\beta$ , já foi estudada através de cristalografia de raios-X (PDB ID 1HNF) (Bodian *et al.*, 1994) e seu domínio de adesão, que se localiza na porção N-terminal da proteína, foi também avaliada através de técnicas de RMN (PDB ID 1GYA) (Wyss *et al.*, 1995).

A proteína CD2 possui um único sítio de N-glicosilação, localizado no domínio de adesão desta glicoproteína. Estudos prévios indicam que o oligossacarídeo adicionado à Asn65 dessa proteína é do tipo oligomanose (Wyss *et al.*, 1995), e é crucial no desempenho das funções de adesão da molécula (Recny *et al.*, 1992). No entanto, embora se conheça a região da proteína CD2 que interage com seu receptor alvo (Arulanandam *et al.*, 1993), localizada próxima a um conjunto de folhas  $\beta$  antiparalelas, a maneira como a glicosilação afeta as funções da proteína ainda não está compreendida.

### 1.2.3 CD59 humana

CD59 é uma glicoproteína de superfície celular que está associada: (1) à proteção de células plasmáticas e do endotélio vascular contra lise mediada pelo sistema complemento, impedindo a ligação do fator C9 com os complexos C5b-8 e C5b-9 (Farkas *et al.*, 2002); (2) atividade regulatória sobre células T e sobre a capacidade dessas células em produzir citocinas (Longhi *et al.*, 2005; Longhi *et al.*, 2006); (3) atividades relacionadas à reprodução e desenvolvimento humano (Kimberley *et al.*, 2007).

Essa proteína possui entre 18 e 20kDa, com um único sítio de N-glicosilação (Asn18) e vários sítios de O-glicosilação. Pertence à família de proteínas Ly-6, ricas em cisteínas e ancoradas à membrana através de GPI (Kimberley *et al.*, 2007). Nesse contexto, sua estrutura tridimensional foi descrita previamente através de técnicas de RMN (PDB ID 1CDQ) (Fletcher *et al.*, 1994) e cristalografia de raios-X (PDB ID 2J8B) (Leath *et al.*, 2007), demonstrando que CD59 é estabilizada por cinco pontes dissulfeto e possui um enovelamento com elevado conteúdo de folhas  $\beta$  (Figura 5). Adicionalmente, parte de seu oligossacarídeo N-ligado foi observado em estudos de RMN, na qual dissacarídeos (PDB ID 1CDS) ou trissacarídeos (PDB ID 1CDR) foram detectados em uma isoforma solúvel da proteína (Fletcher *et al.*, 1994), que é excretada via urina.

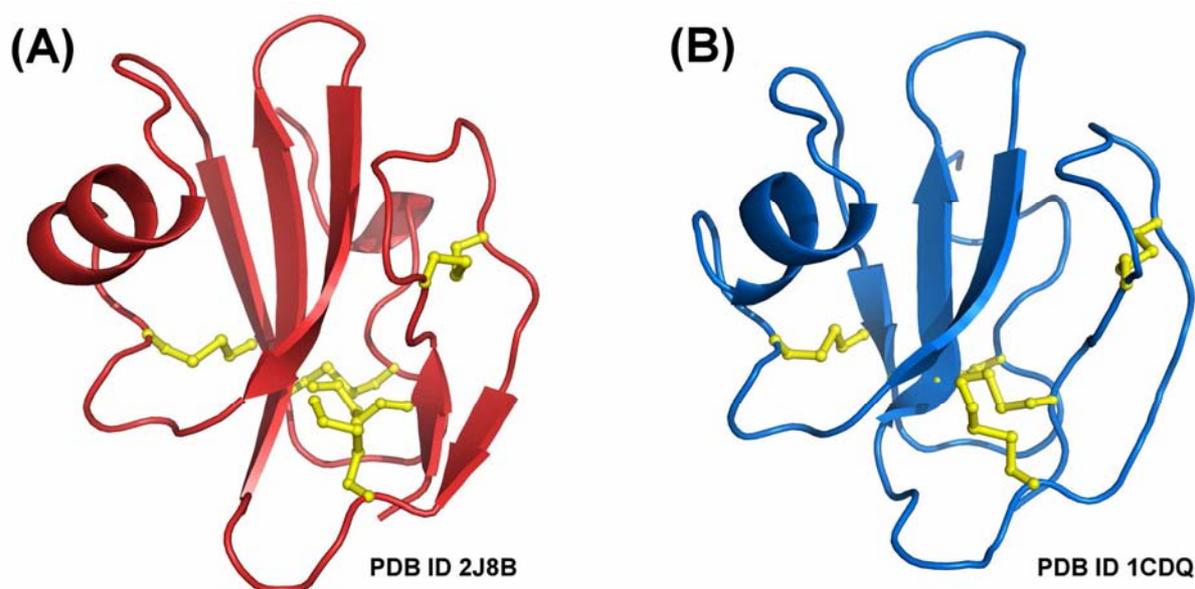


Figura 5: Estruturas da proteína CD59 obtidas a partir de (A) cristalografia de raios-X e (B) RMN, mostrando suas cinco pontes dissulfeto (linhas amarelas), e seu enovelamento, rico em folhas  $\beta$  com uma  $\alpha$ -hélice lateral.

#### 1.2.4 hCG

A gonadotrofina coriônica humana é um hormônio composto por duas subunidades glicoprotéicas, denominadas  $\alpha$ -hCG e  $\beta$ -hCG que, em mulheres grávidas, possui a função de estimular o corpo lúteo, presente nos ovários, a produzir progesterona (Talwar, 1979). Enquanto a subunidade  $\alpha$  possui 92 aminoácidos e é N-glicosilada em duas posições, Asn52 e Asn78, a subunidade  $\beta$  possui 145 aminoácidos, sendo N-glicosilada nas posições Asn13 e Asn30 e O-

glicosilada nas posições Ser121, Ser127, Ser132 e Ser138 (Kessler *et al.*, 1979a; Kessler *et al.*, 1979b). A interação dessas duas glicoproteínas forma um complexo rico em folhas  $\beta$  (Figura 6A), estabilizadas, principalmente, por interações hidrofóbicas, como determinado em estudos de cristalografia de raios X (PDB ID 1HCN e 1HRP) (Lapthorn *et al.*, 1994; Wu *et al.*, 1994).

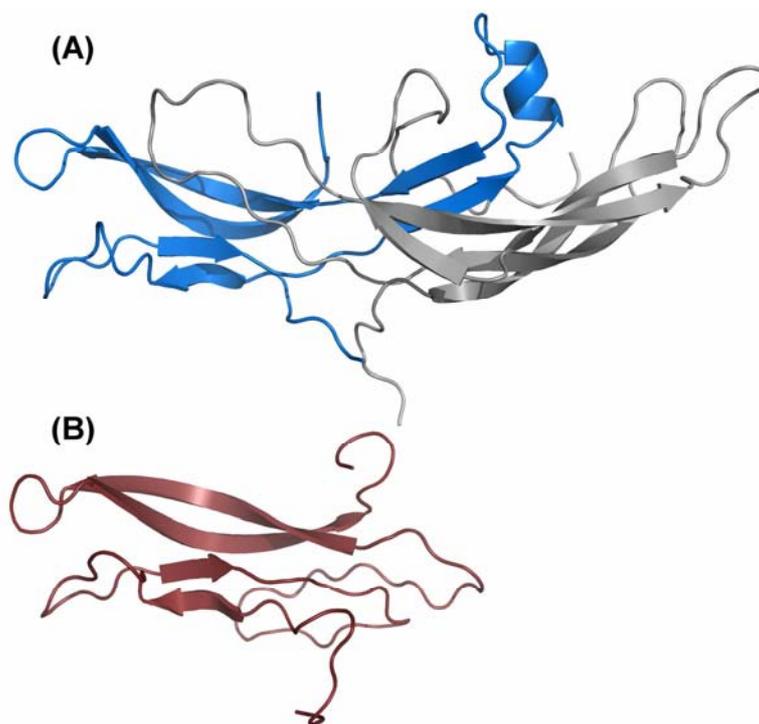


Figura 6: Representação das estruturas (A) cristalográfica do hormônio hCG (PDB ID 1HCN), com as subunidades  $\alpha$  (azul) e  $\beta$  (cinza) ligadas, formando a molécula completa, e da provável estrutura da (B) subunidade  $\alpha$  isolada (vermelho), com reduzido conteúdo de estrutura secundária.

Os hormônios glicoprotéicos heterodiméricos, que compreendem, além da molécula supracitada, os hormônios luteinizante, folículo-estimulante e tireo-estimulante, possuem como característica a presença de uma mesma cadeia  $\alpha$ , sendo diferenciados, portanto, por suas cadeias  $\beta$ , que são codificadas por uma série de genes altamente homólogos arranjados *in tandem*. Adicionalmente, a glicosilação dessas proteínas foi relacionada com a regulação da combinação dessas subunidades (Blithe *et al.*, 1995). Nesse contexto, também são atribuídas à subunidade  $\alpha$  isolada atividades independentes, tais como indução de produção de prolactina (Begeot *et al.*, 1984). Sua estrutura já foi determinada por técnicas de

RMN (PDB ID 1HD4) (Erbel *et al.*, 2000), indicando, juntamente a outros estudos (de Beer *et al.*, 1996), que no estado não complexado tal proteína possui um reduzido conteúdo de estrutura secundária em relação ao seu estado ligado às cadeias de  $\beta$ -hCG (Figura 6B).

### **1.3 Carboidratos em processos biológicos**

O papel de carboidratos sobre fenômenos biológicos não se restringe aos compostos mencionados anteriormente. Nesse sentido, embora grande parte dos eventos biológicos esteja envolvida com estruturas protéicas, estruturas sacarídicas vêm sendo identificadas como moléculas importantes em alguns desses processos (Varki, 1993; Dwek, 1996). Além da participação no metabolismo de glicídios, que gera grande parte da energia – na forma de trifosfato de adenosina, ou ATP – a ser utilizada pelo organismo, carboidratos podem desempenhar papéis importantes no metabolismo, sem envolvimento na geração de energia, tais como compor a parede celular de bactérias, fungos e plantas (Fincher, 2009; Latgé, 2007), auxiliar na reação acrossômica do esperma de espécies de ouriço-do-mar (Alves *et al.*, 1997; Alves *et al.*, 1998; Vilela-Silva *et al.*, 1999; Vilela-Silva *et al.*, 2002) e participar na mediação do reconhecimento intercelular (Misevic & Burger, 1990a; Misevic & Burger, 1990b).

Por exemplo, galactanas e fucanas são polissacarídeos sulfatados que estão entre os carboidratos mais abundantes de origem não-mamífera encontrados na natureza (Pereira *et al.*, 2002b; Melo *et al.*, 2004). São constituídas de unidades de repetição compostas, respectivamente, por galactose e fucose, que podem apresentar variação em seus estados anoméricos e isomeria (D- ou L-), tendo como principais fontes de obtenção as algas marinhas. Nesses organismos, estes polissacarídeos são componentes da parede celular (Lahaye, 2001), possuindo estruturas heterogêneas e complexas, tornando-os, portanto, compostos de difícil elucidação estrutural (Patankar *et al.*, 1993; Mulloy *et al.*, 1994). Em contrapartida, outra fonte de obtenção de galactanas e fucanas são os invertebrados marinhos, principalmente espécies de ouriço-do-mar, tais como *Glyptocidaris crenulatis* e *Echinometra lucunter*, em que esses polissacarídeos apresentam estruturas simples, lineares e regulares (Figura 7), ou seja, com unidades de repetição, padrões de sulfatação e posição de ligações glicosídicas bem definidas. Em ouriços-do-mar, tais polissacarídeos encontram-se relacionados com a reação acrossômica do esperma,

sendo, inclusive, indutores espécie-específicos desse evento (Alves *et al.*, 1997; Alves *et al.*, 1998; Vilela-Silva *et al.*, 1999; Vilela-Silva *et al.*, 2002). Adicionalmente, alguns destes polissacarídeos possuem promissora atividade anticoagulante quando testados em modelos animais (Pereira *et al.*, 1999).

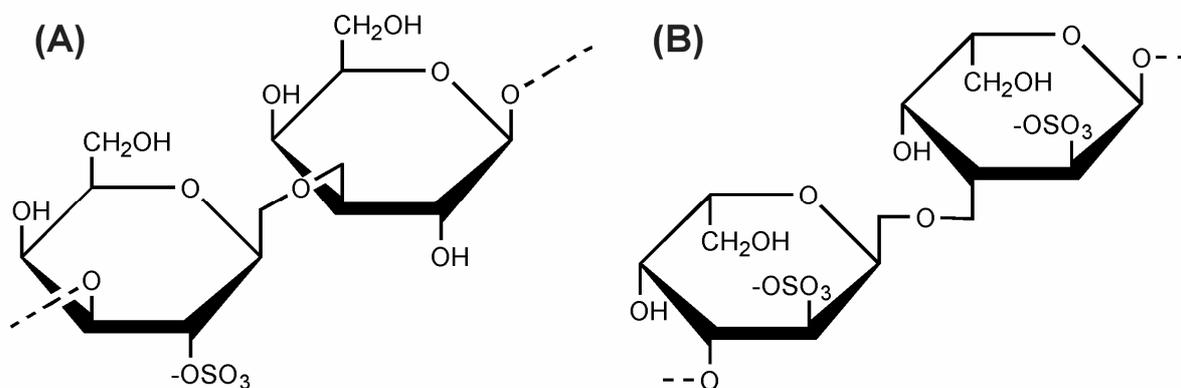


Figura 7: Estrutura de galactanas obtidas de diferentes espécies de ouriço-do-mar: (A) *G. crenularis*, onde a unidade de repetição consiste em um dissacarídeo contendo resíduos de β-D-galactose unidos através de ligações glicosídicas β1→3 e sulfatados na posição C-2 intercaladamente; e (B) *E. lucunter*, em que α-L-galactoses sulfatadas na posição C-2 estão unidas através de ligações glicosídicas α1→3.

Outro exemplo envolve os GAGs, uma família de polissacarídeos sulfatados lineares, compostos por unidades de repetição dissacarídicas que podem variar na composição básica de resíduos, em termos de ligação glicosídica e padrões de acetilação e sulfatação (Gama & Hsied-Wilson, 2005; Handel *et al.*, 2005). Tais moléculas são geralmente encontradas covalentemente ligadas a proteínas na superfície celular ou na matriz extracelular (Imberty *et al.*, 2007). Adicionalmente, podem estar relacionados com uma série de processos biológicos, tais como desenvolvimento neuronal, sinalização celular, facilitação de oligomerização, prevenção à proteólise e adesão celular (Handel *et al.*, 2005). Podem ser divididos em quatro tipos, classificados de acordo com as suas unidades de repetição: (1) condroitina/dermatan sulfato; (2) queratan sulfato; (3) ácido hialurônico; e (4) heparina/heparan sulfato. Nesse contexto, a heparina, um dos componentes desse último grupo, atraiu grande atenção pelo seu potencial terapêutico, sendo composta por resíduos de ácido idurônico (IdoA), ácido glicurônico (GlcA) e glicosamina (GlcN)

(Silva & Dietrich, 1975), que podem apresentar diferentes padrões de sulfatação dependendo do tecido e do organismo de origem deste carboidrato (Nader *et al.*, 2001). Tal polissacarídeo vem sendo utilizado clinicamente há mais de 40 anos como agente anticoagulante e antitrombótico (Jacques, 1979), atuando através da potenciação da atividade da antitrombina sobre proteases da cascata de coagulação (Choay *et al.*, 1983; Danielsson *et al.*, 1986) (Figura 8).

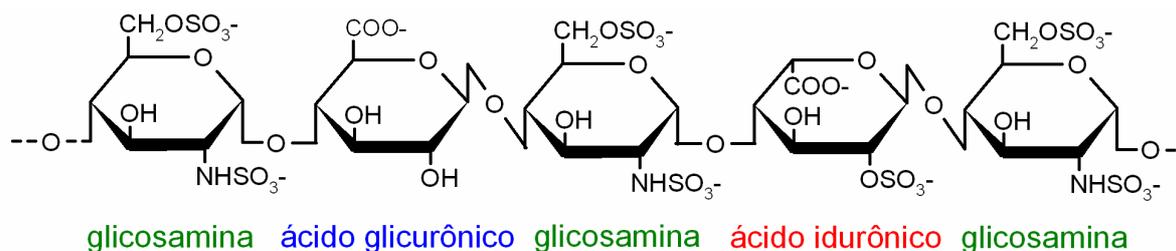


Figura 8: Representação do pentassacarídeo mínimo necessário para que a heparina tenha atividade potencializadora sobre AT. Nessa estrutura, o oligossacarídeo é composto por resíduos de IdoA 2-sulfatado, glicosamina 2,6 dissulfatada e ácido glicurônico não sulfatado.

#### 1.4 Estudos da estrutura e conformação de glicanas e polissacarídeos

A conformação de carboidratos é usualmente descrita de acordo com as recomendações da IUPAC (IUPAC-IUB, 1980; IUPAC-IUB, 1983), segundo as quais o arranjo espacial dos átomos de um monossacarídeo piranosídico, em sua forma cíclica, é determinado por uma letra maiúscula, em itálico, designando a forma do anel, e por números, responsáveis pela distinção das possíveis variantes conformacionais para um mesmo resíduo, como nas formas de cadeira  ${}^4C_1$  e  ${}^1C_4$  (Figura 9A e 9B). Adicionalmente, monossacarídeos podem assumir os estados configuracionais dextrógiro (D-) ou levógiro (L-), que diferenciam-se pela organização espacial de seus substituintes (Figura 9C e 9D), e podem ser encontrados, ainda, nos estados anoméricos  $\alpha$  e  $\beta$ , no que tange à posição axial ou equatorial de seu substituinte na posição C1 (Figura 9D e 9E).

Várias técnicas experimentais vêm sendo aplicadas no estudo de glicoproteínas e polissacarídeos. 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).

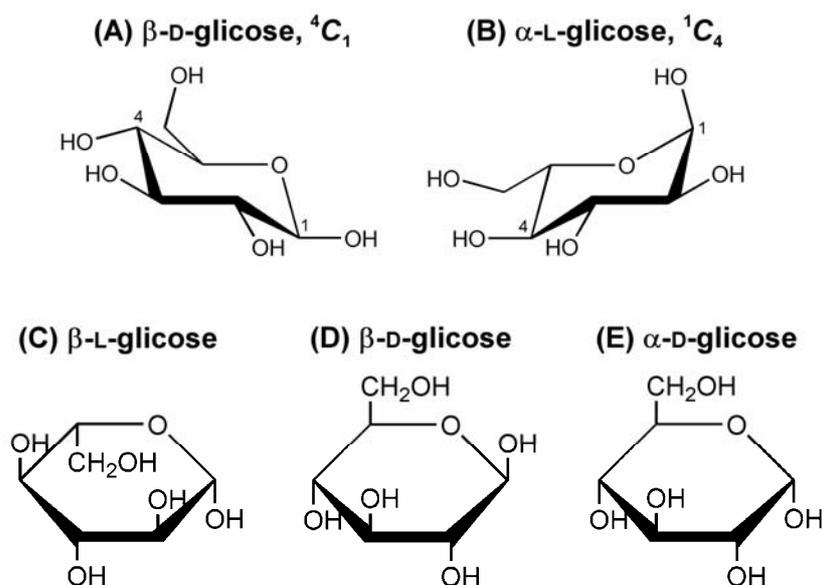


Figura 9: Variantes conformacionais e configuracionais associados a monossacarídeos: conformação do anel piranosídico (A e B), estados levógiro (C) ou dextrógiro (D) e estado anomérico  $\alpha$  ou  $\beta$  (D e E). Nessa representação, a glicose é usada como modelo.

Por outro lado, com o emprego de técnicas de RMN são obtidas 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 interresíduos 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 (Tabela 2). 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.

Tabela 2: Estruturas tridimensionais de glicoproteínas e polissacarídeos determinados por RMN.

Geral	Moléculas	Número de monossacarídeos <sup>b</sup>	Código PDB
	Específica <sup>a</sup>		
Glicoproteína	$\alpha$ -hCG	9	1HD4
	Domínio de adesão – CD2	9	1GYA
	Calcitonina	8	1BZB
	Domínio extracelular – CD152	4	1AH1
	CD59	3	1CDR
	Conotoxina – <i>Conus textile</i>	2	1WCT
	Domínio EGF – fVIIa	1	1FF7
	Domínio lectina – Latrofilina	1	2JXA
	Domínio – Fibronectina	1	1E88
	Domínio – Fibronectina	1	2FN2
Polissacarídeo	Toxina – <i>Corynespora cassicola</i>	1	2HGO
	Domínio EGF – Trombomodulina	1	1DQB
	Heparina	12	1HPN
	Ácido hialurônico	8	2BVK

<sup>a</sup> Referências, em ordem de aparecimento: Erbel *et al.*, 2000; Wyss *et al.*, 1995; Hashimoto *et al.*, 1999; Metzler *et al.*, 1997; Fletcher *et al.*, 1994; Rigby *et al.*, 1999; Kao *et al.*, 1999; Vakonakis *et al.*, 2008; Pickford *et al.*, 2001; Sticht, *et al.*, 1998; Barthe *et al.*, 2007; Wood *et al.*, 2000; Mulloy *et al.*, 1993; Almond *et al.*, 2006;

<sup>b</sup> Número de monossacarídeos unidos em um único oligossacarídeo.

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

O uso de simulações de DM, um procedimento de simulação que consiste na computação do movimento dos átomos em uma molécula, iniciou-se há mais de 30 anos para sistemas biológicos, com o estudo de um sistema protéico envolvendo um inibidor de tripsina pancreática bovina (McCammon *et al.*, 1977). Atualmente,

consiste em uma ferramenta amplamente difundida, que é empregada na investigação da estrutura e dinâmica de biomoléculas em geral, abrangendo estudos que envolvem desde a interação de compostos às suas proteínas-alvo, desnaturação e enovelamento protéicos (Ponder & Case, 2003), até o estudo de exopolissacarídeos bacterianos (Sampaio-Nogueira *et al.*, 2005).

A integração da equação de movimento de Newton  $d^2r_i(t)/dt^2 = F_i / m_i$ , sendo  $d^2r_i(t)/dt^2$  a aceleração,  $m_i$  a massa e  $F_i$  a força sobre um determinado átomo  $i$ , é a característica fundamental dos cálculos de DM que, quando realizada sucessivamente e sobre todos os átomos do sistema, gera uma trajetória de movimento das moléculas em estudo, ou seja, uma seqüência de diferentes posições dos átomos em função do tempo (Leach, 2001).

Esta integração é realizada de forma que uma força  $F_i$  acarreta uma aceleração sobre um determinado átomo  $i$  e, em conseqüência, causa uma mudança de sua posição num intervalo de tempo  $\Delta t$  relativo à aceleração. No entanto, considerando somente tal equação, não é possível determinar o 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 (Leach, 2001). Dessa forma, tais parâmetros são calculados em função de mudanças na energia potencial entre a posição atual e a posição seguinte (a que representará o próximo passo da simulação) sobre cada átomo separadamente. Esta superfície de energia potencial representa a energia de cada molécula, sendo descrita pelo denominado *Campo de Força*.

O campo de força pode ser definido como um conjunto de funções e parametrizações usadas em cálculos de mecânica molecular (de Sant'Anna, 2002). Estas funções (Figura 10) definem as energias de estiramento de ligação e de distorção de ângulo de ligação (tanto de valência quanto de diedro) de uma molécula quando comparadas com a sua conformação não tensionada (aquela caracterizada pelos valores-padrão de comprimentos e de ângulos de ligação). Nesse contexto, os campos de força, juntamente com possíveis termos adicionais de, por exemplo, interação entre átomos não ligados, de efeitos eletrostáticos, de ligação de hidrogênio e de outros efeitos estruturais, expressam o somatório das funções de energia potencial de cada átomo (Leach, 2001) e calculam a energia dos sistemas em função das posições dos núcleos dos átomos, representados por esferas, estando os covalentemente ligados, por sua vez, unidos por molas.

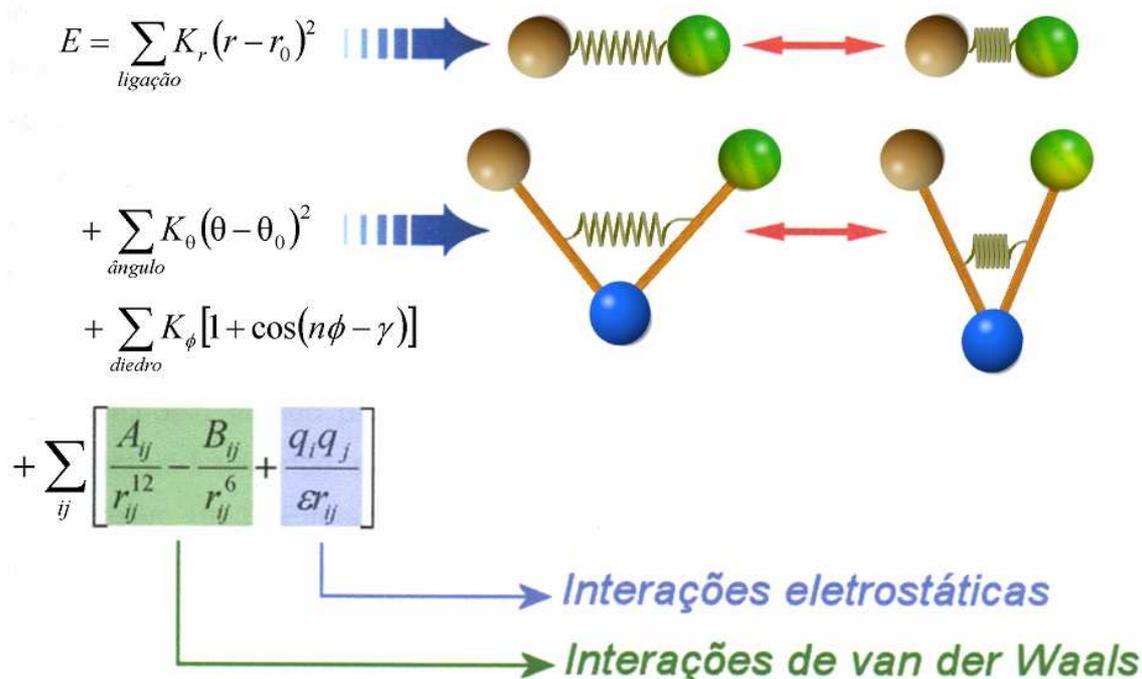


Figura 10: Funções de energia que compõe campos de força. De cima para baixo estão representadas as equações que descrevem o estiramento de ligações químicas, ângulos de ligação, diedros e interações intermoleculares (interações eletrostáticas/Coulômbicas e interações de van der Waals, também identificadas como potencial de Lennard-Jones). Adaptado de Verli, 2004.

Diversos campos de força estão disponíveis para simulações de DM, estando a maioria parametrizada para o estudo de proteínas. Dentre estes, merecem destaque, por seu amplo uso na literatura, AMBER (Case *et al.*, 2005), CHARMM (MacKerell *et al.*, 1998), CVFF (Kitson & Hagler, 1988), TRIPOS (Clark *et al.*, 1989), OPLS (Jorgensen *et al.*, 1988) e GROMOS96 (van Gunsteren *et al.*, 1996). No entanto, embora a maioria dos campos de força disponíveis seja capaz de lidar com proteínas, ou seja, levar uma estrutura conhecida a uma evolução temporal, existe uma carência de parâmetros específicos para outras classes de biomoléculas, tais como ácidos nucleicos, lipídios, compostos sintéticos e carboidratos. 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 dificultam a parametrização de campos de força para estes compostos (Pérez *et al.*, 1998). Adicionalmente, estudos comparando alguns dos principais campos de força para carboidratos atualmente

disponíveis indicam que nenhum deles é consistentemente melhor que os demais, mas que abordagens específicas, seguidas por cada um deles, podem ser melhores para estudar diferentes aspectos desses compostos (Martín-Pastor *et al.*, 1996; Pérez *et al.*, 1998; Corzana *et al.*, 2003; Hemmingsen *et al.*, 2004).

Nesse sentido, o Grupo de Bioinformática Estrutural vem se dedicando ao uso e desenvolvimento de parâmetros para campos de força de MM para a descrição e caracterização de carboidratos em geral. Esse empenho centra-se na parametrização do campo de força GROMOS96, que teve início com a escolha de um esquema de cargas adequado para a descrição desses compostos (Verli & Guimarães, 2004; Becker *et al.*, 2005) e possibilitou a realização de simulações por DM da heparina (Verli & Guimarães, 2004). Após estes estudos, criaram-se as condições necessárias para a descrição do complexo entre um pentassacarídeo sintético, derivado de heparina, e antitrombina (Verli & Guimarães, 2005), da mesma forma que a realização de estudos envolvendo a identificação das forças responsáveis pelo equilíbrio conformacional do resíduo IdoA, presente na estrutura da heparina (Pol-Fachin & Verli, 2008). Adicionalmente, outros trabalhos nessa linha incluem a caracterização da conformação de  $\alpha$ -fucanas e  $\alpha$ -galactanas de ouriço-do-mar (Becker *et al.*, 2007).

Embora o desenvolvimento e aprimoramento de campos de força para carboidratos estejam crescendo gradativamente nos últimos anos, a adequação desses parâmetros no estudo de glicoconjugados, tais como glicoproteínas, ainda encontra-se bastante incipiente. Isso pode ser evidenciado pelo pequeno número de estudos envolvendo DM de glicoproteínas (Tabela 3), incluindo: (1) uma lectina de *Erythrina corallodendron* (Naidoo *et al.*, 1997) e glicopeptídeos derivados da proteína viral gp120 (Huang *et al.*, 1997), utilizando o campo de força CHARMM (Ha *et al.*, 1988); (2) MHC de classe I (Mandal & Mukhopadhyay, 2001), através dos parâmetros do CVFF (Hwang *et al.*, 1998); e (3) prions de camundongo (Wong *et al.*, 2000) e humano (Zuegg & Gready, 2000), uma glicoproteína anticongelamento (Nguyen *et al.*, 2002), o fator da coagulação humana VII (Perera *et al.*, 2002), glicopeptídeos derivados de hemaglutinina (Bosques *et al.*, 2004) e mucina humana (Rubinsten *et al.*, 2004), utilizando campos de força AMBER (Woods *et al.*, 1995; Kirschner *et al.*, 2008).

Tabela 3: Estudos envolvendo DM de glicoproteínas

Glicoproteína / glicopeptídeo <sup>a</sup>	Glicosilação	Campo de Força
Lectina de <i>E. corallodendron</i>	N-	CHARMM
Glicopeptídeos derivados de gp120	N- e O-	CHARMM
MHC de classe I	N-	CVFF
Prion de camundongo	N-	AMBER
Prion humano	N-	AMBER - GLYCAM
Glicoproteína anticongelamento 8	O-	AMBER - GLYCAM
fVII da coagulação humana	O-	AMBER - GLYCAM
Glicopeptídeo derivado de hemaglutinina	N-	AMBER - GLYCAM
Glicopeptídeo derivado de mucina humana	O-	AMBER - GLYCAM

<sup>a</sup> Referências, em ordem de aparecimento: Naidoo *et al.*, 1997; Huang *et al.*, 1997; Mandal & Mukhopadhyay, 2001; Wong *et al.*, 2000; Zuegg & Gready, 2000; Nguyen *et al.*, 2002; Perera *et al.*, 2002; Bosques *et al.*, 2004; Rubinsten *et al.*, 2004.

Embora exista uma série de campos de força sendo utilizados para descrever a conformação de glicoproteínas e polissacarídeos, com destaque para os parâmetros AMBER – GLYCAM, a utilização desses campos de força geralmente exige uma grande capacidade computacional, impedindo uma busca suficientemente ampla pelo espaço conformacional dessas moléculas. Por consequência, a descrição conformacional desses sistemas fica prejudicada e, assim, a confiabilidade das predições realizadas. Nesse sentido, o desenvolvimento e/ou aprimoramento de campos de força, capazes de gerar respostas mais rápidas e suficientemente precisas em relação a dados experimentais prévios, consiste em uma promissora estratégia para a descrição e predição conformacional de carboidratos e glicoconjugados através de simulações de DM. Adicionalmente, no caso específico do campo de força GROMOS96, tais parâmetros estão baseados em ferramentas disponíveis gratuitamente, já tendo demonstrado, em trabalhos prévios, que podem ser utilizados na determinação de GAGs complexos, da mesma forma que na interação dos mesmos com suas moléculas alvo e na descrição da conformação de proteínas.

## 2 Objetivos

A partir do exposto, o presente trabalho se insere no contexto da caracterização conformacional de polissacarídeos e glicoproteínas através de técnicas de MM. Nesse contexto, o trabalho visa comparar o comportamento conformacional de glicoproteínas, determinadas por RMN, com o obtido através de simulações de DM, empregando 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. Ainda, o trabalho visa confrontar a geometria das ligações glicosídicas dos oligossacarídeos que compõe tais glicoproteínas com as obtidas através de simulações de DM dos dissacarídeos isolados. Espera-se, a partir dos resultados obtidos, validar a estratégia empregada e, desta forma, subsidiar futuros estudos de outros sistemas envolvendo glicoproteínas, na ausência de dados experimentais prévios. Dessa forma, as seguintes metas foram estabelecidas:

- Caracterização das propriedades conformacionais de uma  $\beta$ -galactana em comparação com as de uma  $\alpha$ -galactana, ambas de ouriço-do-mar;
- Viabilização da descrição computacional de glicoproteínas unidas a oligossacarídeos N- e O-ligados;
- Viabilização da predição conformacional de oligossacarídeos ligados a proteínas através de DM;

### 3 Metodologia

#### 3.1 Programas utilizados

Diversas metodologias de modelagem molecular foram utilizadas no presente trabalho, incluindo MM, DM e métodos *ab initio*. Os protocolos referentes a cada um destes métodos estão descritos em detalhes a seguir.

Os programas utilizados incluem:

- Ferramentas de visualização de moléculas: VMD v1.8.6 (Humphrey *et al.*, 1996), PyMol v0.99 (DeLano, 2002) e MOLGEN (Schaffenaar & Noordik, 2000);
- Programa para cálculo *ab initio*: GAMESS (Schmidt *et al.*, 1993);
- Programa para construção de mapas de contorno e simulações de dinâmica molecular: GROMACS v3.3.3 (van der Spoel *et al.*, 2005);
- Programa para geração de topologias: PRODRG *Beta* v2.5 (Schuettelkopf & van Aalten, 2004);
- Programas para análise de estrutura secundária: PROCHECK (Laskowski *et al.*, 1993).

#### 3.2 Cálculos utilizando métodos *ab initio*

Uma importante etapa na parametrização de campos de força de mecânica molecular é a obtenção de cargas atômicas para as moléculas em estudo. Nesse sentido, nosso grupo de pesquisas vem aplicando, com sucesso, um procedimento baseado em cargas atômicas de Löwdin, obtidas por cálculos Hartree Fock, na descrição conformacional de polissacarídeos sulfatados (Verli & Guimarães, 2004; Becker *et al.*, 2005) e monossacarídeos componentes de glicanas (Becker, 2007).

Tendo-se em vista o elevado custo computacional associado aos métodos *ab initio*, as moléculas utilizadas para a obtenção destas cargas atômicas foram simplificadas para incluírem somente os resíduos em suas formas isoladas e quando envolvidos nas ligações glicosídicas entre aminoácidos (Asn e Ser) e monossacarídeos (GlcNAc e Fuc). Dessa forma, tais resíduos foram construídos utilizando o programa Molden (Schaffenaar, 1997) e, posteriormente, submetidos a cálculos de minimização de energia utilizando a base HF/3-21G no programa GAMESS (Schmidt *et al.*, 1993). As conformações de mínimo de energia então

obtidas foram empregadas como geometrias de entrada para cálculos *single point* na base HF/6-31G\*\*, de forma a gerar cargas atômicas segundo o esquema de Löwdin (no item de anexos 9.3, encontram-se as cargas atômicas obtidas para os grupamentos estudados até então).

### 3.3 Construção de mapas de contorno por mecânica molecular

As preferências conformacionais de dissacarídeos foram avaliadas através de mapas de contorno, análogos aos mapas de Ramachandran, utilizando campos de força de mecânica molecular. Nesse contexto, utilizando como base os parâmetros do campo de força GROMOS96 (van Gunsteren *et al.*, 1996) e as cargas atômicas de Löwdin, na base HF/6-31G\*\* (Verli & Guimarães, 2004; Becker *et al.*, 2005), 15 mapas de contorno foram calculados a partir da rotação dos diedros componentes de 13 ligações glicosídicas de dissacarídeos (Tabela 4).

Tabela 4: Dissacarídeos estudados

Ligação Glicosídica	Mínimos de energia refinados em solução	Tempo total de simulação
$\alpha$ -D-Man-(1→2)- $\alpha$ -D-Man	2	0.2 $\mu$ s
$\beta$ -D-GlcNAc-(1→2)- $\alpha$ -D-Man	2	0.2 $\mu$ s
$\alpha$ -L-Gal-(1→3)- $\alpha$ -L-Gal	3	0.3 $\mu$ s
$\alpha$ -L-Gal2S-(1→3)- $\alpha$ -L-Gal2S	1	0.1 $\mu$ s
$\beta$ -D-Gal-(1→3)- $\beta$ -D-Gal	5	0.5 $\mu$ s
$\beta$ -D-Gal2S-(1→3)- $\beta$ -D-Gal	2	0.2 $\mu$ s
$\beta$ -D-Gal-(1→3)- $\beta$ -D-Gal2S	2	0.2 $\mu$ s
$\alpha$ -D-Man-(1→3)- $\alpha$ -D-Man	2	0.2 $\mu$ s
$\beta$ -D-Gal-(1→4)- $\beta$ -D-GlcNAc	2	0.2 $\mu$ s
$\beta$ -D-Man-(1→4)- $\beta$ -D-GlcNAc	2	0.2 $\mu$ s
$\beta$ -D-GlcNAc-(1→4)- $\beta$ -D-GlcNAc	4	0.4 $\mu$ s
$\alpha$ -L-Fuc-(1→6)- $\beta$ -D-GlcNAc	3	0.3 $\mu$ s
$\alpha$ -D-Man-(1→6)- $\alpha$ -D-Man	2	0.2 $\mu$ s
<b>13 dissacarídeos</b>	<b>32 confôrmers</b>	<b>3.2 <math>\mu</math>s</b>

Dessa forma, no caso de ligações 1→2, 1→3 e 1→4, os ângulos  $\Phi$  e  $\Psi$  foram rotados de  $-180^\circ$  a  $150^\circ$ , em passos de  $30^\circ$ , gerando um total de 144 confôrmers para cada dissacarídeo. No caso de ligações 1→6, como a ligação glicosídica é

composta por três diedros:  $\Phi$ ,  $\Psi$  e  $\omega$ , foram gerados 144 confôrmeros para cada par de ângulos, ou seja, 144 para o par  $\Phi$  versus  $\Psi$  e 144 para o par  $\Psi$  versus  $\omega$ . Para todos esses tipos de ligação, durante a geração dos confôrmeros, utilizam-se restrições (diedros impróprios) somente para os diedros da ligação glicosídica em estudo, de maneira que tais ângulos são fixados em cada uma das 144 possíveis combinações dos dois diedros em estudo, enquanto o restante da molécula encontra-se livre para ter sua energia minimizada. O protocolo utilizado para tal envolve uma minimização de energia utilizando o algoritmo de gradiente conjugado e, em seguida, uma DM com duração de 20 ps a temperatura de 10K e tempo de integração de 0,5 fs a fim de reforçar a busca pelo arranjo mais estável da conformação avaliada (Becker *et al.*, 2007; Pol-Fachin & Verli, 2008). Os valores de energia das 144 conformações de cada par de diedros são então empregados na construção de um mapa de contorno, considerando-se a estabilidade relativa à conformação mais estável de todas.

### **3.4 Simulações de DM**

#### **3.4.1 Protocolo de simulação**

O protocolo geral de simulação foi baseado em procedimentos previamente descritos (de Groot & Grubmüller, 2001). Todas as simulações foram realizadas à temperatura fisiológica (310K), com duração variável, de acordo com o sistema em estudo (de 50 a 100 ns) (Tabelas 4 e 5). Cada molécula (proteína, glicoproteína ou dissacarídeo) foi solvatada numa caixa retangular 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 (PME, Darden *et al.*, 1993), utilizando raios de corte de *Coulomb* e de van der Waals de 9 Å. 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 constante dielétrica do meio foi tratada como  $\epsilon = 1$ .

Tabela 5: Proteínas e glicoproteínas estudadas

Proteína / Glicoproteína	Modelos simulados	Tempo total de simulação
Domínio EGF-like não glicosilado – PDB 1F7E	1	0.05 $\mu$ s
CD59 não glicosilado – PDB 1CDQ	1	0.05 $\mu$ s
Domínio de adesão CD2 não glicosilado – PDB 1CDB	1	0.05 $\mu$ s
$\alpha$ -hCG não glicosilada – PDB 1DZ7	1	0.05 $\mu$ s
Domínio EGF-like glicosilado – PDB 1FF7	1	0.05 $\mu$ s
CD59 glicosilada – PDB 1CDR	2	0.1 $\mu$ s
CD59 glicosilada – PDB 1CDS	1	0.05 $\mu$ s
Domínio de adesão CD2 glicosilado – PDB 1GYA	2	0.1 $\mu$ s
$\alpha$ -hCG glicosilada – PDB 1HD4	2	0.1 $\mu$ s
<b>9 moléculas</b>	<b>12 modelos</b>	<b>0.6 <math>\mu</math>s</b>

As simulações por DM podem ser divididas em três etapas: *termalização*, *equilíbrio* e *recolhimento de dados*. A primeira delas envolve o aquecimento gradativo do sistema, visando uniformizar as energias contidas na estrutura cristalográfica ou de RMN e, desta forma, evitar deformações nas moléculas em estudo. Nesta etapa, após 1 ps de restrição de posição, cada sistema foi aquecido lentamente de 50 K a 310 K, de maneira que, em cada um dos seis passos de 5 ps, há o aumento da temperatura em 50 K (Figura 11). Após a *termalização* do sistema, a simulação prossegue na temperatura de equilíbrio de 310 K, pelo tempo estipulado.

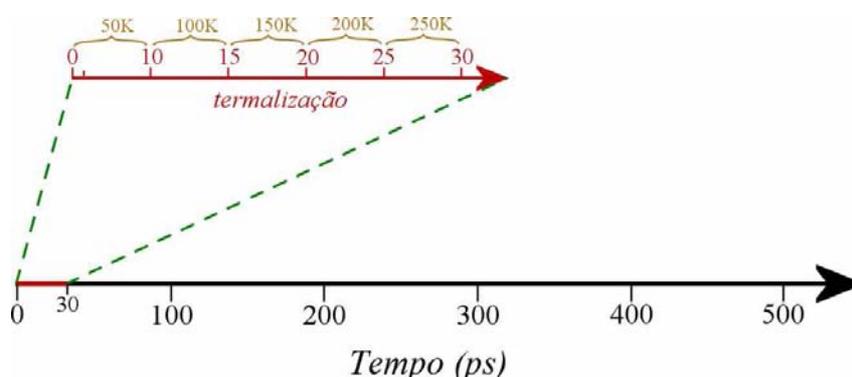


Figura 11: Esquema das etapas que compõe as simulações de DM a 310K. De 0ps a 30ps, uma série de etapas, de 5ps cada, aquece lentamente o sistema em incrementos de 50K. Após 30ps, o sistema está a 310K e se mantém nesta temperatura pelo resto da simulação (Verli, 2005).

### 3.4.2 Construção de topologias

A topologia é um arquivo contendo parâmetros que descrevem as propriedades atômicas de uma determinada estrutura química por meio da descrição da força e tipo de cada ligação química, ângulos de ligação e de diedros, assim como os parâmetros de Lennard Jones e as cargas atômicas. Desta forma, e considerando a ausência de topologias previamente existentes, no campo de força GROMOS96 43a1, para a maioria dos resíduos de carboidratos empregados neste estudo, e visando viabilizar os objetivos traçados, utilizou-se uma ferramenta de construção de topologias denominada PRODRG (Schuettelkopf & van Aalten, 2004).

O PRODRG é um servidor que funciona na *World Wide Web*, gerando um arquivo de topologia para os sistemas de interesse a partir de um arquivo PDB que contenha sua estrutura. Esta ferramenta, no entanto, gera somente um conjunto de parâmetros iniciais, que necessitam de alguns refinamentos, que incluem o acréscimo de cargas atômicas, calculadas por métodos *ab-initio* (descrito no item 3.2), e a adição de diedros impróprios (Tabela 6), necessários para preservar os estados conformacionais ( ${}^4C_1$  ou  ${}^1C_4$ ) dos monossacarídeos em decorrência do emprego do campo de força GROMOS96 43a1. Estes diedros impróprios, além de manter a integridade conformacional do anel em estudo, de acordo com dados experimentais, tornam possíveis também mudanças no estado conformacional de uma piranose ou furanose, por exemplo, de forma a explorar o papel de diferentes estados conformacionais em um determinado fenômeno biológico (Verli & Guimarães, 2005).

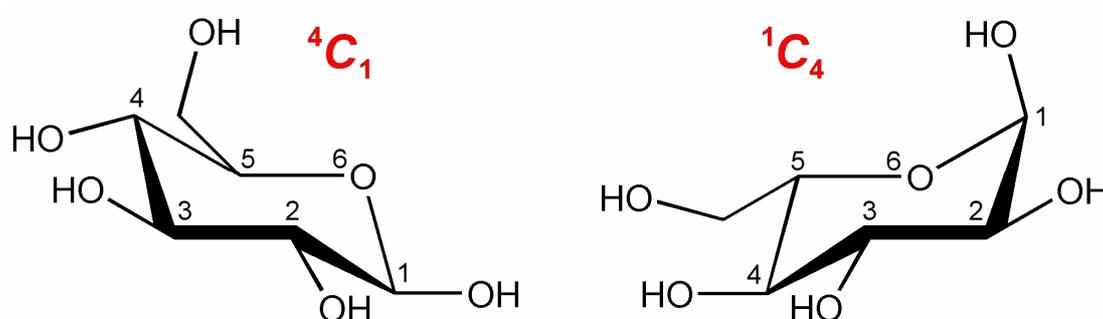
No item 9.4, estão descritas topologias, parametrizadas em função desta dissertação, para uma série de monossacarídeos que compõe glicoproteínas e para os aminoácidos em que a parte sacarídica dessas moléculas está ligada.

### 3.4.3 Validação das simulações de DM

A confiabilidade das predições baseadas em cálculos de DM depende, principalmente, da precisão do campo de força utilizado e do tamanho da amostragem do espaço conformacional realizada durante a simulação, que necessita ser suficientemente ampla para descrever as propriedades do sistema (van Gunsteren & Mark, 1998). De modo geral, o teste final do modelo teórico produzido reside na sua comparação com propriedades experimentais, ou seja, a precisão do método é julgada a partir da observação ou constatação, nas

simulações realizadas, de quão precisas são reproduzidas quantidades conhecidas (van Gunsteren & Berendsen, 1990; Karplus & Petsko, 1990). Dessa forma, a validação dos resultados obtidos nas simulações de DM foi realizada através da comparação com dados experimentais prévios, com destaque para estruturas cristalográficas, dados bioquímicos e, principalmente, conformações oriundas de experimentos de RMN, assim como através da observação da estabilidade dos sistemas estudados, ou seja, pela manutenção, durante as simulações, de propriedades que incluem estrutura secundária, energia, densidade e volume.

Tabela 6: Definição dos diedros impróprios utilizados para definir a conformação dos monossacarídeos contidos nos sistemas simulados.



Seqüência de átomos definindo o diedro <sup>a</sup>	Ângulos (em graus)	
	<sup>4</sup> C <sub>1</sub>	<sup>1</sup> C <sub>4</sub>
5 – 2 – 4 – 1	2,0	-2,0
5 – 2 – 3 – 1	23,0	-23,0
5 – 2 – 3 – 6	-2,0	2,0

<sup>a</sup> Na figura, a glicose é utilizada como modelo.

## 4 Resultados

### 4.1 *Preâmbulo*

Os resultados obtidos serão apresentados a seguir na forma dos trabalhos submetidos ou publicados durante a realização da presente Dissertação. Estes trabalhos estão citados abaixo, assim como uma breve descrição sobre cada um.

- Michelle O. de Castro, Vitor H. Pomin, Livia L. dos Santos, Ana Cristina E. S. Vilela-Silva, Noritaka Hirohashi, **Laércio Pol-Fachin**, Hugo Verli, Paulo Antônio de Souza Mourão: 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.

Este trabalho apresenta a aplicação de ferramentas de modelagem molecular no estudo conformacional de um polissacarídeo sulfatado,  $\beta$ -galactana, isolado a partir de ovos do ouriço-do-mar *Glyptocidaris crenularis*.

- **Laércio Pol-Fachin**, Cláudia Lemelle Fernandes, Hugo Verli: GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations. *Carbohydr. Res.*, **2009**, *344*; 491-500.

Este artigo avalia a capacidade do campo de força GROMOS96 43a1, adicionado por cargas atômicas de Löwdin, na descrição conformacional de glicoproteínas utilizando simulações de DM, tendo como base de comparação dados experimentais de RMN.

- 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.

Por fim, 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 seguir, serão apresentados um resumo e a íntegra de cada manuscrito.

## 4.2 Trabalho I

Este trabalho descreve a identificação de uma  $\beta$ -galactana, a partir de ovos da espécie de ouriço-do-mar *Glyptocidaris crenularis*, que consiste em um polissacarídeo sulfatado composto por resíduos de D-galactose. Até então, haviam sido encontradas apenas  $\alpha$ -galactanas e  $\alpha$ -fucanas nesses organismos, as quais foram atribuídas atividades relacionadas à reação acrossômica desses animais. Dessa forma, testes realizados com a  $\beta$ -galactana recém identificada demonstraram que tal polissacarídeo bloqueia a capacidade de fertilização do esperma em diferentes espécies de ouriço-do-mar, enquanto  $\alpha$ -galactanas e  $\alpha$ -fucanas ativam esse processo. Neste contexto, foram realizados estudos de modelagem molecular destas  $\alpha$ - e  $\beta$ -galactanas de forma a elucidar as razões moleculares para atividades antagônicas como as mencionadas.

Os resultados obtidos sugerem uma maior flexibilidade da  $\beta$ -galactana, em comparação com a da  $\alpha$ -galactana, de forma que as unidades dissacarídicas que a compõe podem assumir, potencialmente, quatro sub-estados conformacionais em solução, enquanto que os dissacarídeos que formam a  $\alpha$ -galactana parecem adotar uma única conformação prevalente. Adicionalmente, as conformações adotadas pelos dissacarídeos de  $\beta$ -galactana não são observadas nos dissacarídeos da  $\alpha$ -galactana.

### **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**

Michelle Oliveira de Castro, Vitor Hugo Pomin, Livia Loiola dos Santos, Ana Cristina Espirito Santo de Vilela Silva, Noritaka Hirohashi, **Laercio Pol-Fachin**, Hugo Verli,  
Paulo Antônio de Souza Mourão

*Journal of Biological Chemistry*, **2009**, 284; 18790-18800

# A Unique 2-Sulfated $\beta$ -Galactan from the Egg Jelly of the Sea Urchin *Glyptocidaris crenularis*

## CONFORMATION FLEXIBILITY VERSUS INDUCTION OF THE SPERM ACROSOME REACTION\*

Received for publication, March 25, 2009, and in revised form, April 8, 2009. Published, JBC Papers in Press, April 29, 2009, DOI 10.1074/jbc.M109.005702

Michelle O. Castro<sup>‡§</sup>, Vitor H. Pomin<sup>‡§</sup>, Livia L. Santos<sup>‡§</sup>, Ana-Cristina E. S. Vilela-Silva<sup>‡¶</sup>, Noritaka Hirohashi<sup>||</sup>, Laércio Pol-Fachin<sup>\*\*</sup>, Hugo Verli<sup>\*\*\*††</sup>, and Paulo A. S. Mourão<sup>‡§1</sup>

From the <sup>‡</sup>Laboratório de Tecido Conjuntivo, Hospital Universitário Clementino Fraga Filho, <sup>§</sup>Instituto de Bioquímica Médica, and <sup>¶</sup>Instituto de Ciências Biomédicas, Universidade Federal do Rio de Janeiro, Caixa Postal 68041, Rio de Janeiro, RJ 21941-590, Brazil, the <sup>||</sup>Genetic Counseling Course, Graduate School of Humanities and Sciences, Ochanomizu University, Tokyo 112-8610, Japan, and the <sup>\*\*</sup>Programa de Pós-Graduação em Biologia Celular e Molecular, Centro de Biotecnologia, and <sup>††</sup>Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Caixa Postal 15005, Porto Alegre, RS 91500-970, Brazil

Sulfated polysaccharides from the egg jelly of sea urchins act as species-specific inducers of the sperm acrosome reaction, which is a rare molecular mechanism of carbohydrate-induced signal-transduction event in animal cells. The sea urchin polysaccharides differ in monosaccharide composition (L-fucose or L-galactose), glycosylation, and sulfation sites, but they are always in the  $\alpha$ -anomeric configuration. Herein, structural analysis of the polysaccharide from the sea urchin *Glyptocidaris crenularis* surprisingly revealed a unique sulfated  $\beta$ -D-galactan composed by (3- $\beta$ -D-Galp-2(OSO<sub>3</sub>)-1 $\rightarrow$ 3- $\beta$ -D-Galp-1)<sub>n</sub> repeating units. Subsequently, we used the *G. crenularis* galactan to compare different 2-sulfated polysaccharides as inducers of the acrosome reaction using homologous and heterologous sperm. We also tested the effect of chemically over-sulfated galactans. Intriguingly, the anomeric configuration of the glycosidic linkage rather than the monosaccharide composition (galactose or fucose) is the preferential structural requirement for the effect of these polysaccharides on sea urchin fertilization. Nuclear magnetic resonance and molecular dynamics indicate that sulfated  $\alpha$ -galactan or  $\alpha$ -fucan have less dynamic structural behavior, exhibiting fewer conformational populations, with an almost exclusive conformational state with glycosidic dihedral angles  $\Phi/\Psi = -102^\circ/131^\circ$ . The preponderant conformer observed in the sulfated  $\alpha$ -galactan or  $\alpha$ -fucan is not observed among populations in the  $\beta$ -form despite its more flexible structure in solution. Possibly, a proper spatial arrangement is required for interaction of the sea urchin-sulfated polysaccharides with the specific sperm receptor.

The evolution of barriers to inter-specific hybridization is a crucial step in the fertilization of free-spawning marine invertebrates. In sea urchins the molecular recognition between sperm and egg ensures species recognition. The jelly coat sur-

rounding sea urchin eggs is not a simple accessory structure; it is considerably complex on a molecular level and intimately involved in gamete recognition. It contains sulfated polysaccharides, sialoglycans, and peptides.

Structural changes in the sulfated polysaccharide from the egg jelly of sea urchins modulate cell-cell recognition and species specificity leading to exocytosis of the acrosomal vesicle, the acrosome reaction. This is a crucial event for the recognition between male and female gametes, leading to the fertilization success, and is also what prevents intercrosses. The sulfated polysaccharide from the egg jelly recognizes its specific receptor present in the sperm. Apart from the sialoglycans that act in synergy with the sulfated polysaccharides, other components of the egg jelly do not possess acrosome reaction-inducing activity (1). The sulfated polysaccharide-mediated mechanism of sperm-egg recognition co-exists with that of bindin and its receptor in the egg (2–4).

The sulfated polysaccharides from sea urchin show species-specific structures composed of repetitive units (mono-, tri-, and tetrasaccharides) that differ in the monosaccharide backbone (L-fucose or L-galactose), glycosidic linkage (3- or 4-linked), and sulfation (2- and/or 4-sulfation). However, they are always in the  $\alpha$ -enantiomeric configuration (4, 5). Previous studies from our laboratory have demonstrated that sea urchin-sulfated polysaccharides induce the acrosome reaction in a species-specific way. In some cases the sperm from a certain species of sea urchin recognizes the sulfated polysaccharide containing a similar structure from a different species. For example, the egg jelly from *Strongylocentrotus franciscanus* contains a 2-sulfated, 3-linked  $\alpha$ -fucan, but the sperm from this species recognizes a heterologous 2-sulfated, 3-linked  $\alpha$ -galactan from *Echinometra lucunter* (6).

We now extended our studies to the sulfated polysaccharides of the sea urchin *Glyptocidaris crenularis* (7). Surprisingly, we observed that this species contains a unique sulfated  $\beta$ -D-galactan composed of repetitive disaccharide units alternating 2-sulfated and non-sulfated 3-linked units. This polymer is markedly distinct from all other sea urchin-sulfated polysaccharides described so far that are composed of units on  $\alpha$ -L-configuration. Furthermore, this sea urchin does not contain sialoglycans, which are commonly found in the echinoderm egg jelly.

\* This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico, Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

<sup>1</sup> To whom correspondence should be addressed: Instituto de Bioquímica Médica, Universidade Federal do Rio de Janeiro, Caixa Postal 68041, Rio de Janeiro, RJ 21941-590, Brazil. Tel.: 55-21-25622090; E-mail: pmourao@hucff.ufrj.br.

We used this new sulfated  $\beta$ -galactan to investigate the acrosome reaction in a further molecular detail using homologous and heterologous sperm. We tested three 2-sulfated polysaccharides that differ in their conformation ( $\alpha$  or  $\beta$ ) and monosaccharide composition (galactose or fucose) as inducers of the sperm acrosome reaction. We aimed to establish the structure versus biological activity of the echinoderm polysaccharides, including structural features at a conformational level.

## EXPERIMENTAL PROCEDURES

**Extraction of Sulfated Galactan from Egg Jelly of Sea Urchins**—Mature females of *G. crenularis* were harvested by dredging in Noheji Bay, Japan. Eggs were spawned into sea water by intracelomic injection of 0.5 M KCl. The crude egg jelly was isolated by pouring eggs repeatedly through nylon mesh, prepared as a 30,000  $\times$  g supernatant, and stored at  $-20^\circ\text{C}$ , lyophilized, and dialyzed against distilled water (8). The acidic polysaccharides were extracted from the jelly coat by papain digestion and partially purified by ethanol precipitation, as described previously (9).

**Purification of Sulfated Galactan**—Approximately 20 mg of the crude polysaccharides from *G. crenularis* was applied to a Mono Q fast protein liquid chromatography (FPLC) column (HR5/5; Amersham Biosciences) equilibrated with 20 mM Tris-HCl (pH 8.0) and coupled to a FPLC system. The column was washed with 10 ml of the same solution and then eluted with a linear gradient of 0–3 M NaCl. Fractions of 0.5 ml were collected at a flow rate of 0.45 ml/min and checked for hexose (10), sialic acid (11), and metachromatic property (12). The NaCl concentration was estimated by conductivity. Fractions were pooled, dialyzed against distilled water, and lyophilized.

**Chemical Analyses**—After acid hydrolysis of the polysaccharide (5.0 M trifluoroacetic acid for 5 h at  $100^\circ\text{C}$ ), sulfate was measured by the  $\text{BaCl}_2$ /gelatin method (13). The presence of hexoses and 6-deoxyhexoses in the acid hydrolysates was estimated by paper chromatography in 1-butanol/pyridine/water (3:2:1, v/v) for 48 h. In addition, alditol acetate derivatives were analyzed by gas-liquid chromatography/mass spectrometry (14).

**Determination of D or L Configuration of Galactose**—The enantiomeric form of the galactose was assigned based on the analysis of the acetylated (–)-2-butyl glycoside, as described (15). Galactose obtained after acid hydrolysis of the polysaccharide from *G. crenularis* (1 mg, see above) was mixed with 0.5 ml of (–)-2-butanol (Aldrich) containing 1 M HCl. After butanolysis for 18 h at  $80^\circ\text{C}$ , the solution was neutralized with  $\text{Ag}_2\text{CO}_3$ , and the supernatant was concentrated and dissolved in 50  $\mu\text{l}$  of distilled water. Thereafter, alditol acetate derivative was prepared (14) and analyzed on a DB-5 GLC column. The temperature of the column was programmed to increase in a linear gradient of 120 to  $240^\circ\text{C}$  at  $2^\circ\text{C}/\text{min}$ . The injector and detector temperatures were 220 and  $260^\circ\text{C}$ , respectively. Appropriate controls of acetylated (–)-2-butyl-D- and L-galactosides were analyzed under the same conditions.

**Agarose and Polyacrylamide Gel Electrophoresis**—The sulfated galactan was analyzed by agarose gel electrophoresis, as described previously (16, 17). About 15  $\mu\text{g}$  of the purified sulfated polysaccharide was applied to a 0.5% agarose gel and run

for 1 h at 110 V in 0.05 M 1,3-diaminopropane acetate (pH 9.0). The sulfated polysaccharides in the gel were fixed with 0.1% *N*-cetyl-*N,N,N*-trimethylammonium bromide solution. After 12 h, the gel was dried and stained with 0.1% toluidine blue in acetic acid:ethanol:water (0.1:5:5, v/v).

The average molecular mass of the sulfated galactan was estimated by comparison with the electrophoretic mobility of standard compounds (18). The sulfated polysaccharides ( $\sim 10$   $\mu\text{g}$  of each) were applied to a 1-mm-thick 10% polyacrylamide slab gel in 0.02 M sodium barbital (pH 8.6). After electrophoresis (100 V for 30 min), the sulfated polysaccharides were stained with 0.1% toluidine blue in 1% acetic acid and washed for about 1 h in 1% acetic acid.

**Desulfation and Oversulfation Procedures**—Desulfation of the sulfated galactan was performed as described previously (16, 19). About 20 mg of the polysaccharide was dissolved in 5 ml of distilled water and mixed with 1 g (dry weight) of Dowex 50-W ( $\text{H}^+$ , 200–400 mesh). After neutralization with pyridine, the solution was lyophilized. The resulting pyridinium salt was dissolved in 2.5 ml of dimethyl sulfoxide/methanol (9:1, v/v). The mixture was heated at  $80^\circ\text{C}$  for 4 h, and the desulfated product was exhaustively dialyzed against distilled water and lyophilized. The extent of desulfation was estimated by the molar ratio of sulfate/total sugar. This method allowed us to detect desulfation up to a molar ratio of  $\leq 0.1$  sulfate/total sugar. About 5 mg of desulfated polysaccharide was obtained at the end of the reaction. Over-sulfated galactans were prepared through chemical sulfonation of the native polysaccharides (6). In a control reaction we followed sulfonation of chondroitin 4-sulfate. The reaction reached about 72.1% of the total sites available for sulfation on the entire chondroitin sulfate backbone, as indicated by NMR analysis. It was not possible to run NMR spectra of the over-sulfated galactans due to scarce material.

**NMR Experiments**— $^1\text{H}$  and  $^{13}\text{C}$  one- and two-dimensional spectra of the native sulfated galactan and of its desulfated derivative were recorded using a Bruker DRX 400 MHz apparatus with a triple resonance probe, as detailed previously (18). About 5 mg of each sample was dissolved in 0.5 ml of 99.9% deuterium oxide (Cambridge Isotope Laboratory, Cambridge, MA). All spectra were recorded at  $50^\circ\text{C}$  with HOD suppression by presaturation. One-dimensional  $^1\text{H}$  NMR spectra were recorded with 16 scans. Two-dimensional  $^1\text{H}/^1\text{H}$  COSY,<sup>2</sup> TOCSY, NOESY, and  $^{13}\text{C},^1\text{H}$  HSQC spectra were recorded using states-TPP1 (states-time proportion phase incrementation) for quadrature detection in the indirect dimension. TOCSY spectra were run with  $4046 \times 400$  points with a spinlock field of 10 kHz and a mixed time of 80 ms. NOESY spectra were recorded with a mixing time of 100 ms.  $^{13}\text{C},^1\text{H}$  HSQC spectra were run with  $1024 \times 256$  points and GARP (globally optimized alternating phase rectangular pulses) for decoupling. Chemical shifts are relative to external trimethylsilylpropionic acid at 0 ppm for  $^1\text{H}$  and to methanol for  $^{13}\text{C}$ .

<sup>2</sup> The abbreviations used are: COSY, correlation spectroscopy; NOESY, nuclear overhauser effect (NOE) spectroscopy; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single quantum coherence; MD, molecular dynamics.

## Sulfated $\beta$ -Galactan from Sea Urchin

**Molecular Dynamics**—All calculations were performed using GROMACS simulation suite (20) and GROMOS96 force field (21). Briefly, structures of disaccharide units containing  $\alpha$ -L-Galp-(1 $\rightarrow$ 3)- $\alpha$ -L-Galp,  $\alpha$ -L-Galp-2(SO<sub>4</sub>)-(1 $\rightarrow$ 3)- $\alpha$ -L-Galp-2(SO<sub>4</sub>),  $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\beta$ -D-Galp,  $\beta$ -D-Galp-2(SO<sub>4</sub>)-(1 $\rightarrow$ 3)- $\beta$ -D-Galp, and  $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\beta$ -D-Galp-2(SO<sub>4</sub>) were submitted to the PRODRG server (22), and the initial geometries and crude topologies were retrieved. These topologies were supplied with Löwdin HF/6-31G\*\* atomic charges (23, 24) and submitted to conformational analysis by varying the  $\phi$  and  $\psi$  dihedral angles from  $-180$  to  $180^\circ$ , with a  $30^\circ$  step, in a total of 144 conformers for each linkage. Each conformation was further refined in a 20-ps molecular dynamics at 10 K, with an integration step of 0.5 fs (25). The relative stabilities of the conformations were used to construct relaxed energy contour plots. The minimum energy conformations described in these plots were submitted to 0.1  $\mu$ s molecular dynamics (MD) simulations in aqueous solutions using the SPC water model (26) following a protocol previously described (23, 24, 27). A triclinic water box under periodic boundary conditions was employed using a 10-Å minimum distance from solute to the box faces. Counter ions (Na<sup>+</sup>) were added to neutralize the system. The Lincs method (28) was applied to constraint covalent bond lengths, allowing an integration step of 2 fs after an initial energy minimization using the Steepest Descents algorithm. All simulations applied the Particle-Mesh Ewald method (29). Temperature and pressure were kept constant by coupling carbohydrate, ions, and solvent to external temperature and pressure baths with coupling constants of  $t = 0.1$  and  $0.5$  ps, respectively (26). The reference temperature was adjusted to 310 K. The relative orientation of a pair of contiguous carbohydrate residues is described by two torsional angles at the glycosidic linkage, denoted  $\phi$  and  $\psi$ , as follows:  $\phi = \text{O-5-C-1-O-1-C-3'}$ , and  $\psi = \text{C-1-O-1-C-3'-C-2'}$ .

**Fertilization Block by Sperm Preincubation with Sulfated Galactans**—*G. crenularis* sperm were collected as undiluted semen, stored on ice, and diluted in ice-chilled seawater shortly before use. Sulfated galactans (20  $\mu$ l) were dissolved in sea water at the final concentration of 4 mg/ml and placed in a test tube. Thereafter, 100-fold diluted sperm suspension (5  $\mu$ l) was added and incubated for 10 min at room temperature. Unfertilized *G. crenularis* eggs spawned by intercelomic injection with 0.5 M KCl were placed in a 48-well culture dish filled with 150  $\mu$ l of seawater. After a 10-min incubation of sperm with sulfated galactan, such sperm (5  $\mu$ l) were inseminated for 30 min at 20 °C. The percentage of fertilization (judged by the presence/absence of the fertilization envelope) was scored under the microscope by counting  $\sim 200$  eggs. Three independent experiments were performed.

**Acrosome Reaction Assays**—The method used to access the acrosome reaction was slightly modified from Su and Vacquier (30). Briefly, sperm were spawned by intracelomic injection of 0.5 M KCl (0.5 ml/animal), collected undiluted, and stored on ice before dilution. They were diluted 1:5 in 10 mM HEPES-buffered sea water (pH 7.9). Immediately after dilution,  $\sim 25$   $\mu$ l of the sperm suspension were mixed with 50  $\mu$ l of the polysaccharide solution. The sugar content of these solutions was previously quantified by the phenol-sulfuric acid assay (10). After 5

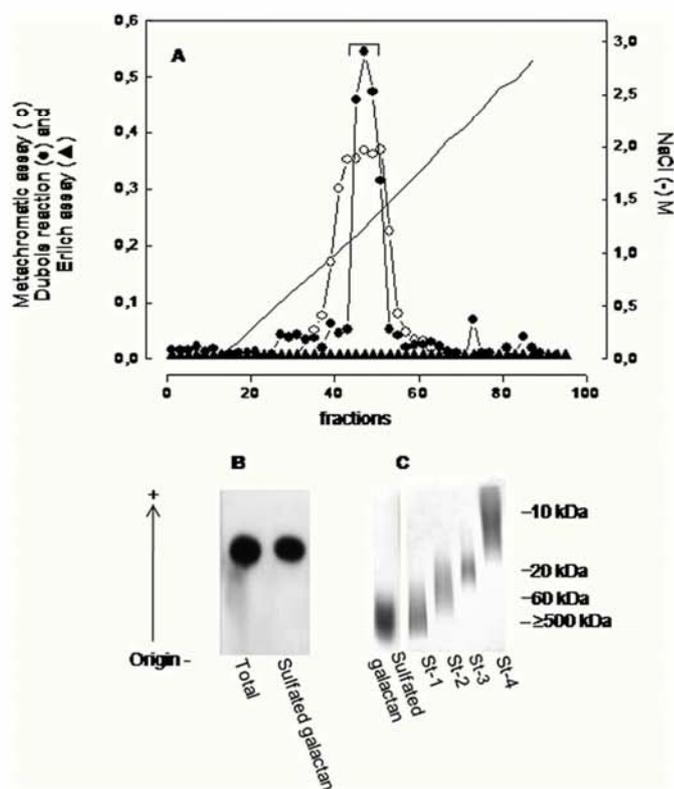
min on ice, sperm were fixed in 350  $\mu$ l of 3.7% formaldehyde in seawater for 30 min, washed 2 times with 500  $\mu$ l of phosphate-buffered saline and stained for at least 2 h with 1 unit of rhodamine phalloidin (Molecular Probes R415, Invitrogen) in 50  $\mu$ l of 0.1 M glycine, 1 mg/ml bovine serum albumin, 0.02% sodium azide in phosphate-buffered saline (pH 7.4). The solution was then washed twice in 500  $\mu$ l of phosphate-buffered saline and incubated with 30  $\mu$ l of 4'6'-diamino-2-phenylindole (Sigma) for 6 min. The cells were washed twice in 500  $\mu$ l of phosphate-buffered saline and then re-suspended in 30  $\mu$ l of 70% glyceraldehyde in the same solution and mounted in a thin layer, and the coverslip was sealed. Sperm were scored blindly using a Zeiss Axioskop 2 plus fluorescent microscope. Photos were acquired with a Zeiss LSM 510 Meta confocal microscope (Jeha, Germany) in red (phalloidin), blue (4'6'-diamino-2-phenylindole), and transmitted light channels.

**Measurements of Increases in Intracellular Ca<sup>2+</sup>**—Intracellular calcium was measured by a methodology slightly modified from Hirohashi and Vacquier (8). Briefly, sperm from *E. lucunter* were collected on ice and in darkness and immediately used. Undiluted semen was suspended in 4 volumes of dye loading buffer (artificial sea water containing 10 mM HEPES, 1 mM CaCl<sub>2</sub>, and 0.1 mg/ml soybean trypsin inhibitor at pH 7.0), placed in dimethyl sulfoxide (final concentration 0.6%) containing fura-2/AM at a final concentration of 12  $\mu$ M, and incubated for 4 h in darkness at 4 °C. The cells were washed twice with phosphate-buffered saline and grounded at  $430 \times g$  for 5 min. The final pellet was resuspended in fresh dye loading buffer without soybean trypsin inhibitor. Finally, 50  $\mu$ l of fura-2-loaded sperm were placed in an 11-mm-diameter glass tube containing 1.5 ml of dye loading buffer without soybean trypsin inhibitor and mounted in a FP-6300 spectrofluorimeter from Jasco at 16 °C at continual agitation. The fluorescence intensity was measured at Ex/Em 340/500 nm to follow the effect of the sulfated polysaccharide in the sperm. The sugar content of the polysaccharide solutions was quantified by the phenol-sulfuric acid assay (10).

## RESULTS AND DISCUSSION

**The Egg Jelly of the Sea Urchin *G. crenularis* Contains a Sulfated D-Galactan**—The polysaccharides extracted from the egg jelly of *G. crenularis* were purified by anion-exchange chromatography on a Mono Q column coupled to a fast protein liquid chromatography system, and fractions were monitored for hexose and metachromasia (Fig. 1A). We observed a single peak of sulfated polysaccharide, eluted from the column with  $\sim 1.3$  M NaCl, as indicated by the hexose assay (closed circles in Fig. 1A) and metachromasia (open circles). Different from other species of sea urchins (17, 31–33), no sialoglycan was detected in the egg jelly of this sea urchin, as indicated by the negative Erlich reaction (closed triangles in Fig. 1A).

The purified sulfated polysaccharide from *G. crenularis* showed a single component on agarose gel electrophoresis (Fig. 1B), with a high molecular mass (Fig. 1C), as already observed for the sulfated polysaccharides from the egg jelly of other species of sea urchins (17, 31–33). Chemical analysis revealed the occurrence of galactose and sulfate, exclusively. After derivatization with butyl alcohol, the galactoside derivatives obtained

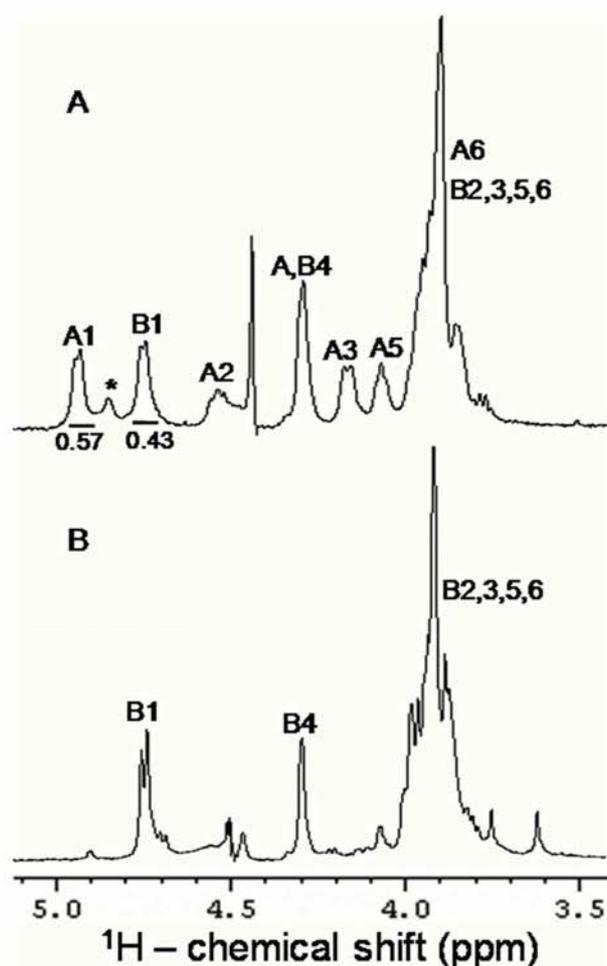


**FIGURE 1. Purification and electrophoretic mobility of the sulfated galactan from the egg jelly of *G. crenularis*.** A, total sulfated polysaccharide from the egg jelly (20 mg) was applied to a Mono Q fast protein liquid chromatography column (HR5/5) equilibrated with 20 mM Tris-HCl (pH 8.0). The column was developed by a linear gradient of 0–3.0 M NaCl in the same solution. Fractions were assayed by metachromasia using 1,9-dimethylmethylene blue (○), the Dubois reaction for hexose (●), and the Ehrlich assay for sialic acid (▲). The NaCl concentration was estimated by conductivity (–). Fractions containing the sulfated galactan (indicated by the horizontal bar) were pooled, dialyzed against distilled water, and lyophilized. B, total polysaccharides from the egg jelly and the purified sulfated galactan (15  $\mu$ g of each) were applied to a 0.5% agarose gel, and electrophoresis was carried out for 1 h at 110 V in 0.05 M 1,3-diaminopropane:acetate (pH 9.0). Gels were fixed with 0.1% *N*-cetyl-*N,N,N*-trimethylammonium bromide solution. After 12 h, the gel was dried and stained with 0.1% toluidine blue in acetic acid/ethanol/water (0.1:1:5, v/v). C, the purified sulfated galactan (10  $\mu$ g) were run on 6% polyacrylamide gels in 0.2 M sodium barbital (pH 8.6) and stained with 0.1% toluidine blue in 1% acetic acid. Molecular mass markers used were dextran sulfate (St-1, average molecular mass  $\geq$ 500 kDa), chondroitin 6-sulfate (St-2, average molecular mass  $\sim$ 60 kDa), chondroitin 4-sulfate (St-3, average molecular mass  $\sim$ 40 kDa), and heparin (St-4, average molecular mass  $\sim$ 10 kDa).

showed the same retention times and peak areas as D-galactose standard, indicating that the galactose occurs in the *G. crenularis* exclusively in the D-enantiomeric form.

Thus, the egg jelly of the sea urchin *G. crenularis* contains a single fraction of sulfated D-galactan. It differs from the sulfated polysaccharides found in other species of sea urchins that contain fucose or galactose always in the L-configuration (17, 31–33).

*The Sulfated Galactan from G. crenularis Has a Regular Disaccharide Repeating Unit Composed of Alternating 2-Sulfated and Non-sulfated 3-Linked  $\beta$ -Galactopyranosyl Units*—For a detailed structural analysis of the sulfated galactan from *G. crenularis*, we employed one- and two-dimensional NMR spectroscopy of the native polysaccharide and of its desulfated derivative (Figs. 2–4).



**FIGURE 2. One-dimensional  $^1\text{H}$  NMR spectra at 400 MHz of the native sulfated  $\beta$ -galactan from *G. crenularis* (A) and its desulfated derivative (B).** About 5 mg of each polysaccharide were dissolved in 0.5 ml of  $\text{D}_2\text{O}$ , and the one-dimensional NMR spectra were recorded at 50  $^\circ\text{C}$ . The residual water signal was suppressed by presaturation.  $^1\text{H}$  chemical shifts are relative to external trimethylsilylpropionic acid at 0 ppm. The signals designated with A and B correspond to the 2-sulfated and non-sulfated  $\beta$ -D-galactopyranosyl units, respectively, and numbers correspond to each hydrogen of the hexose ring. Numbers in panel A below the spectrum indicate integrals of A1 and B1 signals. The peak marked by the asterisk corresponds to a contaminant.

The  $^1\text{H}$  one-dimensional spectrum (Fig. 2A) and mainly the  $^{13}\text{C}$ ,  $^1\text{H}$  HSQC spectrum (Fig. 3A) of the native sulfated galactan showed equimolar proportions of two anomeric signals, denominated as A1 and B1. A1 exhibited a downfield shift of  $\sim$ 0.2 ppm, possibly because of sulfation (34). This conclusion was reinforced by analysis of the  $^{13}\text{C}$ ,  $^1\text{H}$  HSQC spectrum of the native polysaccharide (Fig. 3A) and the disappearance of signal A1 after the desulfation reaction (Figs. 2B and 3B). These observations suggest that the sulfated galactan from *G. crenularis* contains equimolar proportions of non-sulfated and sulfated galactoses.

Correlation peaks in the two-dimensional COSY (Figs. 4, A and D) and TOCSY spectra (Figs. 4, B and E) allowed us to trace the entire spin systems of A and B signals in the native sulfated galactan and of B signal in the desulfated derivative. The COSY spectrum of the native sulfated galactan showed two correlation signals (A1–2 and B1–2) because of scalar-couplings between  $^1\text{H}$ -anomeric proton and H2 for both residues (Fig.

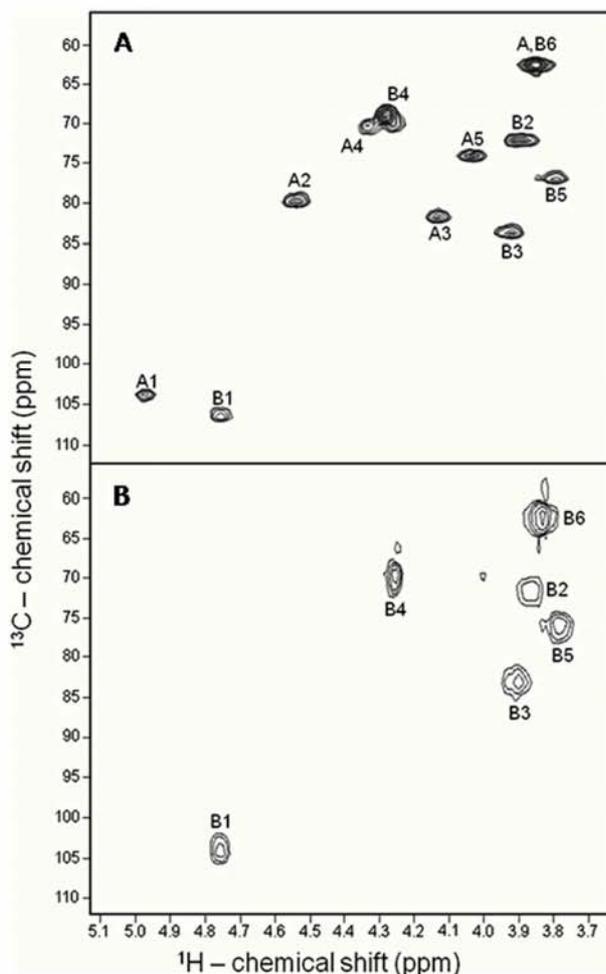
Sulfated  $\beta$ -Galactan from Sea Urchin

FIGURE 3.  $^{13}\text{C}$ ,  $^1\text{H}$  HSQC spectra of the sulfated  $\beta$ -galactan from *G. crenularis* (A) and its desulfated derivative (B). Chemical shifts are relative to external trimethylsilylpropionic acid at 0 ppm for  $^1\text{H}$  and methanol for  $^{13}\text{C}$ . The spin systems were denoted as A and B for 2-sulfated and non-sulfated  $\beta$ -galactopyranosyl units, respectively. Spectra were recorded at 50  $^\circ\text{C}$ .

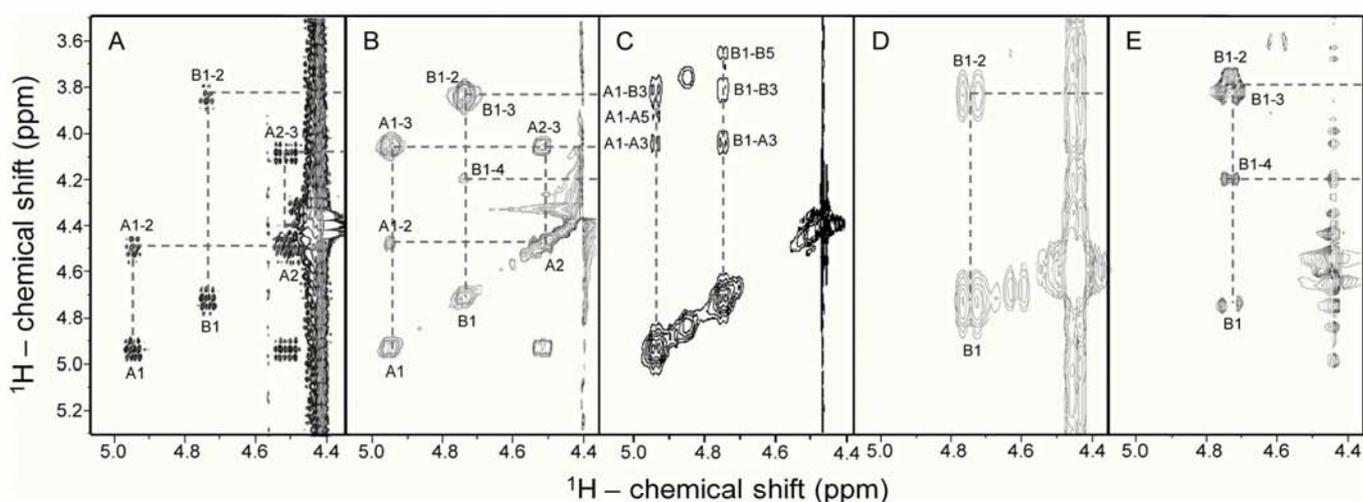


FIGURE 4. Strips of the anomeric regions (expansions from 4.4 to 5.1 ppm) from the COSY (A and D), TOCSY (B and E), and NOESY (C) spectra of the sulfated  $\beta$ -galactan (A–C) from *G. crenularis* and its desulfated derivative (D and E). About 5 mg of each polysaccharide were dissolved in 0.5 ml of  $\text{D}_2\text{O}$ , and the two-dimensional NMR spectra were recorded at 50  $^\circ\text{C}$  at 400 MHz. Chemical shifts are relative to external trimethylsilylpropionic acid at 0 ppm for  $^1\text{H}$ . The spin systems were denoted as A and B for 2-sulfated and non-sulfated  $\beta$ -galactopyranosyl units, respectively. The peak at 4.85/3.85 ppm in the NOESY spectrum (C) corresponds to a contaminant.

4A). A single B1–2 correlation signal is observed in the spectrum of the desulfated derivative (Fig. 4D). After identification of the cross-peaks in the COSY spectra, the TOCSY experiments (Figs. 4, B and E) allowed up to assign unequivocally the whole spin systems A and B and to obtain the  $^1\text{H}$  chemical shifts, as indicated in Table 1. Based on the  $^1\text{H}$  chemical shifts, we assigned the peaks of correlation in the  $^{13}\text{C}$ ,  $^1\text{H}$  HSQC spectra (Fig. 3) and obtained the values of  $^{13}\text{C}$  chemical shifts shown in Table 1. Analysis of these  $^{13}\text{C}$  chemical shifts revealed that the galactan contains 3-linked  $\beta$ -galactopyranosyl residues, as indicated by the typical low-field shift of carbon ( $\sim 10$  ppm) in sites of glycosylation (Table 1). This  $^{13}\text{C}$  shift was also seen in reference compounds (Table 1).

The spin systems A and B traced for the native sulfated galactan differ mainly because of typical  $^1\text{H}$  low-field shift ( $\sim 0.6$  ppm) of H2 that indicates 2-sulfation (Table 1). The neighbors H1 and H3 protons showed characteristic  $\sim 0.2$ -ppm low-field shifts. The spin systems B traced for native sulfated galactan and for its desulfated derivative show almost similar chemical shifts for  $^1\text{H}$  and  $^{13}\text{C}$ .

In summary, these results indicate that sulfated galactan from *G. crenularis* has equimolar proportions of 2-sulfated and non-sulfated  $\beta$ -galactopyranosyl units. It remains to clarify whether these units alternate along the polysaccharide chain or occur in a random distribution or even as clusters in the molecule. This aspect was investigated using two-dimensional NOESY spectrum of the native sulfated galactan (Fig. 4C). The spectrum revealed intraresidue NOEs between  $^1\text{H}$ -anomeric and H3 and H5, as typically found in  $\beta$ -galactopyranose residues in their regular chair conformation (35). Protons H1, H3, and H5 are on the same plane (equatorially or axially) and commonly show  $< 5$  Å space contacts, which allows detection of their NOE. But more significantly, the NOESY spectrum also showed inter-residue NOEs between H1 and H3 of the adjacent residue (A1–B3 and B1–A3). These observations indicated that the 2-sulfated and non-sulfated units intercalate along the polysaccharide chain as a repeating disaccharide

TABLE 1

<sup>1</sup>H and <sup>13</sup>C chemical shifts (ppm) of the sulfated and desulfated  $\beta$ -D-galactan from *G. crenularis* and standard compounds

ND, not determined.

Polysaccharide	Structure	<sup>1</sup> H and <sup>13</sup> C chemical shift <sup>a</sup>					
		H-1 C-1	H-2 C-2	H-3 C-3	H-4 C-4	H-5 C-5	H-6 C-6
Sulfated $\beta$ -D-galactan from <i>G. crenularis</i>	3- $\beta$ -D-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1 (unit A)	4.94	<b>4.52</b>	<i>4.11</i>	4.37	4.02	3.82
	3- $\beta$ -D-Galp-1 (unit B)	104.1	<b>80.2</b>	<i>81.8</i>	70.5	74.3	62.5
Desulfated $\beta$ -D-galactan from <i>G. crenularis</i>	3- $\beta$ -D-Galp-1 (unit B)	4.73	3.87	<i>3.90</i>	4.24	3.75	3.82
	3- $\beta$ -D-Galp-1 (unit B)	107.2	72.0	<i>83.5</i>	69.1	77.0	62.5
Desulfated $\beta$ -D-galactan from <i>Codium isthmocladum</i> <sup>b</sup>	3- $\beta$ -D-Galp-1	4.79	3.90	<i>3.93</i>	4.29	3.82	3.87
	3- $\beta$ -D-Galp-1	103.5	69.9	<i>81.9</i>	68.3	74.8	60.5
Sulfated $\beta$ -D-galactan from <i>Meretrix petechialis</i> <sup>c</sup>	3- $\beta$ -D-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1	4.81	3.64	<i>3.92</i>	4.39	3.86	3.94-3.85
	3- $\beta$ -D-Galp-1	102.6	73.8	<i>82.9</i>	67.1	75.2	60.3
Sulfated $\alpha$ -L-galactan from <i>E. lucunter</i> <sup>d</sup>	3- $\alpha$ -L-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1	4.62	3.82	<i>3.89</i>	4.38	4.09	4.36-4.01
	3- $\alpha$ -L-Galp-1	103.1	70.1	<i>70.2</i>	67.0	73.9	69.9
Desulfated $\alpha$ -L-galactan from <i>E. lucunter</i> <sup>d</sup>	3- $\alpha$ -L-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1	4.83	<b>4.45</b>	<i>3.97</i>	ND	ND	ND
	3- $\alpha$ -L-Galp-1	5.47	<b>4.66</b>	<i>4.23</i>	4.33	4.35	3.82
Desulfated $\alpha$ -L-galactan from <i>E. lucunter</i> <sup>d</sup>	3- $\alpha$ -L-Galp-1	97.2	<b>76.2</b>	<i>75.9</i>	72.5	69.5	63.8
	3- $\alpha$ -L-Galp-1	5.26	4.08	<i>4.14</i>	4.32	4.24	3.82
		98.1	73.5	<i>77.2</i>	69.5	68.5	63.9

<sup>a</sup> Chemical shifts are relative to external trimethylsilylpropionic acid at 0 ppm for <sup>1</sup>H and methanol for <sup>13</sup>C. Values in boldface indicate sulfate position and in italics indicate glycosylated positions.

<sup>b</sup> Values are from Farias *et al.* (43).

<sup>c</sup> Data are from Amornrut *et al.* (38).

<sup>d</sup> Values are from Alves *et al.* (17).

unit. Overall, NMR analysis indicated that the sulfated galactan from the egg jelly of *G. crenularis* has the following disaccharide repeating units: (3- $\beta$ -D-Galp-2(OSO<sub>3</sub><sup>-</sup>)-1)→3- $\beta$ -D-Galp-1)<sub>n</sub> (Fig. 5A).

*The Anomeric Configuration of the Glycosidic Linkage Is an Important Structural Feature for Recognition of Sulfated Polysaccharides by Sea Urchin Sperm*—Previously, we tested a variety of sulfated polysaccharides with well defined structures as inducers of the acrosome reaction in sea urchin sperm. These studies indicated that sulfated polysaccharides show species specificity in inducing the sperm acrosome reaction, which is regulated by the structure of the saccharide chain and its sulfation pattern (3, 31–33). However, the glycosidic linkages in all these previously tested polysaccharides were in the  $\alpha$ -anomeric configuration. The sulfated  $\beta$ -D-galactan from *G. crenularis* extends the possibility of determining the event at a further molecular detail.

Investigation of the acrosome reaction using sperm from *G. crenularis* was not feasible. There was no obvious difference in morphology and phalloidin staining between sperm treated with and without sulfated  $\beta$ -D-galactan or 10  $\mu$ M ionomycin. Thus, the acrosome in this species is too small and difficult to visualize using microscopic methods. Attempts to determine the increase in intracellular Ca<sup>2+</sup> using fura-2 loaded sperm were also unsuccessful. As an alternative, we used a fertilization inhibition assay. Preincubation of the *G. crenularis* sperm with homologous sulfated  $\beta$ -galactan blocks fertilization, possibly because of a premature acrosome reaction (see Table 2). In contrast, preincubation with the sulfated  $\alpha$ -galactan from *E. lucunter*, composed of (3- $\alpha$ -L-Galp-2(OSO<sub>3</sub><sup>-</sup>)-1)<sub>n</sub> (Fig. 5B) (17), did not affect the fertilization capacity of the *G. crenularis* sperm. Clearly, these results indicate that the effect of sulfated galactans on the sea urchin fertilization depends mostly on the anomeric configuration of the glycosidic linkage rather than charge density.

We further investigate the effect of sulfated galactans on sea urchin fertilization using sperm from *E. lucunter*, which

express in their egg jelly a sulfated  $\alpha$ -galactan (Fig. 5B) (17). Sperm from *E. lucunter* were equally sensitive to homologous 2-sulfated  $\alpha$ -galactan and to heterologous 2-sulfated  $\alpha$ -fucan from *S. franciscanus* (see structure in Fig. 5C) but not to the 2-sulfated  $\beta$ -galactan from *G. crenularis*, even when the polysaccharide was tested at high concentrations (Fig. 6A). This indicates that the receptor for egg jelly of *E. lucunter* sperm does not differentiate between the CH<sub>2</sub>OH of L-galactose and CH<sub>3</sub> of L-fucose at position 6. These two polysaccharides present the same sulfation pattern and position of glycosylation but differ in the sugar moieties. In a similar way sperm from *S. franciscanus* were sensitive to the homologous sulfated  $\alpha$ -fucan and to the heterologous *E. lucunter*-sulfated  $\alpha$ -galactan (6, 36). Despite that, *E. lucunter* sperm markedly distinct between  $\alpha$ - and  $\beta$ -galactans.

Another plausible reason for the absence of effect of the sulfated  $\beta$ -galactan on *E. lucunter* sperm is its reduced charge density compared with the homologous polysaccharide (0.5 versus 1.0 sulfate/monosaccharide). We attempted to investigate this aspect using chemically over-sulfated galactan. In the 2-sulfated, 3-linked  $\alpha$ -galactan from *E. lucunter*, the 4- and 6-positions are the only ones capable of additional sulfation. Over-sulfation of this polysaccharide did not change its responsiveness to homologous sperm (Fig. 6B), indicating that increased sulfates do not increase or inhibit the biological activity. In contrast, over-sulfated  $\beta$ -galactan from *G. crenularis* induced acrosome reaction in *E. lucunter* sperm but at a significantly lower potency compared with the homologous  $\alpha$ -galactan. Thus, the anomeric configuration of the glycosidic linkage is still a preferential structural requirement for the effect of these over-sulfated galactans on the acrosome reaction.

Another point to consider is that the sulfated  $\beta$ -galactan could increase intracellular Ca<sup>2+</sup> of *E. lucunter* sperm but not up to the proper high concentration required to induce the complete acrosome reaction. In fact, a low molecular weight sulfated  $\alpha$ -fucan increases intracellular Ca<sup>2+</sup> and pH, which is enough to induce exocytosis of the acrosome vesicle, but only at

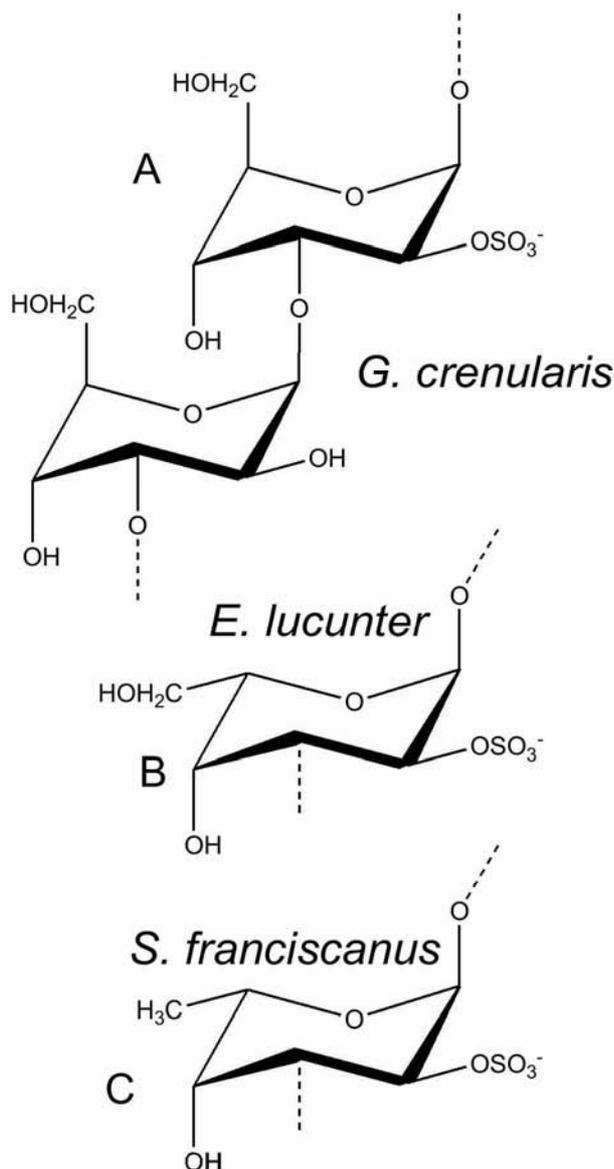
Sulfated  $\beta$ -Galactan from Sea Urchin

FIGURE 5. Structures of 2-sulfated polysaccharides obtained from the egg jelly of different species of sea urchins. A, the sulfated  $\beta$ -D-galactan from *G. crenularis* is composed of the following disaccharide repeating unit: (3- $\beta$ -D-Galp-2(SO<sub>4</sub>)-(1 $\rightarrow$ 3)- $\beta$ -D-Galp-1)<sub>n</sub>. The polysaccharides from *E. lucunter* (B) and *S. franciscanus* (C) are composed of  $\alpha$ -L-galactopyranosyl or  $\alpha$ -L-fucopyranosyl residues, respectively, both 2-sulfated and 3-linked.

**TABLE 2**  
Fertilization block by pre-incubation of *G. crenularis* sperm with homologous and heterologous sulfated galactans

*G. crenularis* sperm were preincubated with sulfated galactan for 10 min. Thereafter, such sperm were inseminated for 30 min with homologous eggs. The percentages of fertilization (indicated by the presence of fertilization envelope) were then scored under microscopy by counting ~200 eggs.

	Fertilization envelope formed			
	Sulfated $\alpha$ -galactan from <i>E. lucunter</i>		Sulfated $\beta$ -galactan from <i>G. crenularis</i>	
	0	4 mg/ml	0	4 mg/ml
	% of total eggs			
Experiment 1	98.7	93.0	98.5	3.3
Experiment 2	99.1	96.7	98.0	0.5
Experiment 3	98.2	98.1	98.3	0.5
Mean $\pm$ S.D.	98.7 $\pm$ 0.5	95.9 $\pm$ 2.6	98.3 $\pm$ 0.3	1.4 $\pm$ 0.9

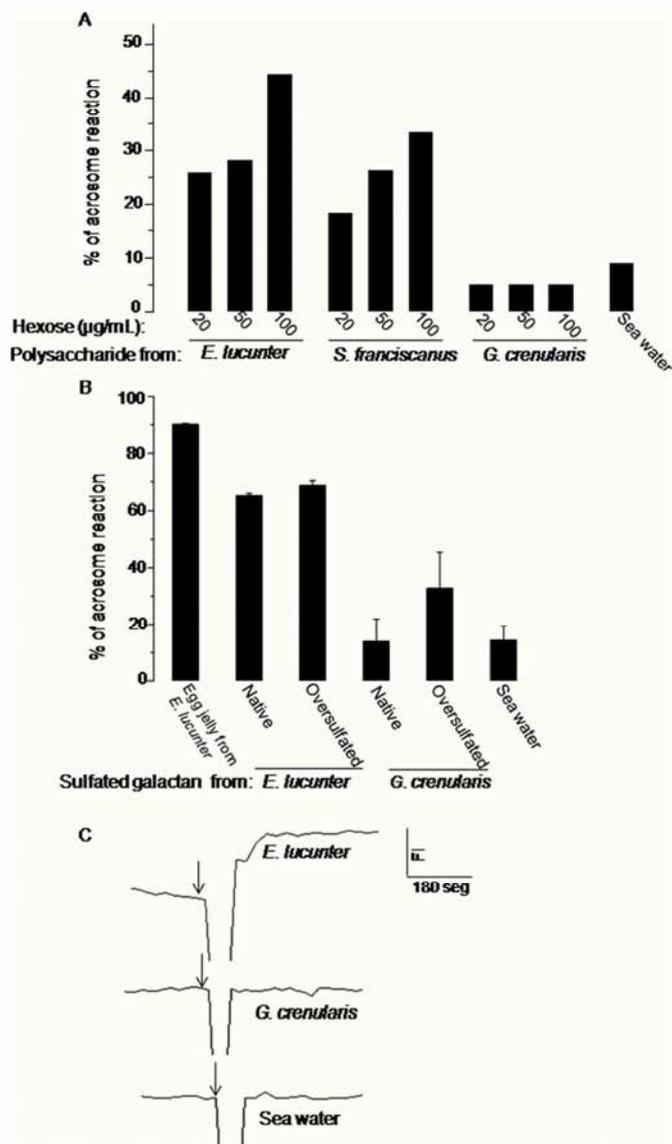


FIGURE 6. Effects of native (A) or over-sulfated (B) 2-sulfated  $\alpha$ -L-galactan,  $\alpha$ -L-fucan, and  $\beta$ -D-galactan from the egg jellies of *E. lucunter*, *S. franciscanus*, and *G. crenularis*, respectively, as inducers of acrosome reaction (A and B) and on increases intracellular Ca<sup>2+</sup> (C) of *E. lucunter* sperm. Sulfated polysaccharides or lyophilized egg jelly were dissolved in sea water and incubated with sperm from *E. lucunter*, and the acrosome reaction was detected using fluorescence phalloidin (see "Experimental Procedures"). Negative control was done with artificial sea water. Approximately 100–150 sperm were scored per data point. The concentrations of polysaccharides were normalized by hexose content. In panel B a fixed concentration of sulfated polysaccharides (100  $\mu$ g/ml) was used in the assays. The results are representative from two experiments and are shown as the average and S.D. C, increases in intracellular Ca<sup>2+</sup>: each sulfated polysaccharide were added to a fura-2 loaded *E. lucunter* sperm suspension (arrows) at a final concentration of 100  $\mu$ g/ml. Scale bars indicate the fluorescence intensity (Fl, y axis) and 100 s (x axis).

a slower kinetics, which is not able to induce the complete acrosome reaction (8, 37). To test for a similar effect of the sulfated  $\beta$ -galactan in *E. lucunter* sperm, we measured the increase in intracellular Ca<sup>2+</sup> using fura-2-loaded sperm after incubation with egg jelly polysaccharides (Fig. 6C). Increase of intracellular Ca<sup>2+</sup> was detected when *E. lucunter* sperm were incubated with homologous polysaccharides but not with the sulfated  $\beta$ -galactan from *G. crenularis* (Fig. 6C).

**TABLE 3**  
 $^3J_{\text{H-H}}$  and  $^1J_{\text{C-H}}$  (Hz) in the NMR spectra of sulfated  $\alpha$ - or  $\beta$ -galactans and  $\alpha$ -fucans  
 ND, not determined.

Polysaccharide	Structure	$^3J_{\text{H-H}}$				
		H1-H2	H2-H3	H3-H4	H4-H5	H5-H6
Sulfated $\beta$ -D-galactan from <i>G. crenularis</i>	3- $\beta$ -D-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1 (unit A)	7.4	7.9	2.7	ND	ND
	3- $\beta$ -D-Galp-1 (unit B)	7.01	ND	ND	ND	ND
Desulfated $\beta$ -D-galactan from <i>G. crenularis</i>	3- $\beta$ -D-Galp-1 (unit B)	7.29	ND	2.83	ND	ND
Sulfated $\beta$ -D-galactan from <i>M. petechialis</i> <sup>a</sup>	3- $\beta$ -D-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1	7.6-7.8	7.8-8.2	<1.5	ND	ND
Sulfated $\alpha$ -L-galactan from <i>E. lucunter</i> <sup>b</sup>	3- $\alpha$ -L-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1	3.0	10.5	ND	ND	4.5
Sulfated $\alpha$ -L-fucan from <i>S. franciscanus</i> <sup>b</sup>	3- $\alpha$ -L-Fucp-2(SO <sub>3</sub> <sup>-</sup> )-1	3.1	9.5	ND	ND	4.9
Sulfated $\alpha$ -L-fucan from <i>Ludwigothurea grisea</i> <sup>c</sup>	3- $\alpha$ -L-Fucp-2(SO <sub>3</sub> <sup>-</sup> )-1	3.5-4.0	≈10.0	<3.0	<3.0	6.7-7.0

		$^1J_{\text{C-H}}$					
		C1-H1	C2-H2	C3-H3	C4-H4	C5-H5	C6-H6
Hz							
Sulfated $\beta$ -D-galactan from <i>G. crenularis</i>	3- $\beta$ -D-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1 (unit A)	164.7	156.8	143.4	151.4	143.0	143.9
	3- $\beta$ -D-Galp-1 (unit B)	163.5	141.2	143.1	150.5	139.9	143.9
Sulfated $\alpha$ -L-galactan from <i>E. lucunter</i> <sup>b</sup>	3- $\alpha$ -L-Galp-2(SO <sub>3</sub> <sup>-</sup> )-1	177.5	151.3	146.2	146.5	146.8	143.3
Sulfated $\alpha$ -L-fucan from <i>S. franciscanus</i> <sup>b</sup>	3- $\alpha$ -L-Fucp-2(SO <sub>3</sub> <sup>-</sup> )-1	177.2	148.7	144.6	140.9	147.1	125.3

<sup>a</sup> Data are from Amornrut *et al.* (38).

<sup>b</sup> See also Pereira *et al.* (39).

<sup>c</sup> Data are from Mulloy *et al.* (44).

In conclusion, studies with the sulfated  $\beta$ -galactan from *G. crenularis* extend the characterization of the induction of sea urchin acrosome reaction to a further molecular detail. It indicates that the echinoderm sperm are sensitive to polysaccharides with the appropriate anomeric configuration.

**The Scalar Coupling Constants  $^3J_{\text{H-H}}$  and  $^1J_{\text{C-H}}$  Differ between Sulfated  $\beta$ -Galactans and Sulfated  $\alpha$ -Galactan or  $\alpha$ -Fucan**—Conformational analysis is an important approach to extend the characterization of the biological effect of the sea urchin polysaccharides at molecular level. The differences in chemical structure may in fact determine spacing between sulfate groups required to match the interval between basic amino acid residues in the protein chain of the receptors.

This aspect was investigated by determining the scalar coupling constant of the 2-sulfated polysaccharides isolated from three species of sea urchins. The sulfated  $\beta$ -galactan from *G. crenularis* showed well defined doublets of  $^1\text{H}$ - $^1\text{H}$  couplings for the anomeric signals in the one-dimensional NMR spectra (Figs. 2, A and B) (Table 3). This pattern of  $^1\text{H}$ - $^1\text{H}$  coupling is evidenced by multiplets in all cross-peaks showed in the COSY spectrum (Fig. 4A) and by the doublets in the other homonuclear two-dimensional experiments (TOCSY and NOESY spectra, Fig. 4, B and C). Interesting, a similar pattern of coupled signals was also observed for another invertebrate sulfated  $\beta$ -galactans (38) but poorly noted in sulfated  $\alpha$ -galactans and in a sulfated  $\alpha$ -fucan (17, 39). These marked spin-spin couplings of the sulfated  $\beta$ -galactan from *G. crenularis* together with the presence of inter-residue NOEs exclusively between protons involved in the glycosidic bond, suggest a polysaccharide with dynamic behavior and the absence of a single preponderant conformation. This proposition is confirmed by the MD simulations, as discussed below.

The scalar-coupling constants  $^3J_{\text{H-H}}$  and  $^1J_{\text{C-H}}$  observed for the sulfated  $\beta$ -galactan differ significantly compared with the values for sulfated  $\alpha$ -galactan and  $\alpha$ -fucan, especially  $^3J_{\text{H-H}}$  (Table 3). Again, these different coupling-constant values reflect distinct conformations for these polysaccharides. The 2-sulfated and non-sulfated units found in the sulfated  $\beta$ -galac-

tan from *G. crenularis* showed similar  $^3J_{\text{H-H}}$  and  $^1J_{\text{C-H}}$  values, with only discrete difference in the  $^1J_{\text{C-H}}$  values (Table 3). It means that 2-sulfation does not significantly alter the glycosidic geometry of the  $\beta$ -galactopyranosyl residues.

**MD of  $\alpha$ - and  $\beta$ -Galactans**—Differences in the scalar coupling constants between sulfated  $\beta$ -galactan and sulfated  $\alpha$ -galactan or  $\alpha$ -fucan suggest distinct conformations. To explore such behavior at the atomic level, we employed conformational calculations based on MD simulations, an important tool for the structural and conformational characterization of carbohydrates (40).

Initially, the conformational preference of the disaccharides  $\alpha$ -L-Galp-(1 $\rightarrow$ 3)- $\alpha$ -L-Galp,  $\alpha$ -L-Galp-2(SO<sub>4</sub>)-(1 $\rightarrow$ 3)- $\alpha$ -L-Galp-2(SO<sub>4</sub>),  $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\beta$ -D-Galp,  $\beta$ -D-Galp-2(SO<sub>4</sub>)-(1 $\rightarrow$ 3)- $\beta$ -D-Galp, and  $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\beta$ -D-Galp-2(SO<sub>4</sub>) was analyzed employing relaxed contour plots (Fig. 7). These maps indicate that the non-sulfated galactans possess more flexible conformational behavior than their sulfated counterparts (Fig. 7, A versus B and C versus D and E), indicating that sulfate groups increase the rigidity of the polysaccharide chain, as already suggested (41). Additionally, the configuration of the glycosidic linkage appears to play an important role on the flexibility of sulfated galactans and fucans in aqueous solution, as the  $\beta$ -D-galactopyranosyl residues showed an increased flexibility when compared with  $\alpha$ -L-galactopyranosyl units (Fig. 7, A and B versus C-E), which did not differ from  $\alpha$ -L-fucopyranose (41).

To refine the analysis obtained by relaxed contour plots data upon the addition of solvent molecules, each galactan minimum energy conformation was further submitted to a 0.1- $\mu$ s MD simulation in aqueous solutions. The analysis of these simulations confirmed the observations that  $\alpha$ -L-galactopyranosyl units have a more rigid structure than  $\beta$ -D-galactopyranoses. Thus, the  $\alpha$ -L-galactopyranose disaccharides, as in the sulfated  $\alpha$ -galactan from *E. lucunter*, present unique prevalent conformation in solution (glycosidic dihedral angles of  $\Phi/\Psi = -102^\circ/131^\circ$ , Fig. 7B). In contrast,  $\beta$ -D-galactopyranosyl disaccharides, as in the sulfated  $\beta$ -galactan from *G. crenularis*, present at least two main solution conformations for each dihedral

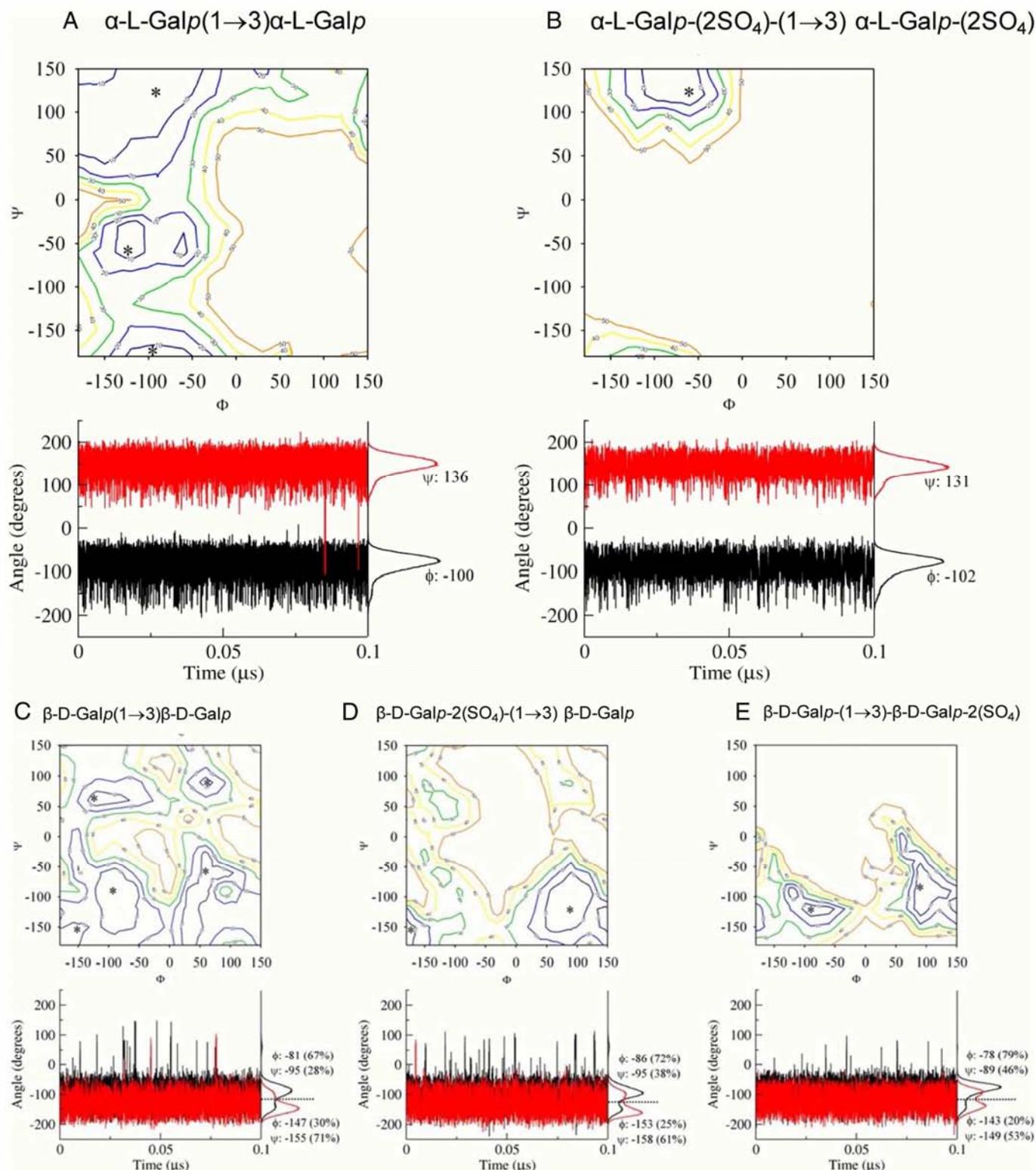
Sulfated  $\beta$ -Galactan from Sea Urchin

FIGURE 7. MD of galactans composed of the following disaccharide units: **A**,  $\alpha$ -L-Galp-(1 $\rightarrow$ 3)- $\alpha$ -L-Galp; **B**,  $\alpha$ -L-Galp-2(SO<sub>4</sub>)-(1 $\rightarrow$ 3)- $\alpha$ -L-Galp-2(SO<sub>4</sub>); **C**,  $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\beta$ -D-Galp; **D**,  $\beta$ -D-Galp-2(SO<sub>4</sub>)-(1 $\rightarrow$ 3)- $\beta$ -D-Galp; **E**,  $\beta$ -D-Galp-(1 $\rightarrow$ 3)- $\beta$ -D-Galp-2(SO<sub>4</sub>). The contour plots are shown at every 10 kJ·mol<sup>-1</sup> from 10 to 50 kJ·mol<sup>-1</sup>. Asterisks (\*) indicate the input minimum energy conformations for MD refinement. The fluctuation and distribution of the  $\phi$  (black) and (red) dihedral angles are also presented together with average and prevalence values (%) for each conformer population. The S.D. for all average values lies in  $\sim 20^\circ$ .

angle (Figs. 7, C–E), indicating four possible conformations co-existing simultaneously in solution. As a consequence, the glycosidic linkage around the sulfated  $\beta$ -galactan has no

prevalent conformation in aqueous solutions, and a series of conformational sub-states occur in equilibrium. Curiously, the conformers experimented by the  $\beta$ -configuration do not

match the preponderant conformer observed for the  $\alpha$ -forms.

MD simulations of the desulfated derivatives from the two enantiomeric forms of the galactans (Figs. 7, *A versus C*) indicated that  $\alpha$ -galactopyranose still has less dynamic molecular behavior compared with  $\beta$ -galactopyranose. Although sulfate groups may promote steric/electrostatic hindrances, which restrict the glycosidic motions, the three-dimensional orders of these sulfated galactans are clearly dominated by the configuration of their glycosidic linkage.

Thus, the sulfated  $\beta$ -galactan from *G. crenularis* is a more dynamic and flexible polysaccharide than the sulfated  $\alpha$ -galactan from *E. lucunter* or the sulfated  $\alpha$ -fucan from *S. franciscanus*. But more significantly, the observation that only  $\alpha$ -polysaccharides induce the acrosome reaction in sperm from *E. lucunter* and *S. franciscanus* can arise from the observation that the preponderant conformer population observed by the active sulfated  $\alpha$ -polysaccharides is not observed among populations of the  $\beta$ -form despite its much more flexible structure in aqueous solution.

**Major Conclusions**—We extended our studies to the sea urchin *G. crenularis*, which inhabits high depth and low temperature seawater (7). The egg jelly of this sea urchin contains a sulfated  $\beta$ -galactan, which is constituted of the disaccharide repeating structure 3- $\beta$ -D-Galp-2(OSO<sub>3</sub>)-1 $\rightarrow$ 3- $\beta$ -D-Galp-1. This is the first report of a sulfated  $\beta$ -galactan with a regular and homogeneous disaccharide structure. The polymer is markedly distinct from all other sea urchin-sulfated polysaccharides described so far that are composed of units on  $\alpha$ -L-configuration. Furthermore, this sea urchin does not contain sialoglycans, commonly found in the echinoderm egg jelly.

Sulfated  $\beta$ -galactans have been reported in marine green algae. In these organisms the polysaccharides have complex structures composed preponderantly of 4-sulfated, 3-linked  $\beta$ -D-galactopyranosyl units but with branching and mostly highly pyruvylated at the non-reducing terminal residues, forming cyclic ketals (42, 43). Red algae contain a linear sulfated galactan made of alternating 3-linked  $\beta$ -D-galactopyranosyl and 4-linked  $\alpha$ -galactopyranosyl residues, but considerable structural variation in these sulfated galactans occurs among different species, including complex sulfation pattern, substitution by methyl groups or pyruvic acid, formation of anhydro sugar, etc. (45).

We used this new sulfated  $\beta$ -galactan to investigate the acrosome reaction in a further molecular detail using homologous and heterologous sperm. We tested three 2-sulfated polysaccharides differing in their conformation ( $\alpha$  or  $\beta$ ) and monosaccharide composition (galactose or fucose) as inducers of the sperm acrosome reaction. We aimed to establish the structure *versus* biological activity of the echinoderm polysaccharides, including structural features at a conformational level. The sperm from *G. crenularis* react to the homologous sulfated  $\beta$ -galactan but not to the sulfated  $\alpha$ -galactan from *E. lucunter*. The species specificity was confirmed as *E. lucunter* sperm react to the homologous sulfated  $\alpha$ -galactan and also to a 2-sulfated  $\alpha$ -fucan but not to the sulfated  $\beta$ -galactan from *G. crenularis*. In a similar way, sperm from *S. franciscanus* react equally to 2-sulfated  $\alpha$ -fucan and  $\alpha$ -galactan (6).

MD and NMR data of  $^3J_{\text{H-H}}$  and  $^1J_{\text{C-H}}$  strongly indicated that  $\alpha$ - and  $\beta$ -isomers of sulfated galactan and fucan exhibit distinct conformational preferences. In particular, the preponderant conformer population experimented by the active sulfated  $\alpha$ -galactan or  $\alpha$ -fucan ( $\Phi/\Psi = -102^\circ/131^\circ$ ) is not observed among populations of the  $\beta$ -form, despite its much more flexible structure in solution. Thus, the anomeric configuration of the glycosidic linkage rather than monosaccharide composition (galactose or fucose) is the main structural requirement to induce the acrosome reaction in *G. crenularis*, *E. lucunter*, and possibly in *S. franciscanus* sperm. Our hypothesis is that sulfated  $\beta$ -galactan from *G. crenularis*, besides being a more flexible structure compared with the  $\alpha$ -polysaccharides, cannot assume the precise conformation necessary for recognition by the *E. lucunter* and *S. franciscanus* sperm. In an opposite way, the conformers found in sulfated  $\alpha$ -galactan are not recognized by *G. crenularis* sperm. In conclusion, our results extend the observation about the structural stringency of the sea urchin polysaccharides as inducers of the sperm acrosome reaction to a conformational level.

**Acknowledgments**—We thank Dr. Ana-Paula Valente for help on the NMR experiments at the Bruker 400 MHz and Adriana A. Piquet for technical assistance. We also thank Dr. Keiichiro Kyoizuka and Masahiko Washio (Research Center for Marine Biology, Asamushi, Tohoku University, Japan) for collection of sea urchin *G. crenularis*.

## REFERENCES

- Hirohashi, N., and Vacquier, V. D. (2002) *J. Biol. Chem.* **277**, 8041–8047
- Vacquier, V. D., and Moy, G. W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2456–2460
- Biermann, C. H., Marks, J. A., Vilela-Silva, A. C., Castro, M. O., and Mourão, P. A. S. (2004) *Evol. Dev.* **6**, 353–361
- Vilela-Silva, A. C., Hirohashi, N., and Mourão, P. A. (2008) *Int. J. Dev. Biol.* **52**, 551–559
- Mourão, P. A. S. (2007) *Braz. J. Med. Biol. Res.* **40**, 5–17
- Hirohashi, N., Vilela-Silva, A. C., Mourão, P. A., and Vacquier, V. D. (2002) *Biochem. Biophys. Res. Commun.* **298**, 403–407
- Hirai, E. (1963) *Sci. Rep. Tôhoku Univ. Ser. IV (Biol.)* **29**, 369–375
- Hirohashi, N., and Vacquier, V. D. (2002) *J. Biol. Chem.* **277**, 1182–1189
- Albano, R. M., and Mourão, P. A. S. (1986) *J. Biol. Chem.* **261**, 758–765
- Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. (1956) *Anal. Chem.* **28**, 350–356
- Kabat, E. A., and Mayer, M. M. (1971) in *Experimental Immunochemistry*, pp. 505–513, Charles C. Thomas Publisher, Springfield, IL
- Farnedale, R. W., Buttle, D. J., and Barrett, A. J. (1986) *Biochim. Biophys. Acta* **883**, 173–177
- Saito, H., Yamagata, T., and Suzuki, S. (1968) *J. Biol. Chem.* **243**, 1536–1542
- Kircher, H. W. (1960) *Anal. Chem.* **32**, 1103–1106
- Gerwig, G. J., Kamerling, J. P., and Vliegthart, J. F. (1979) *Carbohydr. Res.* **77**, 10–17
- Vieira, R. P., Mulloy, B., and Mourão, P. A. S. (1991) *J. Biol. Chem.* **266**, 13530–13536
- Alves, A. P., Mulloy, B., Diniz, J. A., and Mourão, P. A. S. (1997) *J. Biol. Chem.* **272**, 6965–6971
- Pomin, V. H., Pereira, M. S., Valente, A. P., Tollefsen, D. M., Pavão, M. S., and Mourão, P. A. S. (2005) *Glycobiology* **15**, 369–381
- Mourão, P. A. S., and Perlin, A. S. (1987) *Eur. J. Biochem.* **166**, 431–436
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005) *J. Comput. Chem.* **26**, 1701–1718
- Van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hunenberger, P. H.,

## Sulfated $\beta$ -Galactan from Sea Urchin

- Kruger, P., Mark, A. E., Scott, W. R. P., and Tironi, I. G. (1996) *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, University of Groningen, Groningen, The Netherlands and ETH, Zurich, Switzerland
22. Schüttelkopf, A. W., and van Aalten, D. M. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1355–1363
  23. Verli, H., and Guimarães, J. A. (2004) *Carbohydr. Res.* **339**, 281–290
  24. Becker, C. F., Guimarães, J. A., and Verli, H. (2005) *Carbohydr. Res.* **340**, 1499–1507
  25. Pol-Fachin, L., and Verli, H. (2008) *Carbohydr. Res.* **343**, 1435–1445
  26. Berendsen, H. J., Grigera, J. R., and Straatsma, T. P. (1987) *J. Phys. Chem.* **91**, 6269–6271
  27. Verli, H., and Guimarães, J. A. (2005) *J. Mol. Graph. Model.* **24**, 203–212
  28. Hess, B., Bekker, H., Berendsen, H. J., and Fraaije, J. G. (1997) *J. Comput. Chem.* **18**, 1463–1472
  29. Darden, T., York, D., and Pedersen, L. (1993) *J. Chem. Phys.* **98**, 10089–10092
  30. Su, Y. H., and Vacquier, V. D. (2006) *Mol. Biol. Cell* **17**, 114–121
  31. Alves, A. P., Mulloy, B., Moy, G. W., Vacquier, V. D., and Mourão, P. A. (1998) *Glycobiology* **8**, 939–946
  32. Vilela-Silva, A. C., Alves, A. P., Valente, A. P., Vacquier, V. D., and Mourão, P. A. (1999) *Glycobiology* **9**, 927–933
  33. Vilela-Silva, A. C., Castro, M. O., Valente, A. P., Biermann, C. H., and Mourao, P. A. (2002) *J. Biol. Chem.* **277**, 379–387
  34. Pomin, V. H., Valente, A. P., Pereira, M. S., and Mourão, P. A. (2005) *Glycobiology* **15**, 1376–1385
  35. Ravenscroft, N., Gamian, A., and Romanowska, E. (1995) *Eur. J. Biochem.* **227**, 889–896
  36. Hirohashi, N., and Vacquier, V. D. (2002) *Biochem. Biophys. Res. Commun.* **296**, 833–839
  37. Hirohashi, N., and Vacquier, V. D. (2003) *Biochem. Biophys. Res. Commun.* **304**, 285–292
  38. Amornrut, C., Toida, T., Imanari, T., Woo, E. R., Park, H., Linhardt, R., Wu, S. J., and Kim, Y. S. (1999) *Carbohydr. Res.* **321**, 121–127
  39. Pereira, M. S., Vilela-Silva, A. C., Valente, A. P., and Mourão, P. A. (2002) *Carbohydr. Res.* **337**, 2231–2238
  40. Woods, R. J. (1995) *Curr. Opin. Struct. Biol.* **5**, 591–598
  41. Becker, C. F., Guimarães, J. A., Mourão, P. A., and Verli, H. (2007) *J. Mol. Graph. Model.* **26**, 391–399
  42. Bilan, M. I., Vinogradova, E. V., Shashkov, A. S., and Usov, A. I. (2007) *Carbohydr. Res.* **342**, 586–596
  43. Farias, E. H., Pomin, V. H., Valente, A. P., Nader, H. B., Rocha, H. A., and Mourão, P. A. (2008) *Glycobiology* **18**, 250–259
  44. Mulloy, B., Ribeiro, A. C., Alves, A. P., Vieira, R. P., and Mourão, P. A. (1994) *J. Biol. Chem.* **269**, 22113–22123
  45. Usov, A. I., Bilan, M. I., and Shashkov, A. S. (1997) *Carbohydr. Res.* **303**, 93–102

### **4.3 Trabalho II**

Considerando os trabalhos anteriores do grupo, no estudo da conformação de glicosaminoglicanos e polissacarídeos sulfatados, o presente trabalho tem por objetivo avaliar o emprego desta metodologia, baseada no campo de força GROMOS96 43a1 adicionado por cargas atômicas de Löwdin na base HF/6-31G\*\*, em reproduzir o perfil conformacional de carboidratos ligados a proteínas previamente estudadas por RMN e descrever a influência da parte sacarídica sobre a parte protéica de glicoproteínas, validando, assim, o emprego desta metodologia no estudo destas macromoléculas biológicas.

Os resultados obtidos indicam que a metodologia empregada é capaz de reproduzir diferentes aspectos conformacionais de glicoproteínas de acordo com dados experimentais prévios: (1) a conformação de N- e O-ligações glicosídicas, formadas entre aminoácidos e monossacarídeos; (2) a conformação da parte sacarídica de glicoproteínas, de maneira que mais de 95% das geometrias determinadas por RMN para as ligações glicosídicas que compõe tais oligossacarídeos mostraram-se de acordo com os dados de DM; (3) os efeitos da glicosilação sobre a estrutura e conformação da parte protéica de glicoproteínas, conferindo estabilidade conformacional e influenciando propriedades relacionadas a função dessas moléculas. Tais resultados confirmam a capacidade das simulações de DM, utilizando parâmetros do campo de força GROMOS96 e cargas atômicas de Löwdin na base HF/6-31G\*\*, em descrever a conformação de glicoproteínas, contribuindo assim na ampliação do número de protocolos disponíveis para o estudo de carboidratos biológicos e no volume de trabalhos abordando estas biomoléculas através de técnicas de modelagem molecular.

## **GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations**

**Laercio Pol-Fachin, Cláudia Lemelle Fernandes, Hugo Verli**

*Carbohydrate Research*, **2009**, 344; 491-500



Contents lists available at ScienceDirect

Carbohydrate Research

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

## GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations

Laercio Pol-Fachin<sup>a</sup>, Claudia Lemelle Fernandes<sup>a</sup>, Hugo 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 3 September 2008  
Received in revised form 3 December 2008  
Accepted 26 December 2008  
Available online 3 January 2009

#### Keywords:

Glycoproteins  
Molecular dynamics  
Carbohydrates  
Glycosidic linkage  
GROMOS96

### ABSTRACT

Considering the small number of papers assessing the conformational profile of glycoproteins through molecular dynamics (MD) simulations, the current work reports on a systematic analysis of the performance of the GROMOS96 43a1 force field and Löwdin HF/6-31G\*\*--derived atomic charges in the conformational description of glycoproteins. The results substantiate the accuracy of the computational representation of glycoprotein conformational ensembles in aqueous solution based on their agreement to available experimental information, supporting further contributions of computational techniques, mainly MD, in future studies on the characterization of glycoprotein structure and function.

© 2009 Elsevier Ltd. All rights reserved.

### 1. Introduction

Glycoproteins and glycopeptides are known to play important roles in many biological events, such as cell adhesion, cell–cell communication, immune response, intracellular targeting, and protease resistance.<sup>1,2</sup> These glycans can be located at various positions of the protein surface, depending on specific consensus sequences, and are linked only through specific amino acid residues. The linkages most frequently found in nature are the N-glycosidic linkages, involving mainly asparagine (Asn) and N-acetyl-D-glucosamine (GlcNAc) residues, and the O-glycosidic linkages, mostly including serine (Ser) or threonine (Thr) and, for instance, D-galactose (Gal), N-acetyl-D-galactosamine (GalAc), L-fucose (Fuc), or D-mannose (Man) residues. Structurally, these carbohydrate moieties impact several physicochemical properties of proteins, including hydration and polarity, frequently participating in its folding and conformational stabilization.<sup>1,2</sup>

The comprehension of the three-dimensional structure and the dynamical properties of both protein and carbohydrate moieties of glycoproteins is a requirement for a better understanding of the molecular basis of their interaction with each other, with the surrounding environment, and with their molecular targets. In this context, several experimental methods have been applied to study glycosylated proteins and glycans, such as X-ray crystallography and NMR spectroscopy. Concerning the former technique, the con-

formational flexibility of the glycan at the protein surface usually prevents such molecules from being crystallized,<sup>3,4</sup> but when crystals are available, the electron density is affected by the high thermal motion of the glycan moiety, compromising the accuracy of the geometry beyond the rigid core region of N-glycans.<sup>5</sup> NMR methods provide the solution ensemble of conformations in a set of average three-dimensional models, but an unambiguous determination of the complete conformational space of glycoproteins is hampered because only from one to three contacts across a given glycosidic linkage are usually detected.<sup>6</sup>

Considering the difficulties associated with the experimental determination of glycoproteins structure and conformation, MD simulations have been used to clarify the dynamic aspects of macromolecules. In particular, such simulations support the study of these molecules mimicking their natural environment and describe their conformational properties with a reasonable level of accuracy. On the other hand, only a small number of studies have assessed the profile of glycoproteins and glycopeptides through MD simulations, employing different set of parameters for their glycan parts, such as AMBER, CHARMM, and CVFF.<sup>7–9</sup> In comparison to such parameters, our group has been working on MD conformational representation of carbohydrates, polysaccharides, and glycosaminoglycans<sup>10–13</sup> based on the GROMOS96 43a1 force field,<sup>14</sup> enhanced by Löwdin HF/6-31G\*\*--derived atomic charges.<sup>10,12</sup> This approach supports the adequate description of such molecules in aqueous solutions.

Therefore, the current work intends to evaluate the capability of the GROMOS96 43a1 force field and Löwdin HF/6-31G\*\*--derived

\* 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).

atomic charges, and to represent adequately the conformational ensemble of glycoproteins through MD simulations in aqueous solutions. The results of these simulations were compared to experimental data detailing the geometry of the glycosidic linkages between two monosaccharides and between a monosaccharide and an amino acid. Additionally, the effects of glycosylation upon the protein moiety of the glycoproteins were also assessed. Based on the data, the employed methodology may be expected to further contribute to structural and functional studies of glycoconjugates through reliable and accessible methodology.

## 2. Experimental

### 2.1. Computational methods

#### 2.1.1. Nomenclature and software

The nomenclature recommendations and symbols as proposed by IUPAC<sup>15</sup> were used. The relative orientation of a pair of contiguous carbohydrate residues is 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  $\psi$  are defined as shown below:

$$\phi = \text{O5}-\text{C1}-\text{OX}-\text{CX} \quad (1)$$

$$\psi = \text{C1}-\text{OX}-\text{CX}-\text{C}(X-1) \quad (2)$$

For a (1→6) linkage, the  $\omega$  is defined as shown below:

$$\omega = \text{O6}-\text{C6}-\text{C5}-\text{C4} \quad (3)$$

The manipulation of structures was performed with MOLGEN<sup>16</sup> and VMD,<sup>17</sup> the secondary structure content analyses were performed with PROCHECK,<sup>18</sup> and all the MD calculations and remaining analyses were performed using GROMACS simulation suite<sup>19</sup> and GROMOS96 force field,<sup>14</sup> with all saccharide topologies generated with PRODRG.<sup>20</sup>

#### 2.1.2. Topology construction

The oligosaccharide fragments present in the glycoproteins under the PDB codes 1CDR,<sup>21</sup> 1CDS,<sup>21</sup> 1GYA,<sup>22</sup> 1HD4,<sup>23</sup> and 1FF7<sup>24</sup> (Fig. 1) were described in GROMOS96 43a1 force field through PRODRG server-derived topologies,<sup>20</sup> which were further refined by atomic charges, as described previously.<sup>10,12</sup> Briefly, each monosaccharide was submitted to HF/6-31G<sup>+</sup> energy minimization, followed by Hessian matrix analyses to characterize them unequivocally as true minima at the potential energy surface. The residues were divided in charge groups, while improper dihedrals were included as necessary to preserve the hexopyranose conformation of each monosaccharide in accordance with its expected form in aqueous solution: <sup>4</sup>C<sub>1</sub> for D-GlcNAc, 1C<sub>4</sub> for L-Fuc, 4C<sub>1</sub> for D-Man, and <sup>4</sup>C<sub>1</sub> for D-Gal. Additionally, proper dihedrals, as described in GROMOS96 43a1 force field for glucose, were also included in the PRODRG obtained topologies in order to support stable simulations. Concerning the N- and O-linkages between the carbohydrate and amino acid residues, the atomic charges were calculated for the Asn-carbohydrate (GlcNAc C<sub>1</sub> = 0.110, Asn N<sub>δ</sub> = -0.265, and Asn H<sub>Nδ</sub> = 0.155) or Ser-carbohydrate (Fuc C<sub>1</sub> = 0.150, Ser O<sub>γ</sub> = -0.320,

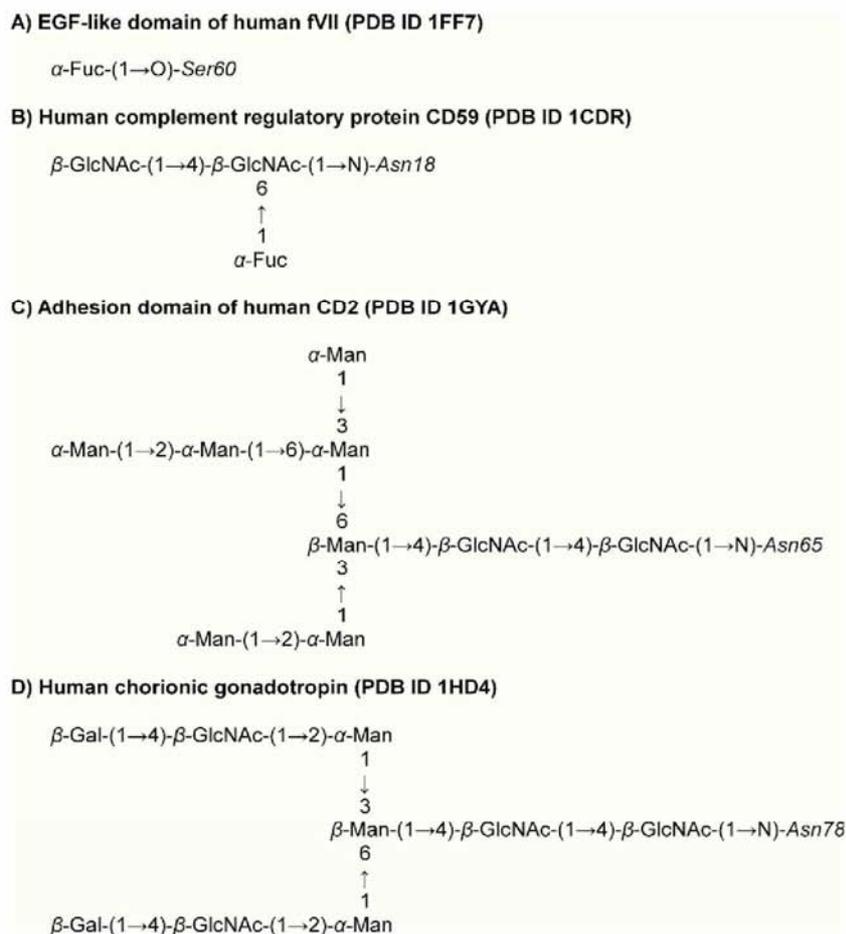


Figure 1. Schematics of the studied glycans and the corresponding glycosylated amino acids in the studied proteins.

and Ser C<sub>β</sub> = 0.170) connection charge groups. Such an approach, based on parameters derived from minimum energy conformations, has been successfully applied in previous work in carbohydrate modeling.<sup>10–13,25,26</sup>

### 2.1.3. MD simulations

The structures of glycoproteins, as obtained from NMR data, were retrieved from the PDB, including the human complement regulatory protein CD59 (CD59), in its non-glycosylated (PDB ID 1CDQ) and glycosylated (PDB codes 1CDR and 1CDS) forms,<sup>21</sup> the  $\alpha$ -subunit of the human chorionic gonadotropin ( $\alpha$ -hCG), in both non-glycosylated (PDB ID 1DZ7)<sup>27</sup> and glycosylated (PDB ID 1HD4) structures,<sup>23</sup> the first epidermal growth factor-like (EGF-like) domain of the human blood coagulation VII in its non-glycosylated (PDB ID 1F7E) and glycosylated (PDB ID 1FF7) forms,<sup>24</sup> and the adhesion domain of human CD2 (hsCD2<sub>105</sub>) in its glycosylated (PDB ID 1GYA)<sup>22</sup> and non-glycosylated (PDB ID 1CDB)<sup>28</sup> forms, in a total of five glycosylated (Fig. 1) and four non-glycosylated proteins. These structures were solvated in a rectangular box using periodic boundary conditions and SPC water model.<sup>29</sup> Amino acids ionization was automatically adjusted to the plasma pH, with histidines being non-ionized and at the N<sub>τ</sub>-H tautomeric form. Counterions (Na<sup>+</sup> and Cl<sup>-</sup>) were added to neutralize the systems, whenever needed. The employed MD protocol was based on previous studies, as described.<sup>10,12</sup> The Lincs method<sup>30</sup> was applied to constrain covalent bond lengths, allowing an integration step of 2 fs after an initial energy minimization using Steepest Descent algorithm. Electrostatic interactions were calculated with Particle Mesh Ewald method.<sup>31</sup> Temperature and pressure were kept constant by coupling glycoproteins, ions, and solvent to external temperature and pressure baths with coupling constants of  $\tau = 0.1$  and 0.5 ps,<sup>32</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 50 ns, without any restraint, while a reference value of 3.5 Å between heavy atoms was considered for a hydrogen bond and a cutoff angle of 30° was used between hydrogen-donor-acceptor.<sup>19</sup> As the NMR selected structures include several models in the same PDB file, a total of eight glycoproteins (one model for 1CDS and 1FF7, and two models for 1CDR, 1GYA, and 1HD4), plus four non-glycosylated counter-parts were simulated to achieve a broader conformational description through MD. Details of the simulated models are included in Table 1. Additional simulations of the non-glycosylated proteins were performed with OPLS-AA force field in order to check the pattern of secondary structure as observed in GROMOS96 MD.

## 3. Results and discussion

### 3.1. Simulation systems

The analysis of GROMOS96 43a1 force field performance on glycoproteins conformational description was based on the MD simulation of four different proteins: (1) the EGF-like domain of human FVII; (2) the human complement regulatory protein CD59; (3) the adhesion domain of human CD2; and (4) the  $\alpha$ -subunit of the human chorionic gonadotropin. These proteins were simulated in both glycosylated and non-glycosylated forms, taking their NMR-derived model as the starting geometry. Among these molecules, three have N-linked glycans and one has an O-linked glycan. These systems were analyzed taking the NMR-obtained ensemble from each macromolecule as reference, to supply the information of GROMOS96 43a1 force field performance on the description of glycan-protein linkage of different glycosidic linkages composing the

glycan structures, and the influence of glycosylation on the protein moiety of each molecule.

Actually, conformational data derived from single crystals of glycoconjugates should be considered carefully since the crystal environment may supply distorted hexopyranose rings and/or be submitted to packing effects, which are also known to occur for other classes of molecules, such as DNA<sup>33</sup> and proteins.<sup>34,35</sup> In contrast, although NMR techniques are not considered to provide an unambiguous determination of the complete conformational space of glycans,<sup>3</sup> it does provide solution conformations for the studied glycans which, in a set of average three-dimensional structures, could provide a realistic sample of glycans-composing disaccharide conformations and, consequently, a better reference for the experimental support of MD in solutions.

### 3.2. N- and O-glycosidic linkage geometry

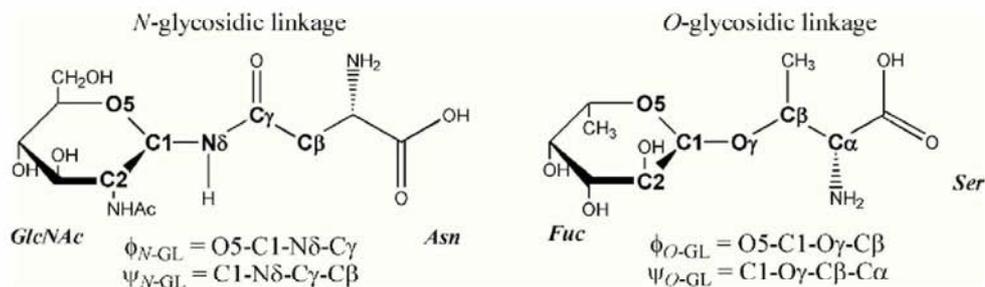
One of the main features determining the orientation of carbohydrates attached to proteins lies in its linkage to specific amino acid residues, such as asparagine and serine/threonine. The N-glycosidic linkages have been extensively studied by both NMR and X-ray methods, while a lower amount of experimental data is available for the O-glycosidic linkages, probably due to the increased flexibility and versatility of these linkages, involving distinct amino acid and monosaccharide residues. Therefore, the performed simulations of glycoproteins were initially analyzed by considering the connection between carbohydrates and proteins, using a set of experimental data as reference, as shown in Table 1 for average geometries and in Figure 2 for the respective conformational behavior of each linkage.

As can be observed, previous analysis of X-ray data indicates that the  $\phi_{N-GL}$  dihedral angle presents higher flexibility compared to the  $\psi_{N-GL}$  dihedral angle, with most X-ray structures presenting the  $\phi_{N-GL}$  with a 80° amplitude (between -140° and -60°) and  $\psi_{N-GL}$  with a 40° amplitude (between 160° and -160°).<sup>36</sup> Additional data, based on ~500 crystal structures obtained from the PDB, indicate a lower flexibility, in which the GlcNAc-(1→N)-Asn linkage exhibits an average  $\phi_{N-GL}$  geometry at ~-100° and  $\psi_{N-GL}$  at ~180°.<sup>37</sup> One of the most recent papers upon such linkages, combining X-ray crystallography and both HF and B3LYP with the 6-31+G\* basis set, on models and analogs of GlcNAc-(1→N)-Asn,<sup>38</sup> also observed similar geometries for such dihedral angles. This large amplitude of the  $\phi_{N-GL}$  also suggests the occurrence of multiple conformer populations co-existing in solution and being sustained depending on the crystallographic environment, as indeed observed in MD (Fig. 2A).

Hence, all glycoproteins presented equivalent conformational patterns around the GlcNAc-(1→N)-Asn linkage during the performed simulations, in a behavior apparently independent on the surrounding protein scaffold (Fig. 2A and Supplementary data). The curve representing such conformational patterns include about 81% of the NMR experimental geometries of  $\phi_{N-GL}$ , as well as 99% in  $\psi_{N-GL}$  (in a total of 226 geometries for each dihedral angle),<sup>21–23,39–42</sup> indicating a good reliability of the simulations to predict the dynamics of N-linked glycans.

The geometry of O-glycosidic linkages was also evaluated and compared to three NMR studies, containing a fucose residue linked to two amino acid residues, Ser<sup>24</sup> or Thr,<sup>43</sup> as well as an additional structure presenting a 3-O-methyl-mannose residue linked to Thr.<sup>44</sup> It may be observed that the  $\phi_{O-GL}$  and  $\psi_{O-GL}$  dihedral angles of these linkages adopt similar geometries regardless of whether Ser and Thr is the involved amino acid residue (Table 1). Different from N-glycosidic linkages, no major flexibility differences could be observed when comparing  $\phi_{O-GL}$  and  $\psi_{O-GL}$  dihedral angles. In addition, the  $\phi_{O-GL}$  presented two major MD conformations, which are both found in NMR data, around -70° (in 1URK and 1FF7 PDB

**Table 1**  
Comparison of the dihedral angles around N- and O-glycosidic linkages, including X-ray, NMR, and MD simulations data



Glycosidic linkage	Structures	Dihedral angle (°) <sup>†</sup>			
		$\phi_{N-GL}$	$\psi_{N-GL}$	$\phi_{O-GL}$	$\psi_{O-GL}$
GlcNAc-(1→N)-Asn	X-ray <sup>a</sup>	-60 to -140	160 to -160	—	—
	X-ray <sup>b</sup>	-99 ± 21	177 ± 12	—	—
	NMR <sup>c</sup>	-91 ± 45	180 ± 7	—	—
	NMR-1CDR <sup>d</sup>	-77 ± 45	180 ± 0	—	—
	MD-1CDR <sup>e,*</sup>	-121 ± 29	-179 ± 14	—	—
	MD-1CDR <sup>f,*</sup>	-126 ± 28	179 ± 13	—	—
	NMR-1CDS <sup>d</sup>	-63 ± 44	180 ± 0	—	—
	MD-1CDS <sup>g,*</sup>	-119 ± 30	-177 ± 14	—	—
	NMR-1GYA <sup>d</sup>	-99 ± 15	179 ± 2	—	—
	MD-1GYA <sup>h,*</sup>	-121 ± 28	178 ± 13	—	—
	MD-1GYA <sup>i,*</sup>	-116 ± 29	-174 ± 14	—	—
	NMR-1HD4 <sup>d</sup>	-86 ± 50	177 ± 6	—	—
	MD-1HD4 <sup>j,*</sup>	-120 ± 30	-174 ± 14	—	—
	MD-1HD4 <sup>k,*</sup>	-116 ± 30	-178 ± 12	—	—
Fuc-(1→O)-Ser	NMR-1FF7 <sup>d</sup>	—	—	-67 ± 10	178 ± 10
	MD-1FF7 <sup>l,*</sup>	—	—	-110 ± 29	178 ± 22
Fuc-(1→O)-Thr	NMR-1URK <sup>d</sup>	—	—	-104 ± 30	-154 ± 76
3-O-methyl-Man-(1→O)-Thr	NMR-2HGO <sup>d</sup>	—	—	72 ± 3	120 ± 22

<sup>†</sup> Values presented as averages ± standard deviation values, obtained from all trajectory points, that is, no distinction of conformer populations was made to support a direct comparison to experimental data.

<sup>a</sup> 50 ns MD averages.

<sup>b</sup> Data from Ref. 36.

<sup>c</sup> Data from Ref. 37, wherein the geometry from most populous conformer is present.

<sup>d</sup> Average dihedral angle values from all models included in PDB codes: 1CDR,<sup>21</sup> 1CDS,<sup>21</sup> 1GYA,<sup>22</sup> 1HD4,<sup>23</sup> 1BZB,<sup>39</sup> 1BYV,<sup>39</sup> 2FN2,<sup>40</sup> 1E88,<sup>41</sup> 1E9J,<sup>23</sup> 2JXA,<sup>42</sup> and 2JX9.<sup>42</sup>

<sup>e</sup> Average dihedral angle values from all models included in PDB codes 1CDR and 1CDS (10), 1URK (15), 1GYA (18), 1FF7 and 2HGO (20) and 1HD4 (26).

<sup>f</sup> Using the model 1 from PDB file 1CDR as a starting structure.

<sup>g</sup> Using the model 8 from PDB file 1CDR as a starting structure.

<sup>h</sup> Using the model 1 from PDB file 1CDS as a starting structure.

<sup>i</sup> Using the model 1 from PDB file 1GYA as a starting structure.

<sup>j</sup> Using the model 12 from PDB file 1GYA as a starting structure.

<sup>k</sup> Using the model 1 from PDB file 1HD4 as a starting structure.

<sup>l</sup> Using the model 11 from PDB file 1HD4 as a starting structure.

<sup>†</sup> Using the model 1 from PDB file 1FF7 as a starting structure. Values of 111 ± 28° were also populated in the simulation.

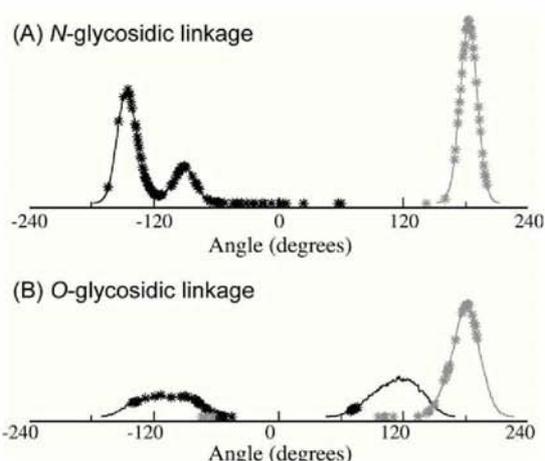
files) or around 70° (in 2HGO PDB file). Such a pattern is also in good agreement with the distribution of the dihedral angles around the Fuc-(1→O)-Ser linkage of the EGF-like domain, where two main peaks are indeed observed around the  $\phi_{O-GL}$  angle (Fig. 2B). Again, good reliability is observed for O-linked glycoproteins, with 97% of the experimental geometries within the distribution curve for  $\phi_{O-GL}$  and 69% for  $\psi_{O-GL}$  (in a total of 45 geometries for each dihedral angle).<sup>24,43,44</sup>

### 3.3. Dynamics of glycosidic linkages that compose the glycan part of glycoproteins

Similar to the conformational behavior observed for the linkage between monosaccharides and amino acids, a given disaccharide showed mostly equivalent conformational profiles, with a minor influence of the protein scaffold on glycan conformational preference (Fig. 3 and Supplementary data). As observed, most of the experimental conformations, defined by NMR data, are within the distribution curves obtained from the MD simulations, indicat-

ing an adequate reproduction of the glycan conformational ensemble. In addition, considering that NMR methods provide average three-dimensional models, giving rise to ensemble-average properties instead of single solution conformations,<sup>6</sup> it is feasible to observe NMR geometries between main conformational states/peaks. Such characteristics also correlate to the observation of some geometries not included in the main regions of MD distribution curves, as it may not represent distinct conformers but may represent average ones.

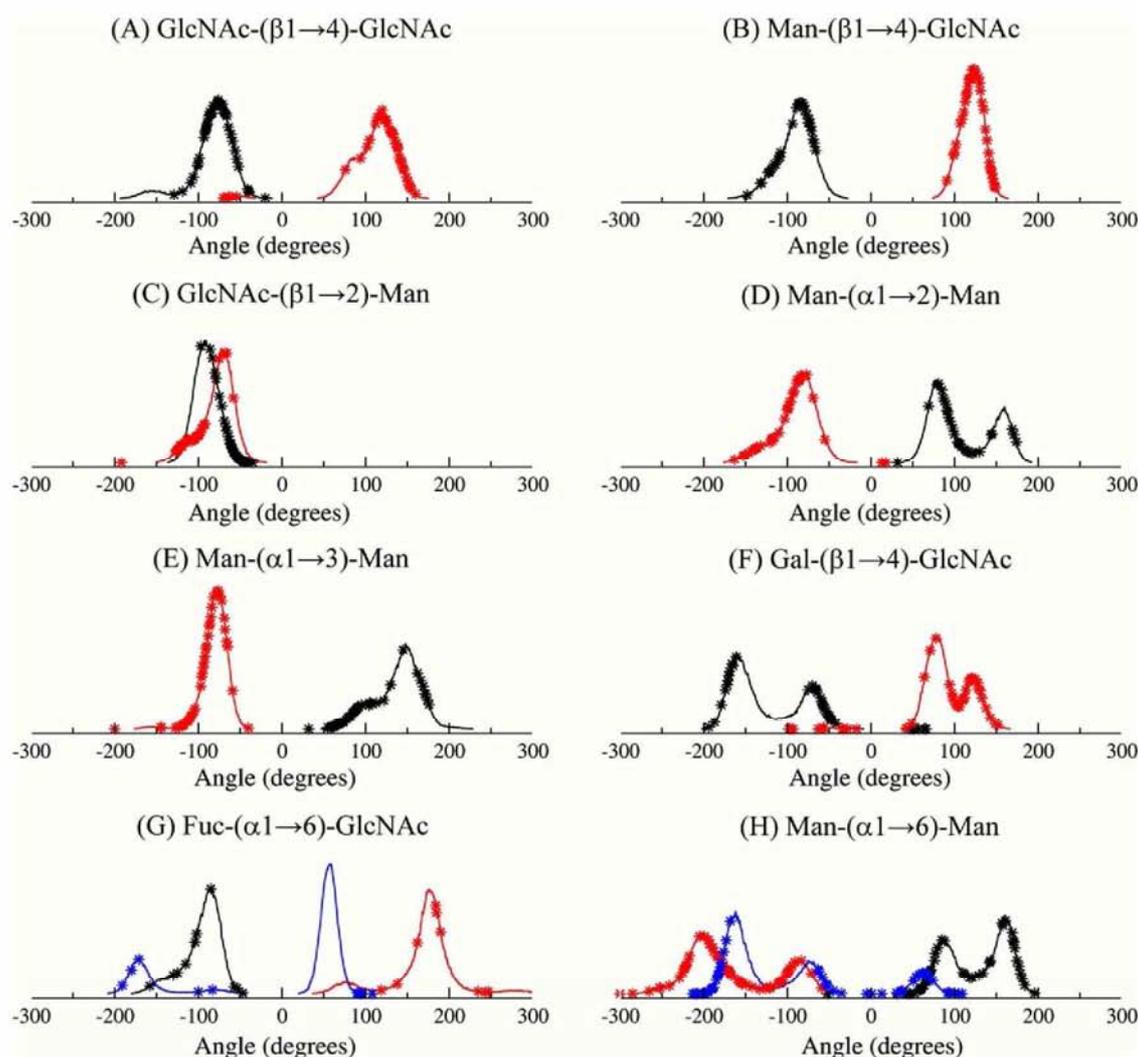
Although the correct reproduction of the oligosaccharide geometry is important, the flexibility associated with disaccharides exposed to different environments must also be taken into account. In this context, it is proposed that the number of possible conformations for the dihedral angles of a given disaccharide, when comprising a glycan moiety of a glycoprotein, is considerably reduced at its core compared to the distal ends.<sup>5</sup> This appears to be caused by the conformational restriction promoted by the surrounding carbohydrate residues, although its geometry is similar to the one present in the free carbohydrate.<sup>2</sup> In accordance with such a



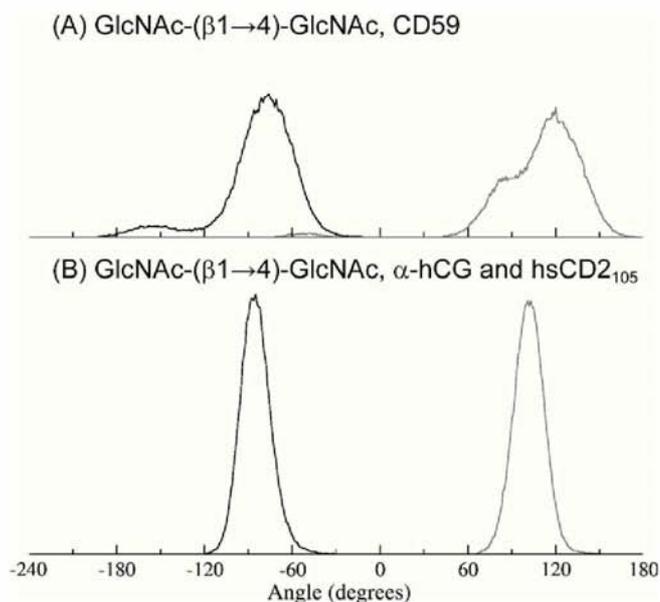
**Figure 2.** Distribution of the  $\phi$  (black) and  $\psi$  (gray) dihedral angles associated with N- and O-glycosidic linkages in 50 ns MD simulations of the studied glycoproteins. Experimental geometries, as obtained from NMR data, are indicated by asterisks (\*).

model, we were able to observe differences in the conformational behavior of the same disaccharide,  $\beta$ -GlcNAc-(1 $\rightarrow$ 4)-GlcNAc, in

different molecular contexts (Figs. 1 and 4). In linear chains, with little or no steric hindrance, as seen in CD59 (PDB code 1CDR, Fig. 4A), a more flexible pattern is observed. In contrast, when such disaccharide is the core of a more complex glycan, as seen hsCD2<sub>105</sub> and  $\alpha$ -hCG (PDB codes 1GYA and 1HD4, respectively, Fig. 4B), additional rigidity occurs, although the main peaks are highly similar, that is, a related set of the main conformational states. Such profiles seem to be influenced by the solvent accessibility around the disaccharides, as a higher solvent accessible surface (SAS) is observed in more flexible glycosidic linkages (Supplementary data). A similar influence of solvent exposure is observed for  $\alpha$ -Man-(1 $\rightarrow$ 3)-Man depending on its localization at the glycan structure, that is, a given disaccharide will have higher flexibility when more exposed to solvent (Supplementary data). On the other hand, the presence of a linked carbohydrate residue alone does not necessarily influence the surrounding residues accessibility to the solvent. For example, the conformational pattern of the  $\beta$ -GlcNAc-(1 $\rightarrow$ 4)-GlcNAc linkage in the two CD59 PDB structures (1CDR and 1CDS) appears to not be influenced by a fucose six-linked to the GlcNAc residue (Supplementary data). This is probably due to the high flexibility of the (1 $\rightarrow$ 6)-linkage, which has less steric hindrance with the remaining glycan residues and, consequently, less influence on the solvation of the surrounding glycan.



**Figure 3.** Distribution of the  $\phi$  (black),  $\psi$  (red) and  $\omega$  (blue) dihedral angles associated with specific disaccharides that compose the carbohydrate moieties of the studied glycoproteins in 50 ns MD simulations. Each curve is representative for the conformational behavior of a given disaccharide through the simulated glycoproteins (see Supplementary data). Asterisks (\*) in the distribution curves indicate the experimental geometries, as obtained from NMR data, in a total of 836 values.



**Figure 4.** Distribution of the  $\phi$  (black) and  $\psi$  (gray) dihedral angles associated with the  $\beta$ -GlcNAc-(1 $\rightarrow$ 4)-GlcNAc disaccharide behavior during 50 ns MD on different glycoproteins: (A) Attached to CD59 (PDB ID 1CDR and 1CDS); (B) Attached to both hsCD2<sub>105</sub> (PDB ID 1GYA) and  $\alpha$ -hCG (PDB ID 1HD4).

Among the studies upon the conformational preferences of specific glycosidic linkages, special attention should be given to the description of the  $\omega$  angle preferred orientations.<sup>45</sup> This bond is reported to adopt three main conformational states, around 180°, 60° and -60°, which is in good agreement with the conformational behavior observed from MD simulations. In contrast, it can be observed that NMR-measured dihedrals with regard to the  $\omega$  angle of both  $\alpha$ -Fuc-(1 $\rightarrow$ 6)-GlcNAc and  $\alpha$ -Man-(1 $\rightarrow$ 6)-Man disaccharides adopt distinct geometries with large amplitudes (Fig. 3). These may indeed represent average values from two or three conformational states, which are observed to occur in solution for such disaccharides through MD.

### 3.4. Effect of glycosylation upon the protein moiety of glycoproteins

Although the biological functions of glycosylation are still not completely understood,<sup>2,46</sup> it is known that the glycan moieties of glycoproteins are fundamental to many biological processes, and are also involved in protein folding and/or stabilization.<sup>1,2</sup> In previous studies, MD simulations were employed to explore the role of glycosylation on peptides and proteins, such as the C-terminal region of the human prion glycoprotein,<sup>47</sup> the MUC1 peptide,<sup>48</sup> and a glycopeptide derived from a hemagglutinin protein fragment,<sup>49</sup> employing the AMBER force field,<sup>7</sup> and the MHC class I glycoprotein,<sup>50</sup> which used CVFF.<sup>9</sup> In this context, to evaluate simultaneously the glycan conformational description in MD, we have also performed an analysis of the influence of glycosylation on the dynamics and conformation of the protein moiety (Figs. 5 and 6), and these results will be discussed below. Additionally, the non-glycosylated molecules were also simulated with the OPLS-AA force field, with a secondary structure pattern equivalent to that observed with GROMOS96 (Supplementary data), which support the reliability of the obtained results.

#### 3.4.1. EGF-like

According to experimental studies,<sup>24</sup> fucosylation does not significantly affect the structure of the EGF-like motif. In agreement

with such an observation, no major differences in the secondary structure pattern between non-fucosylated and fucosylated proteins could be observed in MD simulations, as for RMSD (Fig. 5E) and radius of gyration (Fig. 5I). A high RMSD value for both forms of the EGF-like motif was observed, which appears to be due to the high loop content of this 46 amino acids protein, mainly its C- and N-terminal regions. On the other hand, a difference in flexibility between glycosylated and non-glycosylated proteins is observed around residues 56–59 (Fig. 6A and B), suggesting that the recognition of EGF-like domains by the O-fucosyltransferase may be associated with an entropic driven process, as this region lies within a putative consensus sequence for fucosylation in several EGF-like domains.<sup>24</sup>

#### 3.4.2. CD59

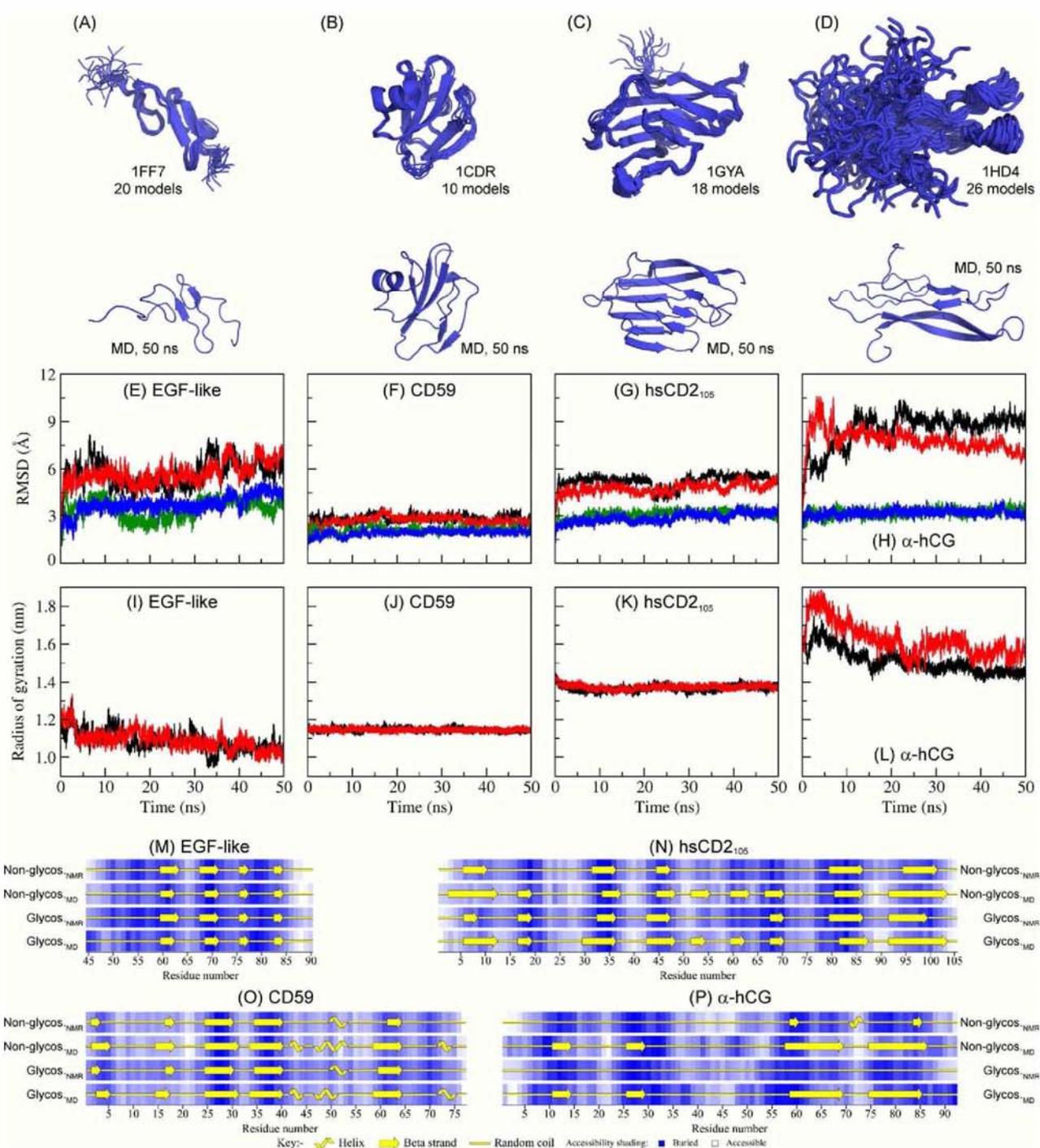
The soluble form of CD59 had been studied by NMR methods, revealing a monomeric protein of 77 residues in a non-glycosylated form or with a disaccharide or trisaccharide N-glycan attached to Asn18.<sup>21</sup> Such data indicated that the global conformation of the protein backbone is essentially unchanged upon inclusion of oligosaccharides in the structure calculations. In accordance to these observations, MD simulations of non-glycosylated and glycosylated forms of CD59 showed no major differences between both forms of the protein for RMSD (Fig. 5F) and radius of gyration (Fig. 5J). Additionally, the NMR data pointed to small changes in the mean dihedrals and angular parameters of the protein component upon inclusion of carbohydrate groups, mainly around the side chain of Asn18.<sup>21</sup> Accordingly, the RMSF analyses of the simulated proteins support a decreased flexibility in the region of the N-glycan attachment, around Asn18 (Fig. 6C and D).

#### 3.4.3. hsCD2<sub>105</sub>

Regarding hsCD2<sub>105</sub>, no major differences could be observed between non-glycosylated and glycosylated proteins based on RMSD and radius of gyration (Fig. 5G and K). However, an increased RMSD occurs in both forms of the protein, probably due to high loop content (mainly C- and N-terminal regions, Fig. 5C and G) and to the formation of several  $\beta$ -strands during MD simulations, known to exist in solution.<sup>22,28,51</sup> However, according to PROCHECK analysis, they are present in lower amounts in the NMR structures of both glycosylated and non-glycosylated forms (Fig. 5N).

Additionally, the NMR data suggest that the N-linked glycan at Asn65 is required for adhesion functions, because it stabilizes protein folding by counterbalancing an unfavorable clustering of five positive charges centered around Lys61.<sup>22</sup> Such an effect appears to be observed in MD simulations, because the region around these residues presents decreased flexibility upon glycosylation, as well as around regions comprising residues 22–32, 49–53, 61–67, and the N-terminal (Fig. 6E and F) domain.

Regarding CD2 function, transmembrane variants with mutations in the consensus N-glycosylation sequence Asn65-Gly66-Thr67,<sup>52,53</sup> which prohibit the attachment of the high-mannose N-glycan at Asn65, could be normally expressed on cell surfaces. However, they have shown neither antibody- nor ligand-binding activity, suggesting that the N-linked glycan on hsCD2<sub>105</sub> plays an important role in the function of this protein. In addition, the NMR studies suggested that protein-carbohydrate interactions may be involved in mediating the CD2 interaction with CD58 which should involve, mainly, the amino acids 32–34, 43–51 and 86–91 of hsCD2<sub>105</sub>.<sup>52,54</sup> Accordingly, the residues around these regions concentrate the main differences in protein flexibility upon glycosylation (Fig. 6E and F), while participating in a series of  $\beta$ -sheets folded in MD apparently promoted by the glycan presence (Fig. 5N), suggesting a possible role of  $\beta$ -strand structures on CD2-CD58 binding.

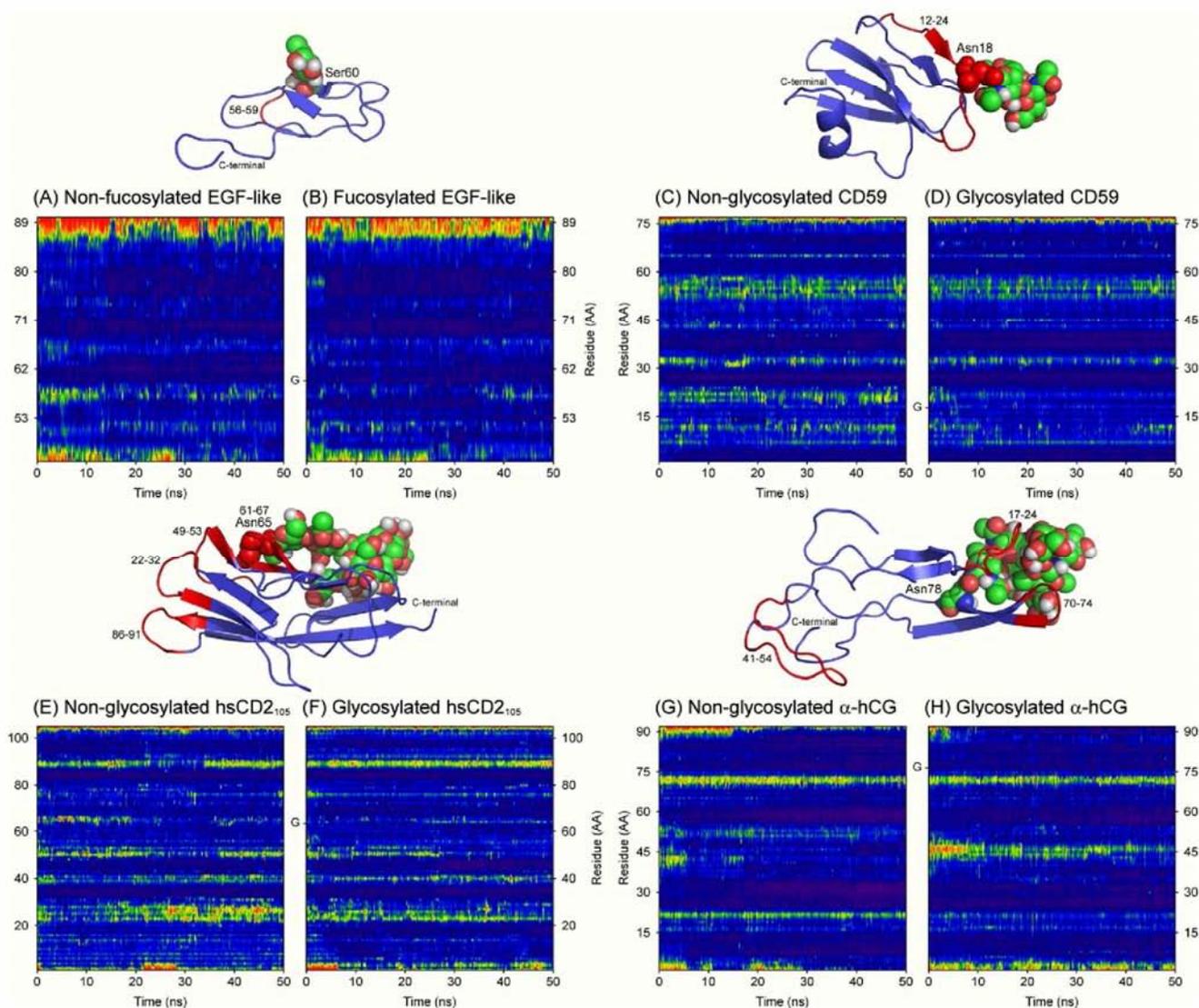


**Figure 5.** Comparative analysis between the studied glycoproteins and their non-glycosylated counterparts. The superimposed NMR models and the final MD conformation of EGF-like (A), CD59 (B), hsCD2<sub>105</sub> (C) and  $\alpha$ -hCG (D) are presented together with root mean square deviation (RMSD) (E, F, G and H, respectively), in relation to the starting NMR model, for its respective non-glycosylated (black) and glycosylated (red) forms considering all atoms and the radius of gyration (I, J, K and L, respectively). The RMSD profile associated only with the atoms included in secondary structure elements of each molecule are shown for the non-glycosylated (green) and glycosylated (blue) forms. A comparative PROCHECK analysis of secondary structure content, obtained from NMR and final MD conformations of non-glycosylated and glycosylated proteins, is also shown (see Section 2 for further details).

#### 3.4.4. $\alpha$ -hCG

The  $\alpha$ -hCG structure has been determined through NMR methods in its glycosylated and non-glycosylated forms and deposited under PDB codes 1HD4<sup>23</sup> and 1DZ7,<sup>27</sup> respectively. Due to its high loop content, increased RMSD values are observed during MD simulations for both forms of the protein (Fig. 5H). This may also be related to the formation of a series of  $\beta$ -sheets in the simulations

(Fig. 5P), which is also observed in OPLS/AA simulations of the non-glycosylated protein (Supplementary data). Additionally, some important differences could be observed between non-glycosylated and glycosylated forms of  $\alpha$ -hCG based on RMSD (Fig. 5H) and radius of gyration (Fig. 5L). These differences seem to be caused by different loop flexibilities during the MD simulations (Fig. 5H), but especially by the large disordered loop between



**Figure 6.** Root mean square fluctuation (RMSF) for the studied proteins, as a function of both residue number and time: EGF-like, in its non-fucosylated (A) and fucosylated (B) forms (from 0.04 to 0.24 nm); CD59, in its non-glycosylated (C) and glycosylated (D) forms (from 0.03 to 0.18 nm); hCD2<sub>105</sub>, in its non-glycosylated (E) and glycosylated (F) forms (from 0.04 to 0.16 nm); and the  $\alpha$ -hCG, in its non-glycosylated (G) and glycosylated (H) forms (from 0.05 to 0.35 nm).

residues 33–58 (Fig. 5P). This loop is known to exist in free  $\alpha$ -hCG, based on NMR studies,<sup>55</sup> and seems to adopt increased flexibility in the glycosylated protein (Fig. 6G and H). Such behavior is reproducible in an additional simulation, taking a different NMR model as starting structure under the same PDB code. As a general feature, glycosylation of the hCG  $\alpha$ -subunit at Asn78 seems to play an important role in the stabilization of the protein<sup>23</sup> by exerting a protective function through shielding the protein surface from the environment,<sup>55</sup> including residues around 23–26 and 68–70. Therefore, from the MD simulations, we could observe a stabilization of such regions, especially 23–26, as observed in RMSF analysis (Fig. 6G and H).

### 3.5. Conformation of glycoproteins and their assessment through MD simulations

In recent years, glycobiology has become a critical facet of post-genomic science<sup>56</sup> as many proteins are post-translationally modified by glycosylation or co-translationally added by N-glycans in the endoplasmic reticulum.<sup>46</sup> Glycosylation may alter and regulate the biological activities of such macromolecules. In this context,

several structural and functional studies on glycans alone, or on glycopeptides or glycoproteins have been carried out, using both experimental and theoretical methods. Concerning the use of the latter techniques, an assortment of approaches has been employed to model glycan structures when its three-dimensional structure is absent or is considered inappropriate. These include energy minimization of an entire oligosaccharide,<sup>47</sup> submitting the glycan to a simulated annealing protocol,<sup>57</sup> or constructing the carbohydrate moiety of glycoproteins from the disaccharide level,<sup>50</sup> that is, minimized structures of isolated disaccharides used to assemble an oligosaccharide tree. Additionally, several force field parameters have been employed to describe and simulate carbohydrate moieties of glycoproteins.<sup>7–9</sup> However, data regarding to the geometry of glycosidic linkages of such biomacromolecules are not always compared to experimental data<sup>58</sup> and, at times, are not analyzed during the simulated time,<sup>47</sup> that is, a proper populational ensemble characterization.

In contrast to most nucleic acids and proteins, which are linear and have a unique type of linkage, carbohydrates can be branched, linked through one of two anomeric configurations,<sup>4</sup> and there are more than 100 possible monomers (whereas  $\sim 10$  in glycan composition),

which may be connected by different atoms, increasing its structural diversity.<sup>59</sup> All these properties, together with environmental factors, such as hydration and target proteins, determine the conformational ensemble of oligosaccharides. They are very flexible molecules, as their minimum-composing units, disaccharides, can adopt a dynamic equilibrium between two or more distinct conformations. All these properties are associated to the many functional roles of carbohydrates, making them one of the most challenging classes of molecules for conformational analysis.<sup>6</sup> Indeed, they are considered to have several orders of magnitude higher potential information content than any other biological macromolecule.<sup>59</sup> In this context, standard experimental techniques face several challenges when employed to obtain atomic-level structural information about carbohydrates. For example, X-ray crystallography may be difficult on these highly flexible systems and NMR spectroscopy mainly provides time-averaged conformational data.<sup>6</sup> Therefore, molecular modeling techniques, such as MD simulations, emerge as a promising tool for structural and conformational representation of carbohydrates, as they can describe, with a high level of accuracy in both spatial and temporal components, their geometry, flexibility and interaction with target proteins.<sup>10–13,50,57,60,61</sup>

Several force field parameters have been modified and/or designed to model and simulate carbohydrates, such as AMBER'93,<sup>7</sup> AMBER'06,<sup>62</sup> CVFF,<sup>9</sup> and CHARMM.<sup>8</sup> In this context, a previous paper pointed out that different approaches may be better for simulating different aspects of saccharides.<sup>63</sup> For example, the AMBER'93 force field seemed to be better than CHARMM for describing glycan dynamics and interactions with proteins, but less accurate for describing glycan hydration.<sup>4</sup> This may indicate that there is not a definitive force field for carbohydrates MD simulations, whereas further improvements were achieved with AMBER'06. On the other hand, this is the first systematic work assessing glycoprotein dynamics employing GROMOS96 43a1 force field parameters, together with the GROMACS simulation suite, which was shown to adequately represent the carbohydrate (and protein) moieties, as their NMR-derived conformations were adequately reproduced in a series of 50 ns MD simulations. As well, the simulation time scale achieved in this work totalized 0.4  $\mu$ s for glycoproteins MD simulations and 0.2  $\mu$ s for their non-glycosylated counterparts.

Although the modeling and adequate reproduction of glycans conformation are crucial for the adequate description of glycoproteins function, the analysis of N- and O-glycosidic linkages is essential. This may be because this relationship determines the orientation of the whole oligosaccharide relative to the protein and, in turn, influences the functional role of this class of biomacromolecules, such as the exposure of the glycan chain on the cell surface.<sup>38</sup> Among the MD studies on glycoproteins, only a few have explored the geometry of such linkages, including the  $\alpha$ -GalNAc-(1 $\rightarrow$ O)-Ser motif in solution,<sup>64</sup> and an antifreeze glycopeptide,<sup>58</sup> both consisting of O-glycosidic linkages with a carbohydrate moiety bound up to a disaccharide. The current work analyzed and properly described the conformations of an O-glycosidic linkage in a fucosylated glycoprotein, and the conformation of N-glycosidic linkages in four different glycoproteins. Among the parameters studied were its geometry and flexibility, indicating that the employed methodology is adequate for the reproduction of glycoprotein conformational ensembles in solution.

#### 4. Conclusions

It is well known that glycoproteins play important roles in many biological events and that their carbohydrate moieties may be directly involved in such functions, including protein stabiliza-

tion, folding, and interaction with target molecules. However, experimental methods present limitations when applied to carbohydrates. For example, X-ray crystallography is difficult on these highly flexible systems, and NMR tends to provide time-averaged conformations. Therefore, MD simulations emerge as a promising tool for complementing both NMR and X-ray experimental data on glycans and glycoproteins, because it provides a reasonable sampling of the conformational behavior of such molecules in solution.

The employment of molecular modeling techniques may be expected to contribute significantly to the understanding of the many events related to this class of biomacromolecules. Indeed, some force field parameters have been designed for modeling and simulating glycans, being shown to adequately represent experimental geometries. In this context, the GROMOS96 43a1 force field, enhanced by Löwdin HF/6-31G\*\*<sup>65</sup>-derived atomic charges for the glycan moiety, adequately reproduced the conformational ensemble of a series of glycoproteins, as well as the effect of glycosylation over the studied proteins, when compared to NMR data.

Finally, the employed methodology appeared to reproduce adequately the conformational ensemble of a series of glycoproteins containing carbohydrate moieties that are involved in different biological roles. Consequently, upcoming studies using MD simulations upon glycoproteins may be expected to provide a significant contribution to the investigation of their involvement in biological events.

#### Acknowledgments

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq #420015/2005-1 and #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.2008.12.025.

#### References

- Varki, R. A. *Glycobiology* **1993**, *3*, 97–130.
- Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
- Bohne-Lang, A.; von der Lieth, C. W. *Pac. Symp. Biocomput.* **2002**, *7*, 285–296.
- Petrescu, A. J.; Wormald, M. R.; Dwek, R. A. *Curr. Opin. Struct. Biol.* **2006**, *16*, 600–607.
- Petrescu, A. J.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **1999**, *9*, 343–352.
- Woods, R. J. *Glycoconjugate J.* **1998**, *15*, 209–216.
- Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832–3846.
- Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207–221.
- Hwang, M. J.; Ni, X.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *Biopolymers* **1998**, *45*, 435–468.
- Verli, H.; Guimarães, J. A. *Carbohydr. Res.* **2004**, *339*, 281–290.
- Verli, H.; Guimarães, J. A. *J. Mol. Graphics Modell.* **2005**, *24*, 203–212.
- Becker, C. F.; Guimarães, J. A.; Verli, H. *Carbohydr. Res.* **2005**, *340*, 1499–1507.
- Pol-Fachin, L.; Verli, H. *Carbohydr. Res.* **2008**, *343*, 1435–1445.
- 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: Switzerland, 1996.
- IUPAC-IUB Commission on Biochemical Nomenclature, *Pure Appl. Chem.* **1983**, *55*, 1269–1272.
- Schaftenaar, G. MOLDEN. CAOS/CAMM Center, University of Nijmegen, Toernooiveld 1, 6525 ED NIJMEGEN, The Netherlands, 1997.
- Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
- Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
- van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

20. Schuettelkopf, A. W.; van Aalten, D. M. F. *Acta Crystallogr., Sect. D* **2004**, *60*, 1355–1363.
21. Fletcher, C. M.; Harrison, R. A.; Lachmann, P. J.; Neuhaus, D. *Structure* **1994**, *2*, 185–199.
22. 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.
23. 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.
24. Kao, Y. H.; Lee, G. F.; Wang, Y.; Starovasnik, M. A.; Kelley, R. F.; Spellman, M. W.; Lerner, L. *Biochemistry* **1999**, *38*, 7097–7110.
25. Lins, R. D.; Hünenberger, P. H. *J. Comput. Chem.* **2005**, *26*, 1400–1412.
26. Kräutler, V.; Müller, M.; Hünenberger, P. H. *Carbohydr. Res.* **2007**, *342*, 2097–2124.
27. Erbel, P. J. A.; Karimi-Nejad, Y.; De Beer, T.; Boelens, R.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.* **1999**, *260*, 490–498.
28. Withka, J. M.; Wyss, D. F.; Wagner, G.; Arulanandam, A. R.; Reinherz, E. L.; Recny, M. A. *Structure* **1993**, *1*, 69–81.
29. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
30. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
31. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
32. Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
33. Jain, S.; Sundaralingam, M. *J. Biol. Chem.* **1989**, *264*, 12780–12784.
34. Eyal, E.; Gerzon, S.; Potapov, V.; Edelman, M.; Sobolev, V. *J. Mol. Biol.* **2005**, *351*, 431–442.
35. Andrec, M.; Snyder, D. A.; Zhou, Z.; Young, J.; Montelione, G. T.; Levy, R. M. *Proteins* **2007**, *69*, 449–465.
36. Imberty, A.; Pérez, S. *Protein Eng.* **1995**, *8*, 699–709.
37. Petrescu, A. J.; Milac, A. L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **2004**, *14*, 103–114.
38. Ali, M. M.; Aich, U.; Varghese, B.; Pérez, S.; Imberty, A.; Loganathan, D. *J. Am. Chem. Soc.* **2008**, *130*, 8317–8325.
39. Hashimoto, Y.; Toma, K.; Nishikido, J.; Yamamoto, K.; Haneda, K.; Inazu, T.; Valentine, K. G.; Opella, S. J. *Biochemistry* **1999**, *38*, 8377–8384.
40. Sticht, H.; Pickford, A. R.; Potts, J. R.; Campbell, I. D. *J. Mol. Biol.* **1998**, *276*, 177–187.
41. Pickford, A. R.; Smith, S. P.; Staunton, D.; Boyd, J.; Campbell, I. D. *EMBO J.* **2001**, *20*, 1519–1529.
42. Vakonakis, I.; Langenhan, T.; Prömel, S.; Russ, A.; Campbell, I. D. *Structure* **2008**, *16*, 944–953.
43. Hansen, A. P.; Petros, A. M.; Meadows, R. P.; Nettesheim, D. G.; Mazar, A. P.; Olejniczak, E. T.; Xu, R. X.; Pederson, T. M.; Henkin, J.; Fesik, S. W. *Biochemistry* **1994**, *33*, 4847–4864.
44. Barthe, P.; Pujade-Renaud, V.; Breton, F.; Gargani, D.; Thai, R.; Roumestand, C.; de Lamotte, F. *J. Mol. Biol.* **2007**, *367*, 89–101.
45. Shefter, E.; Trueblood, K. N. *Acta Crystallogr.* **1965**, *18*, 1067–1077.
46. Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364–2369.
47. Zuegg, J.; Gready, J. E. *Glycobiology* **2000**, *10*, 959–974.
48. Rubinstein, A.; Kinarsky, L.; Sherman, S. *Int. J. Mol. Sci.* **2004**, *5*, 119–128.
49. Bosques, C. J.; Tschampel, S. M.; Woods, R. J.; Imperiali, B. *J. Am. Chem. Soc.* **2004**, *126*, 8421–8425.
50. Mandal, T. K.; Mukhopadhyay, C. *Biopolymers* **2001**, *59*, 11–23.
51. Wyss, D. F.; Withka, J. M.; Knoppers, M. H.; Sterne, K. A.; Recny, M. A.; Wagner, G. *Biochemistry* **1993**, *32*, 10995–11006.
52. Arulanandam, A. R. N.; Withka, J. M.; Wyss, D. F.; Wagner, G.; Kister, A.; Pallai, P.; Recny, M. A.; Reinherz, E. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11613–11617.
53. Recny, M. A.; Luther, M. A.; Knoppers, M. H.; Neidhardt, E. A.; Klandekar, S. S.; Concino, M. F.; Schimke, P. A.; Francis, M. A.; Moebius, U.; Reinhold, B. B.; Reinhold, V. N.; Reinherz, E. L. *J. Biol. Chem.* **1992**, *267*, 22428–22434.
54. Peterson, A.; Seed, B. *Nature* **1987**, *329*, 842–846.
55. de Beer, T.; van Zuylen, C. W. E. M.; Leeftang, B. R.; Härd, K.; Boelens, R.; Kaptein, R.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.* **1996**, *241*, 229–242.
56. Turnbull, J. E.; Field, R. A. *Nat. Chem. Biol.* **2007**, *3*, 74–77.
57. Mukhopadhyay, C. *Biopolymers* **1998**, *45*, 177–190.
58. Nguyen, D. H.; Colvin, M. E.; Yeh, Y.; Feeny, R. E.; Fink, W. H. *Biophys. J.* **2002**, *82*, 2892–2905.
59. Imberty, A.; Pérez, S. *Chem. Rev.* **2000**, *100*, 4567–4588.
60. Becker, C. F.; Guimarães, J. A.; Mourão, P. A. S.; Verli, H. *J. Mol. Graphics Modell.* **2007**, *26*, 391–399.
61. Naidoo, K. J.; Denysyk, D.; Brady, J. W. *Protein Eng.* **1997**, *10*, 1249–1261.
62. Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622–655.
63. Corzana, F.; Motawia, M. S.; Du Penhoat, C. H.; Pérez, S.; Tschampel, S. M.; Woods, R. J.; Engelsen, S. B. *J. Comput. Chem.* **2003**, *25*, 573–586.
64. Corzana, F.; Busto, J. H.; Jiménez-Osés, G.; Asensio, J. L.; Jiménez-Barbero, J.; Peregrina, J. M.; Avenoza, A. *J. Am. Chem. Soc.* **2006**, *128*, 14640–14648.

#### **4.4 Trabalho III**

Considerando a dificuldade de obtenção 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 conformacional de dissacarídeos, obtidos através de simulações de DM em solução aquosa, como guias para a construção de glicanas compondo glicoproteínas. 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.

Os resultados obtidos indicam a co-existência de múltiplos confôrmeros para os dissacarídeos estudados em solução, especialmente para as ligações glicosídicas  $\beta 1 \rightarrow 4$ , de acordo com dados experimentais e teóricos prévios. Estes estados conformacionais mostraram-se equivalentes aos principais estados conformacionais populados pelas mesmas unidades dissacarídicas quando compondo estruturas glicoprotéicas. Adicionalmente, a prevalência dos diferentes estados conformacionais também foi bastante semelhante entre os dissacarídeos isolados e compondo glicoproteínas, 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.

### **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



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-D-glucosamine (D-GlcNAc), D-mannose (D-Manp), L-fucose (L-Fucp), and D-galactose (D-Galp), 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:

$$\phi = O5-C1-OX-CX'$$

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

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

$$\omega = O6-C6-C5-C4$$

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 D-GlcNAc, <sup>1</sup>C<sub>4</sub> for L-Fucp, <sup>4</sup>C<sub>1</sub> for D-Manp, and <sup>4</sup>C<sub>1</sub> for D-Galp. 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>

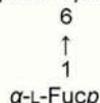
#### 2.1.2. Calculation of energy contours maps

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,

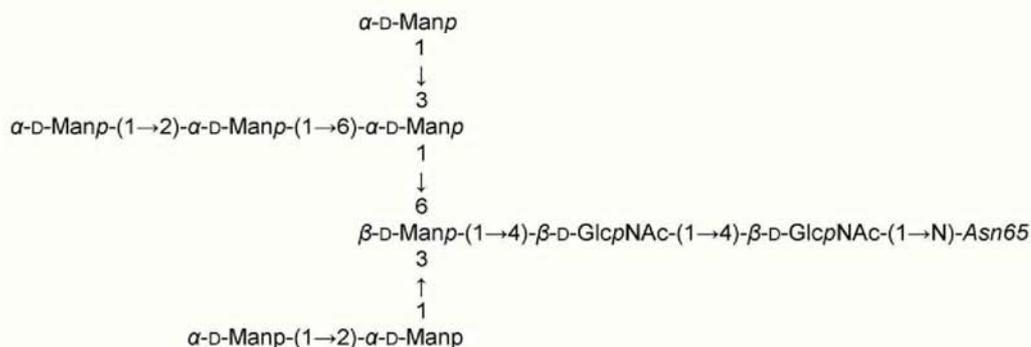
#### (A) Human complement regulatory protein CD59 (PDB ID 1CDS)



#### (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)

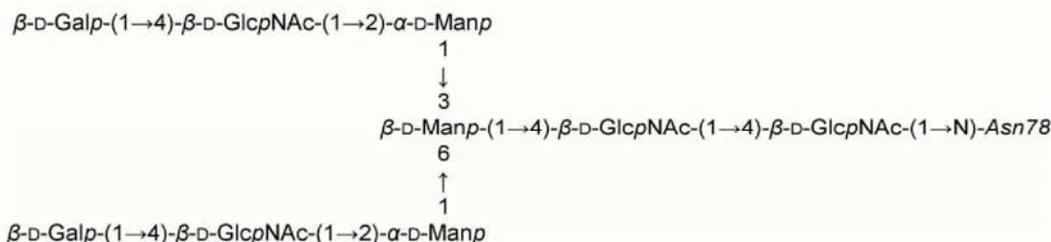


Figure 1. Schematics of the NMR glycan structures used to construct the simulated disaccharides.

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.

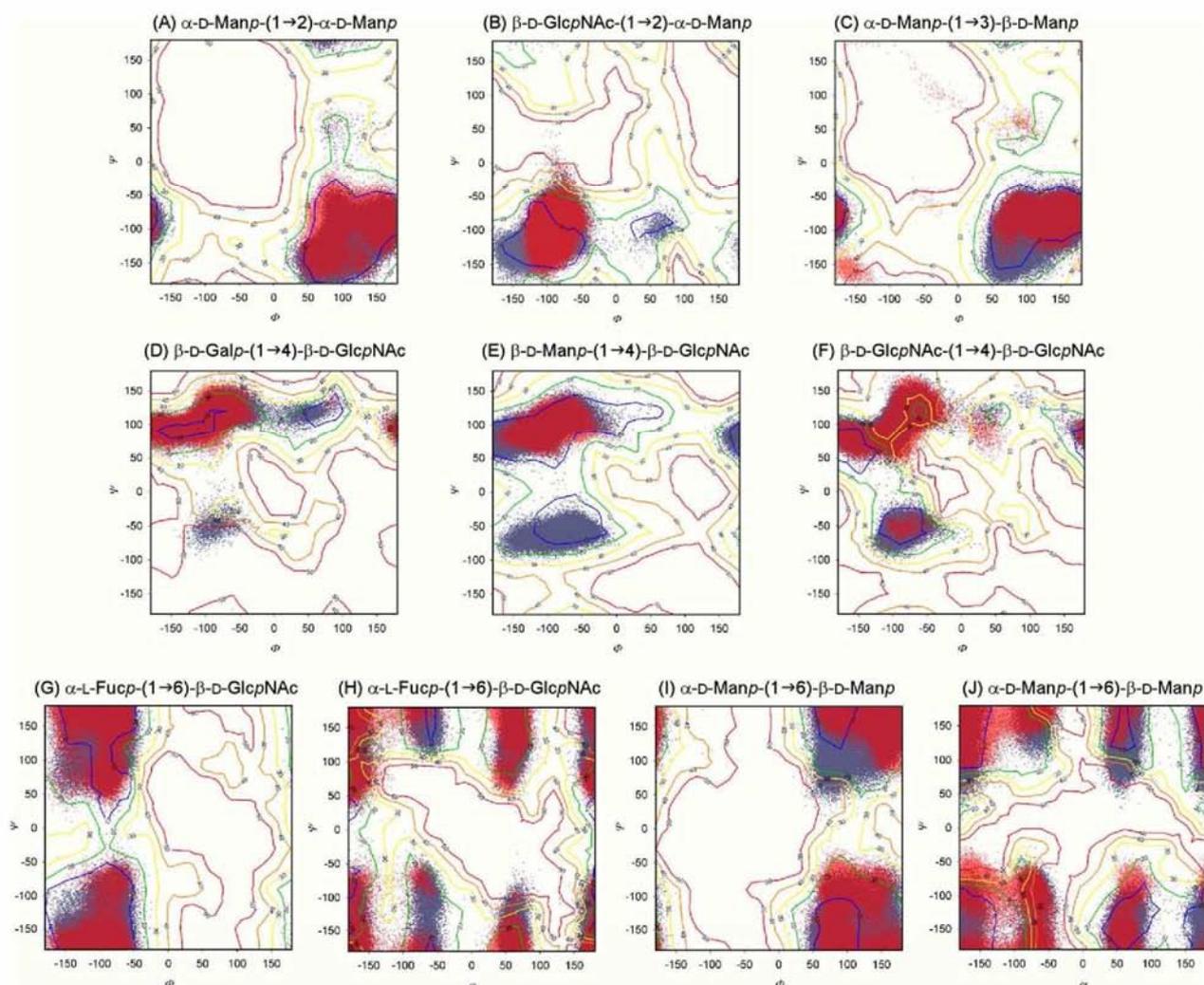
### 2.1.3. MD simulations

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	–	–
$\beta$ -D-Galp-(1→4)-GlcpNAc	<b>44</b>	<b>160 ± 21</b>	<b>-80 ± 21</b>	–
	1.5	<b>-80 ± 29</b>	-50 ± 23	–
$\beta$ -D-Manp-(1→4)-GlcpNAc	<b>86</b>	<b>120 ± 21</b>	–	–
	3	50 ± 29	–	–
$\beta$ -D-GlcpNAc-(1→4)-GlcpNAc	10	-150 ± 17	90 ± 16	–
	23	-70 ± 30	-55 ± 19	–
$\beta$ -D-GlcpNAc-(1→4)-GlcpNAc	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	–
$\alpha$ -D-Fucp-(1→6)-GlcpNAc	15	-100 ± 18	90 ± 14	–
	11	-150 ± 17	–	–
$\alpha$ -D-Fucp-(1→6)-GlcpNAc	14	<b>-90 ± 35</b>	120 ± 27	60 ± 16
	16	-160 ± 32	–	–
$\alpha$ -D-Manp-(1→6)-Manp	12	160 ± 34	-60 ± 17	–
	3	-90 ± 24	–	–
$\alpha$ -D-Manp-(1→6)-Manp	17	90 ± 23	<b>-180 ± 17</b>	–
	<b>38</b>	<b>180 ± 46</b>	–	–
$\alpha$ -D-Manp-(1→6)-Manp	<b>27</b>	<b>80 ± 23</b>	<b>-160 ± 30</b>	<b>60 ± 17</b>
	6	–	110 ± 29	–
$\alpha$ -D-Manp-(1→6)-Manp	6	150 ± 18	-160 ± 31	–
	7	–	110 ± 24	–
$\alpha$ -D-Manp-(1→6)-Manp	5	80 ± 22	-90 ± 21	-70 ± 19
	6	–	170 ± 34	–
$\alpha$ -D-Manp-(1→6)-Manp	1.5	150 ± 19	-90 ± 23	–
	5	–	170 ± 30	–
$\alpha$ -D-Manp-(1→6)-Manp	11	80 ± 24	-125 ± 21	175 ± 19
	11	–	-180 ± 24	–
$\alpha$ -D-Manp-(1→6)-Manp	4	–	90 ± 21	–
	3	150 ± 19	-125 ± 23	–
$\alpha$ -D-Manp-(1→6)-Manp	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.

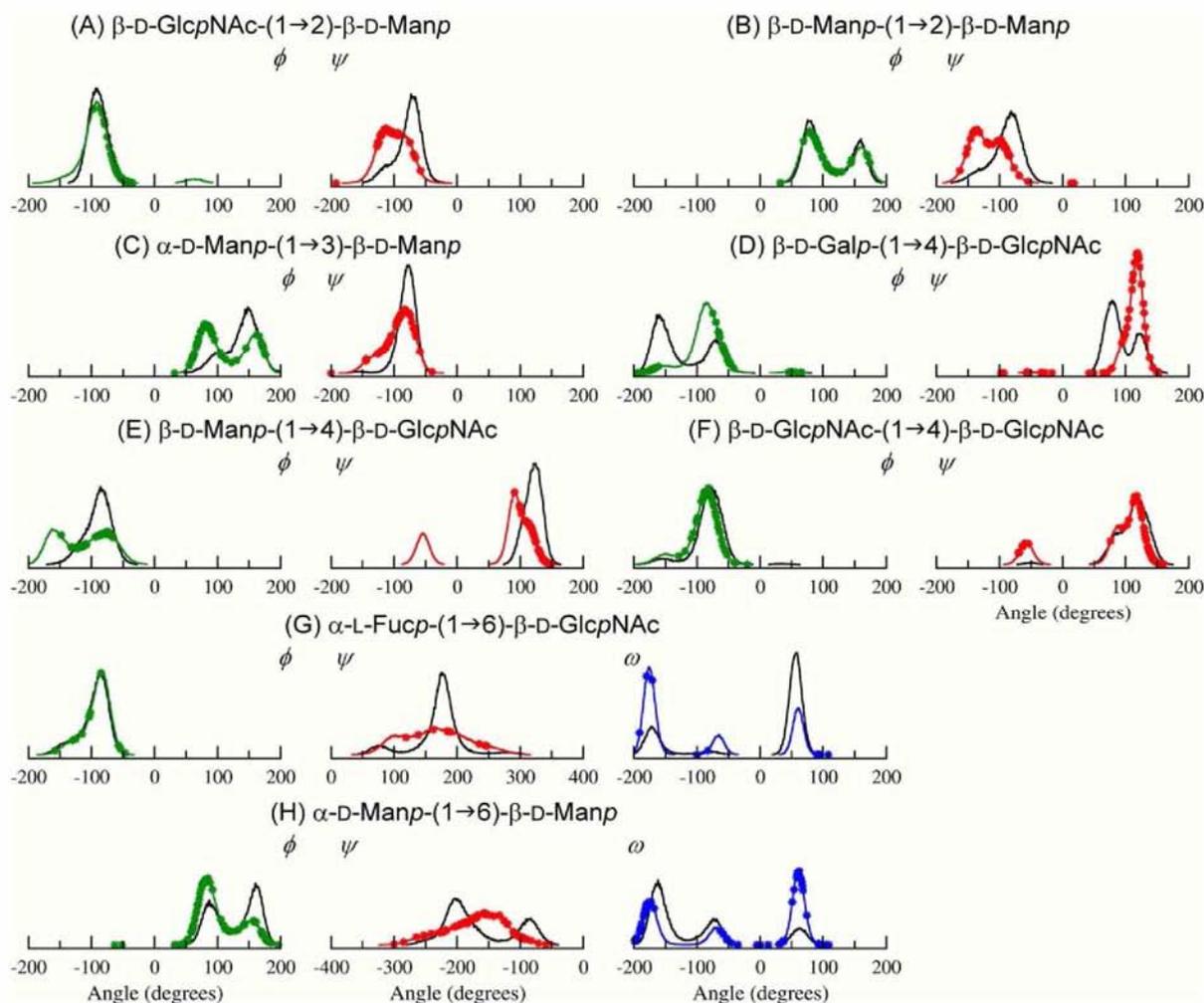
### 3.4. Accuracy in disaccharide use in solution as a conformational reference for building models of glycans in glycoproteins

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-

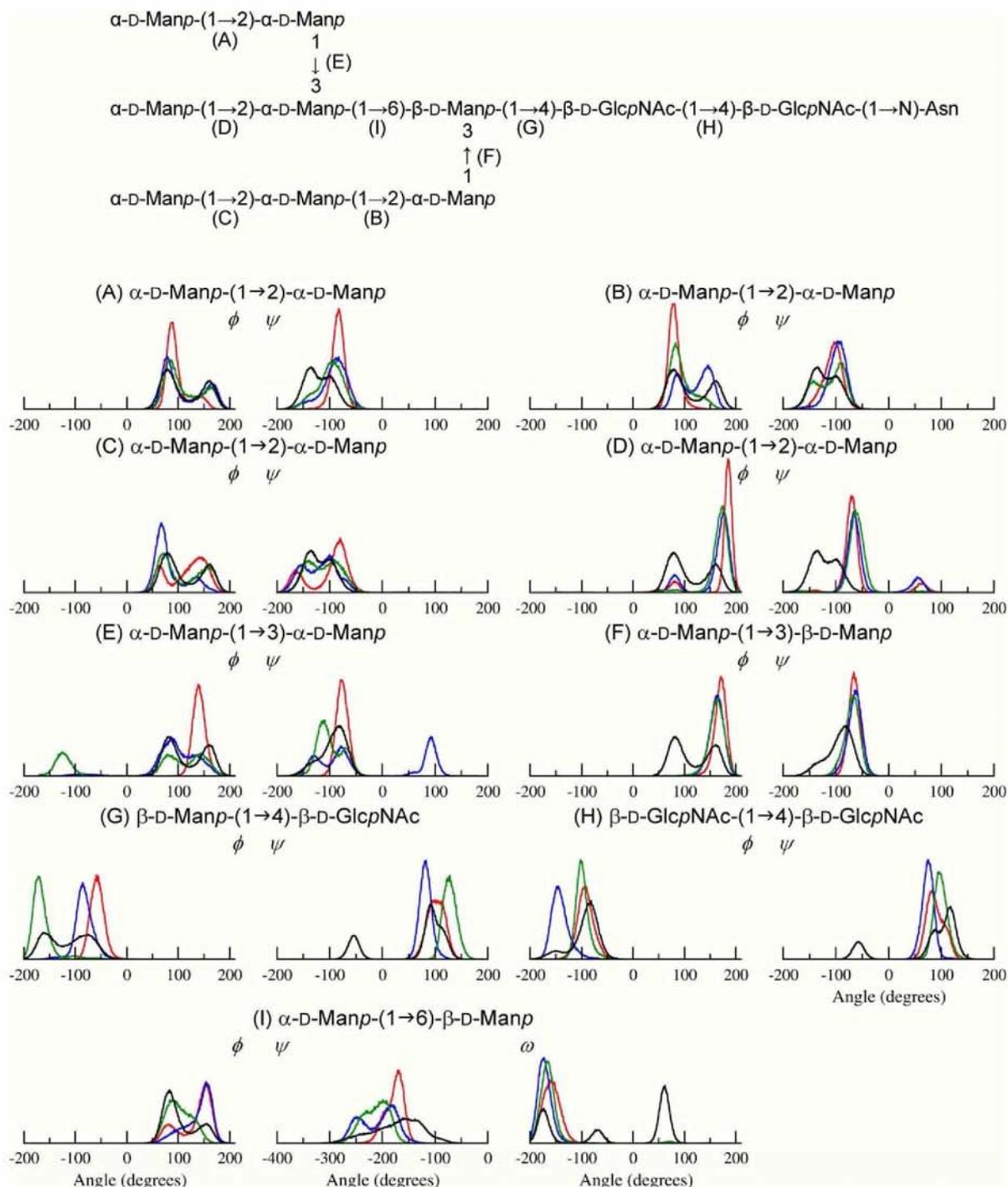
tween the isolated disaccharide and the entire glycoproteins, there were no formations of new conformational peaks, but a modification in the relative abundance of the same conformational state. The major exception was the  $\psi$  angle from  $\alpha$ -D-Manp-(1 $\rightarrow$ 6)- $\beta$ -D-Manp, in 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-

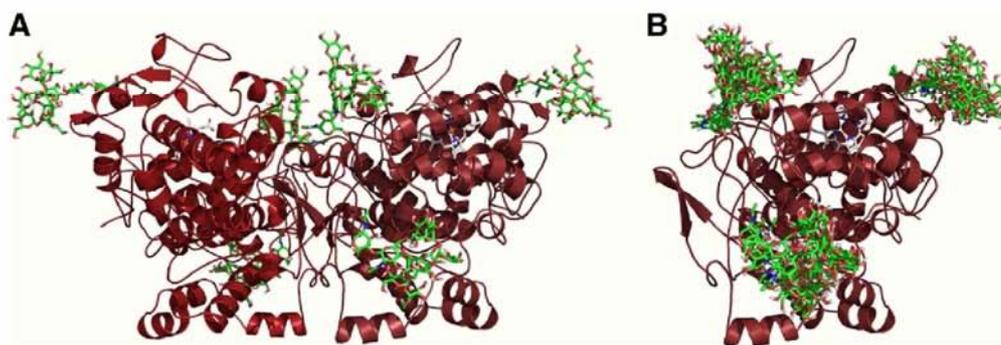
bled in the endoplasmic reticulum.<sup>2,62</sup> In this context, and considering the great therapeutic impact of cyclooxygenase inhibitors, a putative model was built for the glycan moiety of such enzymes (Fig. 4, see Section 2 for details), in agreement with most of the available experimental data 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.; Hünenberger, 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.; Guimaraes, 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.; Pomim, 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.; Lii, 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.; Engels, 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.; Engels, 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.; Verovsky, 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.

## 5 Discussão Geral

### 5.1 Estrutura e conformação de $\alpha$ - e $\beta$ -galactanas

O entendimento da estrutura tridimensional e das propriedades dinâmicas de polissacarídeos constitui-se em um importante pré-requisito para a compreensão das bases moleculares do reconhecimento desses compostos por suas proteínas-alvo (Imberty & Pérez, 2000) e, conseqüentemente, no entendimento de suas funções biológicas. Nesse sentido, poucas estruturas tridimensionais de polissacarídeos sulfatados, estudadas por métodos experimentais como RMN e cristalografia de raios-X, encontram-se disponíveis, tanto livres (Mulloy *et al.*, 1993; Almond *et al.*, 2006) quanto complexados a suas proteínas alvo (Faham *et al.*, 1996; Mikhailov *et al.*, 1996; Mikhailov *et al.*, 1997). Tais estruturas, embora auxiliem no entendimento de interações proteína-ligante e na determinação do arranjo espacial desses carboidratos, consistem em modelos tridimensionais estáticos, incapazes de representar adequadamente as propriedades dinâmicas de tais compostos. Nesse contexto, a MM auxilia na complementação das informações contidas nessas estruturas, descrevendo os carboidratos (e suas possíveis proteínas-alvo) como entidades moleculares flexíveis e dinâmicas, e em soluções mimetizando ambientes fisiológicos.

Dessa forma, baseando-se em parâmetros do campo de força GROMOS96 43a1 e em cargas atômicas de Löwdin obtidas na base HF/6-31G\*\*, estudos envolvendo polissacarídeos sulfatados como heparina (Verli & Guimarães, 2004; Verli & Guimarães, 2005; Pol-Fachin & Verli, 2008), fucanas e galactanas (Becker *et al.*, 2007) vêm sendo realizados pelo nosso grupo de pesquisa. Na presente dissertação, tais protocolos foram aplicados no estudo comparativo de uma  $\alpha$ -galactana de *Echinometra lucunter* e uma  $\beta$ -galactana de *Glyptocidaris crenularis* (Figura 12).

#### 5.1.1 Descrição da geometria das ligações glicosídicas

A diversidade estrutural de carboidratos é considerada maior que a de outras classes de polímeros, tais como ácidos nucléicos e proteínas (Imberty & Pérez, 2000), uma vez que polissacarídeos podem ser ramificados, compostos por mais de cem monossacarídeos conhecidos e podem ligar-se em dois estados anoméricos

distintos (Imberty & Pérez, 2000), através de átomos diferentes (Petrescu *et al.*, 2006). Termos adicionais de diversidade podem ser conferidos a carboidratos através de modificações específicas, tais como sulfatação e acetilação (Petrescu *et al.*, 2006), que podem restringir ou amplificar a variabilidade conformacional de polissacarídeos (Becker *et al.*, 2007; Pol-Fachin & Verli, 2008). Nesse sentido, a fim de avaliar o efeito da presença de grupamentos sulfato sobre a geometria das ligações glicosídicas de  $\alpha$ - e  $\beta$ -galactanas, dissacarídeos sulfatados e não-sulfatados foram submetidos a análise conformacional de suas ligações glicosídicas. Nesse contexto, a construção de mapas de contorno é uma ferramenta útil para descrever a conformação das ligações glicosídicas de carboidratos, avaliando a energia associada a cada conformação gerada pela combinação de diferentes valores de  $\Phi$  e  $\Psi$ , indicando as prováveis conformações mais estáveis para os dissacarídeos estudados (Becker, 2007).

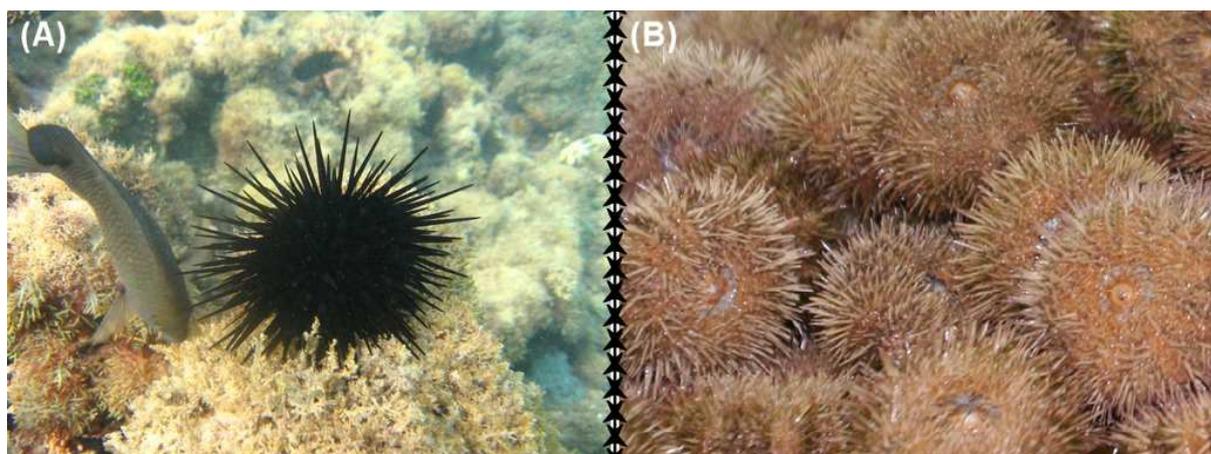


Figura 12: Imagens de espécimes de (A) *Echinometra lucunter* e (B) *Glyptocidaris crenularis*.

A comparação dos mapas de contorno das ligações glicosídicas de dissacarídeos sulfatados e não-sulfatados de  $\alpha$ -galactana indica um aumento do número de regiões de mínimo de energia na molécula não-sulfatada em relação à sulfatada. Da mesma forma, cinco regiões de mínimo de energia podem ser observadas no dissacarídeo não-sulfatado de  $\beta$ -galactana, enquanto apenas duas regiões podem ser observadas nos mapas de energia das ligações glicosídicas dos dissacarídeos sulfatados, indicando um papel dos grupamentos sulfato na promoção

de rigidez em cadeias polissacarídicas, conforme observado em trabalhos prévios do grupo para moléculas semelhantes (Becker *et al.*, 2007).

Estudos baseados em mapas de contorno, embora ofereçam um bom ponto de partida para o entendimento das preferências conformacionais de compostos, se baseiam em cálculos realizados em condições de vácuo ou solvente implícito, ou seja, sem a presença física das moléculas do solvente. Na busca de modelos mais adequados ao entendimento de fenômenos biológicos, uma estratégia promissora envolve a simulação, por DM, de confôrmeros oriundos de mapas de contorno na presença física de moléculas de solvente, ou seja, em condições de solvente explícito. Assim, torna-se possível observar o papel de interações intermoleculares entre soluto e solvente, principalmente ligações de hidrogênio, no comportamento conformacional de biomoléculas e demais compostos de interesse. Adicionalmente, simulações realizadas em temperaturas próximas à fisiológica permitem a observação de transições conformacionais e, potencialmente, a co-existência de múltiplos estados conformacionais em solução, produzindo, portanto, uma imagem mais precisa do comportamento molecular em condições fisiológicas.

Nesse contexto, a adição de solvente explícito na determinação conformacional das moléculas em estudo demonstrou que as conformações populadas pelos dissacarídeos sulfatados e não-sulfatados de  $\alpha$ -galactana em solução aquosa são as mesmas (Figura 13), ou seja, os múltiplos mínimos de energia observados no mapa de contorno da molécula não-sulfatada convergiram para uma mesma região, assim como previamente observado para ligações glicosídicas de heparina (Pol-Fachin & Verli, 2008). Da mesma forma, os diversos mínimos de energia observados nos mapas de energia para os dissacarídeos sulfatados e não-sulfatados de  $\beta$ -galactana convergiram para regiões semelhantes em solução aquosa (Figura 13). No entanto, a geometria das ligações glicosídicas de  $\alpha$ - e  $\beta$ -galactanas são diferentes, especialmente no ângulo  $\Psi$ . Considerando que a atividade biológica desses polissacarídeos é distinta, no que se refere à reação acrossômica de seus organismos, tais divergências conformacionais podem estar envolvidas com esse processo, uma vez que a forma global do polissacarídeo será fundamental no reconhecimento ligante-receptor e, assim, no desenvolvimento de uma resposta biológica. Pode-se observar, portanto, que a solvatação dessas moléculas foi decisiva na determinação conformacional desses dissacarídeos, de maneira a

descartar os mínimos de energia que, apesar de serem estáveis no vácuo, não são prevalentes em solução aquosa, e indicar a provável conformação dos mesmos em seus ambientes naturais, de acordo com dados de RMN (Figura 14).

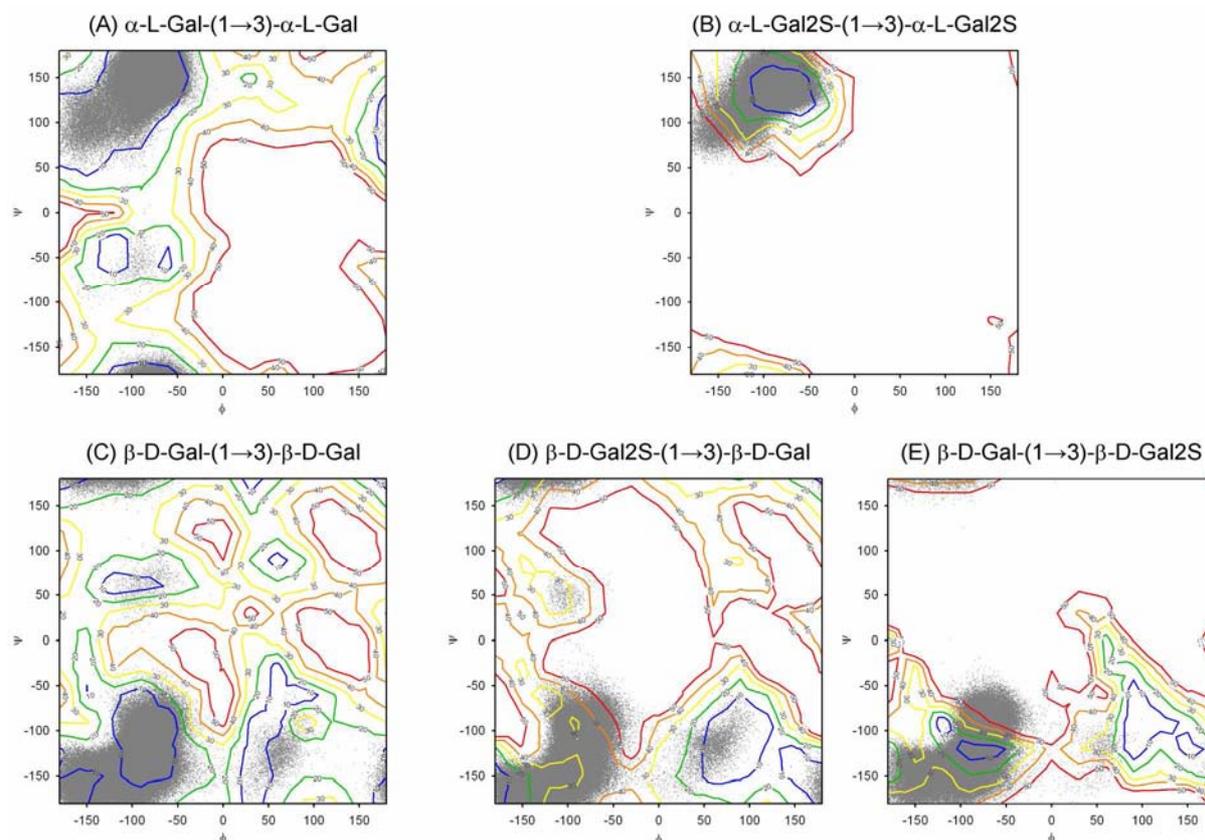


Figura 13: Mapas de contorno para as ligações glicosídicas de  $\alpha$ - e  $\beta$ -galactanas, sulfatadas e não-sulfatadas, sobrepostos aos perfis conformacionais observados nas simulações de DM em solução (pontos em cinza). Os níveis de energia são mostrados a cada  $10 \text{ kJ.mol}^{-1}$ , de  $10$  a  $50 \text{ kJ.mol}^{-1}$ .

Em suma, o refinamento dos mínimos de energia, obtidos de mapas de contorno, por simulações de DM em solução aquosa permitiu a descrição conformacional de dissacarídeos da  $\beta$ -galactana de *Glyptocidaris crenulatis*, de maneira a reproduzir os contatos de NOESY observados para essas moléculas, e auxiliar na determinação da relação entre a conformação e a função biológica de polissacarídeos sulfatados.

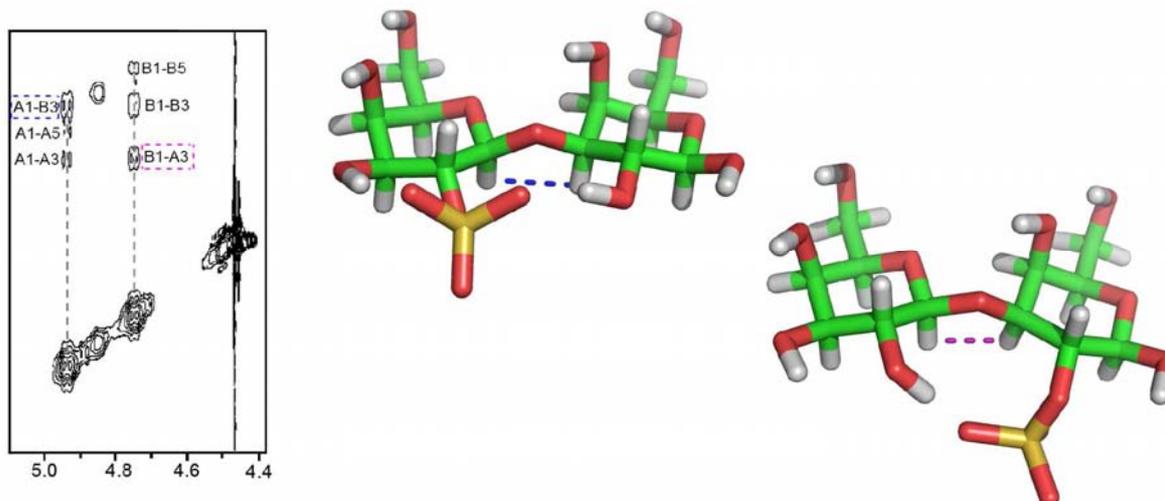


Figura 14: Distâncias entre hidrogênios vicinais entre monossacarídeos, observados através de técnicas de RMN e reproduzidos pelas simulações de DM em solução aquosa para os dissacarídeos de  $\beta$ -galactana.<sup>1</sup>

## 5.2 Reprodução conformacional de glicoproteínas

Os trabalhos envolvendo a descrição conformacional de glicoproteínas, publicados até o presente momento, utilizaram os campos de força AMBER (Renouf & Hounsell, 1993; Woods *et al.*, 1995; Kirschner *et al.*, 2008), CHARMM (Ha *et al.*, 1988) e CVFF (Hwang *et al.*, 1998). Apesar da grande importância de glicoproteínas para vários eventos biológicos (Varki, 1993; Dwek, 1996), poucos são os estudos que avaliaram dinamicamente sua estrutura e conformação (Tabela 3). Nesse sentido, considerando que os parâmetros do campo de força GROMOS96 43a1 se mostraram adequados na reprodução e predição da conformação de carboidratos em geral (Verli & Guimarães, 2004; Verli & Guimarães, 2005; Becker *et al.*, 2005; Becker *et al.*, 2007; Pol-Fachin & Verli, 2008), estruturas de proteínas e glicoproteínas (Tabela 5) determinadas por RMN foram utilizadas como pontos de partida para estudos de DM em solução aquosa a fim de avaliar a capacidade dos parâmetros do campo de força GROMOS96 e das cargas atômicas de Löwdin obtidas na base HF/6-31G\*\* em descrever, de acordo com dados experimentais, a estrutura e conformação dessa classe de moléculas.

<sup>1</sup> Tendo em vista que GROMOS96 é um campo de força do tipo “united-atom” e que, portanto, os hidrogênios vicinais encontram-se unidos aos seus respectivos carbonos apolares, tais átomos foram adicionados a estruturas de dissacarídeos assumindo as conformações mais prevalentes em solução aquosa.

Ao contrário dos estudos conformacionais realizados com dissacarídeos, para os quais protocolo e metodologia já haviam sido estabelecidos em estudos anteriores (Becker *et al.*, 2007; Pol-Fachin & Verli, 2008), estudos envolvendo glicoproteínas não haviam sido realizados até então pelo grupo. Dessa forma, fez-se necessário adicionar novos termos ao campo de força, relacionados com a parte sacarídica de glicoproteínas, assim como parametrizar as N- e O-ligações glicosídicas, envolvendo, portanto, aminoácidos e monossacarídeos.

Nesse contexto, cada monossacarídeo foi descrito junto aos parâmetros do campo de força baseando-se nas topologias obtidas do servidor PRODRG (Schüttelkopf & van Aalten, 2004) e na topologia para resíduos de glicose, já presentes no campo de força. Adicionalmente, as cargas atômicas de Löwdin foram incluídas a tais parâmetros, tanto para os monossacarídeos constituintes da parte sacarídica de glicoproteínas (Becker, 2007), quanto para as cadeias laterais dos aminoácidos que sofreram modificações pós-traducionais, ou seja, asparagina e serina (Figura 15).

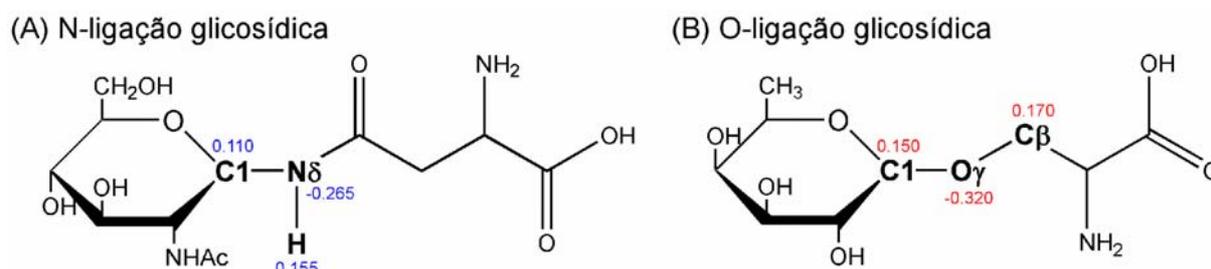


Figura 15: Cargas atômicas obtidas para as ligações glicosídicas entre aminoácidos e monossacarídeos.

A adequada parametrização dos monossacarídeos constituintes da parte sacarídica das glicoproteínas estudadas possibilitou que as simulações de DM fossem estáveis para as quatro glicoproteínas estudadas: o domínio EGF-*like* do fator de coagulação fVII (PDB ID 1FF7) (Kao *et al.*, 1999), a proteína regulatória do sistema complemento (PDB ID 1CDR) (Fletcher *et al.*, 1994), o domínio de adesão do CD2 (PDB ID 1GYA) (Wyss *et al.*, 1995) e a subunidade  $\alpha$  da gonadotrofina coriônica (PDB ID 1HD4) (Erbel *et al.*, 2000), todas humanas. Os resultados obtidos mostraram-se de acordo com dados conformacionais de RMN prévios, tendo sido

avaliadas 836 geometrias de  $\Phi - \Psi - \omega$  para dissacarídeos e 271 geometrias de ligações glicosídicas entre aminoácidos e monossacarídeos.

Sabe-se que a parte sacarídica de uma glicoproteína pode influenciar diversas propriedades da cadeia polipeptídica a que está ligada (Tabela 1). Nesse sentido, a correta descrição conformacional desses oligossacarídeos é essencial no entendimento dos efeitos da glicosilação sobre essas moléculas, tais como na determinação da orientação da cadeia polissacarídica sobre a proteína (Ali *et al.*, 2008). Dessa forma, o perfil conformacional da parte protéica das glicoproteínas estudadas foi comparado ao das suas isoformas não-glicosiladas, por meio de variáveis que incluem o conteúdo de estrutura secundária e a flexibilidade de regiões específicas dessas moléculas, a fim de avaliar o efeito da glicosilação sobre a parte protéica das glicoproteínas estudadas.

Através das simulações por DM, foi possível observar que os oligossacarídeos ligados ao domínio EGF-*like* do fator de coagulação fVII, à proteína regulatória do sistema complemento e à subunidade  $\alpha$  da gonadotrofina coriônica promoveram uma estabilização conformacional em regiões de alça próximas ao sítio de glicosilação ou em regiões espacialmente adjacentes da glicana. Nesse sentido, em um estudo prévio utilizando simulações de DM do fator da coagulação VII completo (Perera *et al.*, 2002), utilizando parâmetros do campo de força AMBER – GLYCAM, efeitos semelhantes da glicosilação sobre o domínio EGF-*like* foram observados. Adicionalmente, o oligossacarídeo do tipo oligomanose ligado ao domínio de adesão do CD2 humano influenciou a flexibilidade de uma série de alças relacionadas à interação dessa proteína com sua proteína alvo, CD58, indicando que a glicosilação, nesse caso, pode estar envolvida com a função da glicoproteína.

Dessa forma, baseando-se em dados estruturais e bioquímicos prévios, foi possível observar que a metodologia empregada se mostrou capaz de descrever a influência da glicosilação sobre a parte protéica de quatro glicoproteínas diferentes. Esses efeitos foram observados independentemente do tipo de glicosilação (N- ou O-) e da estrutura do oligossacarídeo (desde um único monossacarídeo O-ligado até nonassacarídeos N-ligados do tipo oligomanose ou complexo).

### **5.3 Previsão conformacional de oligossacarídeos**

Em decorrência da dificuldade de determinação experimental da geometria e conformação de ligações glicosídicas, a principal fonte de informações estruturais

sobre estes sistemas são cálculos de modelagem molecular, muitos dos quais disponíveis na internet. Estas informações, presentes em bancos de dados como o GLYCOSCIENCES (<http://www.glycosciences.de/>) (Lütteke *et al.*, 2006), baseiam-se em geometrias obtidas para dissacarídeos em mapas de contorno, gerados em condições de vácuo. Tais estruturas são comumente utilizadas na inferência sobre a estrutura de oligossacarídeos, como, por exemplo, na construção da parte sacarídica de glicoproteínas, a fim de gerar um modelo capaz de auxiliar no entendimento dos processos biológicos em que essas moléculas estão envolvidas (Dwek, 1996).

O campo de força MM3 (Allinger *et al.*, 1989; Allinger *et al.*, 1990), utilizado na obtenção desses mapas, embora seja reconhecido por gerar uma representação altamente confiável de aspectos conformacionais de carboidratos (Pérez *et al.*, 1998; Hemmingsen *et al.*, 2004), não possui ampla aplicabilidade a outras classes de moléculas, como proteínas (MacKerell *et al.*, 1998). Nesse sentido, tendo em vista a adequação do protocolo baseado nos parâmetros do campo de força GROMOS96, descrito anteriormente, na reprodução de dados experimentais obtidos por RMN para ambas as porções glicídica e protéica de glicoproteínas, nos voltamos para a busca de estratégias capazes de descrever a estrutura tridimensional dessas moléculas, com destaque para suas cadeias sacarídicas, na ausência de informações experimentais prévias. A estratégia empregada buscou avaliar se conformações obtidas para dissacarídeos, simulados por DM em solução aquosa, se constituiriam em bons pontos de partida para a construção de glicanas complexas e glicoproteínas. Tal procedimento considera que o arcabouço protéico não seria capaz de promover grandes alterações conformacionais na parte sacarídica de glicoproteínas.

Os dissacarídeos avaliados estão contidos nas porções sacarídicas de três glicoproteínas humanas que tiveram sua estrutura determinada por técnicas de RMN: a proteína regulatória do sistema complemento (PDB ID 1CDR) (Fletcher *et al.*, 1994), o domínio de adesão do CD2 (PDB ID 1GYA) (Wyss *et al.*, 1995) e a subunidade  $\alpha$  da gonadotrofina coriônica (PDB ID 1HD4) (Erbel *et al.*, 2000). No total, oito dissacarídeos foram avaliados, incluindo três contendo ligações glicosídicas  $\beta 1 \rightarrow 4$ , dois  $\alpha 1 \rightarrow 6$ , um  $\alpha 1 \rightarrow 2$ , um  $\alpha 1 \rightarrow 3$  e um  $\beta 1 \rightarrow 2$ , que compreendem grande parte das estruturas encontradas em partes sacarídicas N-ligadas a glicoproteínas (do tipo complexo, oligomanose ou híbrido).

Os mapas de contorno obtidos para ligações glicosídicas  $\beta 1 \rightarrow 4$  avaliadas apresentam ampla analogia entre si, incluindo conformações de máximo e mínimo de energia localizadas em regiões similares nos mapas de energia. Da mesma forma, os dissacarídeos Man-( $\alpha 1 \rightarrow 2$ )-Man e Man-( $\alpha 1 \rightarrow 3$ )-Man demonstram essa mesma semelhança, tal como previamente descrito (Ress & Scott, 1971; Sathyanarayana & Rao, 1972; Bluhm *et al.*, 1982; Sathyanarayana & Stevens, 1983). Adicionalmente, de maneira geral, é possível observar que os dissacarídeos contendo ligações  $\alpha$  apresentam um único mínimo de energia principal, enquanto que os dissacarídeos que contem ligações  $\beta$  podem conter dois ou mais mínimos de energia estáveis, o que está de acordo com dados prévios (Lipkind *et al.*, 1984; Lipkind *et al.*, 1985).

Os mapas de contorno presentes no portal GLYCOSCIENCES (Lütteke *et al.*, 2006) possuem definições de RMN para os diedros  $\Phi$ ,  $\Psi$  e  $\omega$  (Tabela 5), ou seja, contendo átomos de hidrogênios vicinais definidos explicitamente. No entanto, o GROMOS96 é um campo de força do tipo “*united-atom*”, em que os átomos de hidrogênio encontram-se unidos aos carbonos apolares a que estariam ligados. Dessa forma, os diedros  $\Phi$ ,  $\Psi$  e  $\omega$  foram definidos de acordo com definições de cristalografia (Tabela 7) para a construção de mapas de contorno utilizando parâmetros do GROMOS96, utilizando, dessa maneira, apenas átomos pesados.

Tabela 7: Definições de RMN e de cristalografia para os diedros componentes de ligações glicosídicas de dissacarídeos.

Diedro	Definição			
	Ligações 1→2, 1→3 e 1→4 <sup>a</sup>		Ligação 1→6	
	RMN <sup>b</sup>	Cristalografia <sup>b</sup>	RMN <sup>b</sup>	Cristalografia <sup>b</sup>
$\Phi$	H1 - C1 - O1 - CX	O5 - C1 - O1 - CX	H1 - C1 - O1 - C6	O5 - C1 - O1 - C6
$\Psi$	C1 - O1 - CX - HX	C1 - O1 - CX - C(X-1)	C1 - O1 - C6 - C5	C1 - O1 - C6 - C5
$\omega$	–	–	O1 - C6 - C5 - H5	O1 - C6 - C5 - C4

<sup>a</sup> Nas definições, X pode ser 2, 3 ou 4, dependendo da ligação glicosídica;

<sup>b</sup> A conversão entre geometrias dessas duas definições pode ser realizada no portal “Torsion Angle Converter for Pyranoses”: <http://www.dkfz-heidelberg.de/spec/ppc/>.

Através da conversão de geometrias obtidas na definição de cristalografia para a definição de RMN, é possível observar que os mínimos de energia dos mapas de

contorno gerados, utilizando parâmetros do campo de força GROMOS96, mostraram-se de acordo com aqueles observados em mapas de energia presentes no portal GLYCOSCIENCES, construídos utilizando o campo de força MM3 (Allinger *et al.*, 1989; Allinger *et al.*, 1990). Esses parâmetros são reconhecidos por gerar uma representação altamente confiável de aspectos conformacionais de carboidratos (Pérez *et al.*, 1998; Hemmingsen *et al.*, 2004), embora a confiabilidade desse campo de força se dê em condições de vácuo (MacKerell, 2004), tendo em vista que os resultados obtidos em estudos avaliando a sua capacidade quando utilizados junto a solvente explícito são menos satisfatórios quando comparados a outros campos de força, tais como CHARMM (MacKerell, 2004). Da mesma forma, os parâmetros MM3 para proteínas parecem não descrever essa classe de moléculas de maneira adequada (MacKerell *et al.*, 1998), o que dificulta a utilização dos dados presentes nos mapas de contorno construídos para dissacarídeos na descrição conformacional dinâmica de glicoproteínas.

A utilização de solvente explícito em simulações de DM permite uma amostragem conformacional que, potencialmente, maximiza as conformações biologicamente relevantes de biomoléculas, ou seja, destaca um dado conjunto de confôrmeros ao invés de outros, mesmo que todos sejam considerados energeticamente equivalentes por cálculos realizados na ausência de solvatação (Yoneda *et al.*, 2006). No contexto de dissacarídeos, a utilização de solvente explícito pode auxiliar na determinação da geometria observada para essas moléculas em seu ambiente natural, ao invés das várias conformações observadas em mapas de contorno para cada ligação glicosídica.

Nesse sentido, assim como observado para os mapas de contorno, as ligações glicosídicas  $\beta 1 \rightarrow 4$  apresentaram perfil conformacional semelhante entre si, da mesma forma que os dissacarídeos Man-( $\alpha 1 \rightarrow 2$ )-Man e Man-( $\alpha 1 \rightarrow 3$ )-Man. Adicionalmente, como previamente estabelecido (Shefter & Trueblood, 1965), o ângulo  $\omega$  apresentou, nos dissacarídeos contendo ligações glicosídicas  $\alpha 1 \rightarrow 6$ , três conformações possíveis, localizadas, aproximadamente, nas geometrias  $60^\circ$ ,  $-180^\circ$  e  $-60^\circ$ . Nesse sentido, foi possível observar que, tanto para Man-( $\alpha 1 \rightarrow 6$ )-Man quanto para Fuc-( $\alpha 1 \rightarrow 6$ )-GlcNAc, dois desses confôrmeros,  $-60^\circ$  e  $-180^\circ$ , apresentam um equilíbrio dinâmico, ou seja, convergem entre si mutuamente, enquanto que o confôrmero de geometria  $60^\circ$  possui um tempo de meia vida maior (Figura 16). Esse comportamento já foi observado em um estudo anterior utilizando outro campo de

força para carboidratos, AMBER – GLYCAM, na qual a definição aplicada para esse diedro, embora também compreendendo apenas átomos pesados, foi diferente da aplicada neste trabalho, que utilizou parâmetros do GROMOS96 (Kirschner & Woods, 2001). Em decorrência desse estudo, em que a orientação do ângulo  $\omega$  foi estudada sem substituintes ligados, uma adequada determinação conformacional e populacional desse diedro não foi reproduzida pelos perfis de energia obtidos em fase gasosa, mas somente quando moléculas de água foram explicitamente incluídas nos cálculos, de maneira que as preferências conformacionais do ângulo  $\omega$  foram atribuídas, portanto, a efeitos de solvatação.

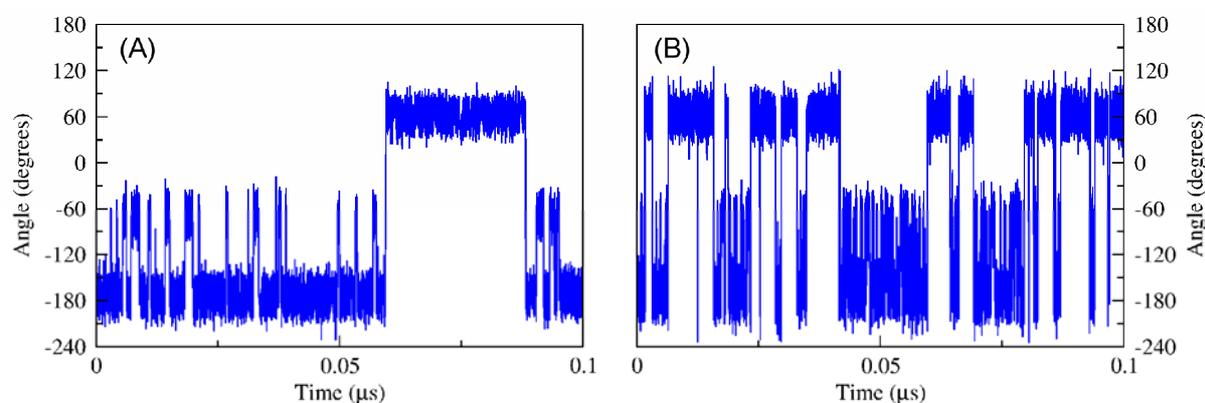


Figura 16: Flutuação do ângulo  $\omega$  nos dissacarídeos (A) Fuc-( $\alpha$ 1 $\rightarrow$ 6)-GlcNAc e (B) Man-( $\alpha$ 1 $\rightarrow$ 6)-Man.

### 5.3.1 Predição da conformação de glicanas a partir de seus dissacarídeos constituintes

À exceção dos ângulos  $\Phi$  de Man-( $\beta$ 1 $\rightarrow$ 4)-GlcNAc e Man-( $\alpha$ 1 $\rightarrow$ 3)-Man e  $\Psi$  de Man-( $\beta$ 1 $\rightarrow$ 4)-GlcNAc que, na simulação dos dissacarídeos isolados em solução, não tiveram sua geometria mais prevalente de acordo com os dados de RMN e com as simulações das glicoproteínas inteiras, todos os demais diedros se mostraram semelhantes nas três metodologias avaliadas. Mesmo nos três ângulos supracitados, a conformação, nos dissacarídeos livres, que se encontra de acordo com a observada pelas técnicas de RMN e pelas simulações de DM das glicoproteínas, e que estão de fato descritas na Tabela 8, é apenas  $\sim$ 10% menos prevalente que a geometria do confôrmero principal. Nesse sentido, os resultados obtidos demonstram que a metodologia empregada para simulações de DM das glicoproteínas foi capaz de reproduzir o perfil conformacional das ligações

glicosídicas da parte sacarídica dessas moléculas, da mesma forma que a geometria dos dissacarídeos, livres em solução, seria capaz de prever a conformação desses oligossacarídeos.

Tabela 8: Comparação da geometria dos ângulos  $\Phi$  e  $\Psi$ , obtidas por DM e RMN, de ligações glicosídicas que compõe a parte sacarídica de glicoproteínas

Ligação Glicosídica	Ângulo	RMN <sup>a</sup>	Dissacarídeos <sup>b</sup>	Glicoproteínas <sup>c</sup>
GlcNAc-( $\beta$ 1 $\rightarrow$ 4)-GlcNAc	$\Phi$	-79 $\pm$ 19	-80 $\pm$ 24	-82 $\pm$ 27
	$\Psi$	133 $\pm$ 16	115 $\pm$ 16	108 $\pm$ 29
Man-( $\beta$ 1 $\rightarrow$ 4)-GlcNAc	$\Phi$	-95 $\pm$ 23	-70 $\pm$ 30	-95 $\pm$ 30
	$\Psi$	126 $\pm$ 15	110 $\pm$ 22	118 $\pm$ 21
Man-( $\alpha$ 1 $\rightarrow$ 3)-Man	$\Phi$	94 $\pm$ 36	80 $\pm$ 24	128 $\pm$ 36
	$\Psi$	-93 $\pm$ 39	-85 $\pm$ 28	-83 $\pm$ 21
Man-( $\alpha$ 1 $\rightarrow$ 6)-Man	$\Phi$	103 $\pm$ 49	80 $\pm$ 23	124 $\pm$ 37
	$\Psi$	-168 $\pm$ 54	-160 $\pm$ 30	-156 $\pm$ 59
GlcNAc-( $\beta$ 1 $\rightarrow$ 2)-Man	$\Phi$	-60 $\pm$ 13	-90 $\pm$ 35	-90 $\pm$ 22
	$\Psi$	-109 $\pm$ 20	-95 $\pm$ 32	-86 $\pm$ 29
Gal-( $\beta$ 1 $\rightarrow$ 4)-GlcNAc	$\Phi$	-104 $\pm$ 57	-80 $\pm$ 29	-114 $\pm$ 44
	$\Psi$	113 $\pm$ 25	120 $\pm$ 21	100 $\pm$ 30
Man-( $\alpha$ 1 $\rightarrow$ 2)-Man	$\Phi$	107 $\pm$ 28	80 $\pm$ 25	116 $\pm$ 39
	$\Psi$	-110 $\pm$ 27	-135 $\pm$ 36	-99 $\pm$ 35
Fuc-( $\alpha$ 1 $\rightarrow$ 6)-GlcNAc	$\Phi$	-98 $\pm$ 37	-90 $\pm$ 35	-109 $\pm$ 22
	$\Psi$	-165 $\pm$ 47	180 $\pm$ 46	169 $\pm$ 65

<sup>a</sup> Dados expressos em média  $\pm$  desvio padrão de todos os valores de RMN medidos;

<sup>b</sup> Picos conformacionais que representam as geometrias de partida para futuros estudos de DM;

<sup>c</sup> Picos conformacionais principais, obtidos de dissacarídeos compondo glicoproteínas;

Em resumo, o comportamento conformacional observado para os dissacarídeos estudados através das simulações de DM em solução aquosa se mostrou de acordo com dados experimentais, da mesma forma que com os perfis obtidos para essas ligações glicosídicas em simulações de DM das glicoproteínas completas. Adicionalmente, a ausência de estruturas determinadas experimentalmente para serem utilizadas como geometria de partida para essas simulações não se mostrou como um fator limitante, uma vez que os mínimos de

energia provenientes dos mapas de contorno dos oito dissacarídeos estudados foram utilizados como estruturas de entrada. Dessa forma, a capacidade da metodologia em reproduzir o perfil conformacional de dissacarídeos qualifica-a para prever a geometria de ligações glicosídicas para os quais não há conhecimento prévio, assim como para serem utilizados na construção da parte sacarídica de glicoproteínas.

#### ***5.4 Implicações para a caracterização conformacional de glicoproteínas e polissacarídeos***

A partir dos resultados obtidos, pode-se destacar a participação de ferramentas de MM, marcadamente a DM utilizando parâmetros do campo de força GROMOS96 e cargas atômicas de Löwdin, no estudo conformacional de glicoproteínas e carboidratos em geral, permitindo a caracterização dessas moléculas em seu ambiente natural (Verli & Guimarães, 2004), complementando técnicas experimentais ou prevendo propriedades até então desconhecidas.

A descrição conformacional de dissacarídeos, utilizando a metodologia de construção de mapas de contorno e subsequente refinamento dos mínimos de energia em solução aquosa, se mostrou de acordo com dados experimentais obtidos em solução (ou seja, através de RMN), possibilitando, inclusive, inferir propriedades funcionais sobre oligossacarídeos, mesmo na ausência de conhecimento conformacional prévio.

Sabe-se das dificuldades em se obter a estrutura de carboidratos por métodos de cristalografia de raios-X, por ocasião da elevada flexibilidade de oligossacarídeos em geral, bem como por RMN, que sugerem apenas um ou uma série de modelos conformacionais que representam propriedades conformacionais médias dessas moléculas. Nesse contexto, as geometrias mais prevalentes, extraídas do perfil conformacional em solução de dissacarídeos, podem ser utilizadas na construção de partes sacarídicas de glicoproteínas, tanto N- quanto O-ligadas, assim como na montagem de outros polissacarídeos de interesse biológico, modelos, esses, que servirão de estruturas de entrada para novos estudos de MM.

Dessa forma, a 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.

## 6 Conclusões

A partir dos objetivos traçados, o presente trabalho permitiu:

- Construção de uma hipótese capaz de explicar o comportamento conformacional e biológico de  $\alpha$ - e  $\beta$ -galactanas, baseada em uma série de mapas de contorno e simulações de DM;
- Reprodução, em mais de 95%, do comportamento conformacional da parte sacarídica de glicoproteínas em solução aquosa, em comparação a dados experimentais de RMN, validando o protocolo empregado para descrição de carboidratos por MM;
- Identificação do potencial relacionado ao refinamento de mínimos de energia obtidos por mapas de contorno em solução aquosa por DM como uma estratégia capaz de agregar qualidade à descrição conformacional de carboidratos, mais próxima do que deve ser esperado em soluções biológicas;
- Consolidação da metodologia de construção de mapas de contorno e refinamento dos mínimos de energia obtidos por DM em solução aquosa como um procedimento adequado para estudos conformacionais de carboidratos em geral;
- Proposição do emprego de estados conformacionais de dissacarídeos em solução como blocos para a construção de glicanas complexas ligadas a glicoproteínas.

Globalmente, os resultados obtidos confirmam o potencial de ferramentas de MM no estudo de sistemas biológicos, especialmente na reprodução e predição de propriedades conformacionais de oligossacarídeos, correlacionando-se com dados experimentais prévios.

## 7 Perspectivas

Considerando-se os procedimentos de simulação propostos e validados no presente trabalho, especialmente os relacionados ao estudo de glicoproteínas, as seguintes perspectivas podem ser traçadas:

- Avaliar o efeito da glicosilação na dinâmica de glicoproteínas relacionadas à cascata de coagulação sanguínea, tais como antitrombina, trombina e fator Xa;
- Construção de mapas de contorno e caracterização conformacional de dissacarídeos componentes da heparina, de maneira a utilizar as geometrias mais prevalentes observadas em solução aquosa para essas ligações glicosídicas na construção dos diversos polissacarídeos que compõe este medicamento;
- Caracterização da interação de um heparinóide com glicoproteínas da cascata de coagulação sanguínea, destacadamente o complexo ternário formado por antitrombina, heparina e trombina/fXa.

## 8 Referências Bibliográficas

### Referências citadas no texto e nos artigos publicados

- Abu-Qarn, M.; Eichler, J.; Sharon, N.: Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Curr. Opin. Struct. Biol.*, **2008**, *18*, 544-550.
- Albano, R. M.; Mourão, P. A. S.: Isolation, fractionation and preliminary characterization of a novel class of sulfated glycans from the tunic of *Styela plicata* (Chordata-Tunicata). *J. Biol. Chem.*, **1986**, *261*, 758-765.
- Ali, M. M.; Aich, U.; Varghese, B.; Pérez, S.; Imberty, A.; Loganathan, D.: Conformational Preferences of the Aglycon Moiety in Models and Analogs of GlcNAc-Asn Linkage: Crystal Structures and ab Initio Quantum Chemical Calculations of N-( $\beta$ -D-Glycopyranosyl)haloacetamides. *J. Am. Chem. Soc.*, **2008**, *130*, 8317-8325.
- 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.
- Almond, A.; Deangelis, P. L.; Blundell, C. D.: Hyaluronan: the local solution conformation determined by NMR and computer modeling is close to a contracted left-handed 4-fold helix. *J. Mol. Biol.*, **2006**, *358*, 1256-1269.
- Alves, A. P.; Mulloy, B.; Diniz, J. A.; Mourão, P. A. S.: Sulfated polysaccharides from the egg jelly are species-specific inducers of acrosomal reaction in sperm of sea urchins. *J. Biol. Chem.*, **1997**, *272*, 6965-6971.
- Alves, A. P.; Mulloy, B.; Moy, G. W.; Vacquier, V. D.; Mourão, P. A. S.: Females of the sea urchin *Strongylocentrotus purpuratus* differ in the structure of their egg jelly sulfated fucans. *Glycobiology*, **1998**, *8*, 939-946.
- Amornrut, C.; Toida, T.; Imanari, T.; Woo, E. R.; Park, H.; Linhardt, R.; Wu, Y. S.; Kim, Y. S.: A new sulfated  $\beta$ -galactan from clams with anti-HIV activity. *Carbohydr. Res.*, **1999**, *321*, 121-127.
- Andrec, M.; Snyder, D. A.; Zhou, Z.; Young, J.; Montelione, G. T.; Levy, R. M.: A large data set comparison of protein structures determined by crystallography

- and NMR: Statistical test for structural differences and the effect of crystal packing. *Proteins*, **2007**, *69*, 449-465.
- Arulanandam, A. R. N.; Withka, J. M.; Wyss, D. F.; Wagner, G.; Kister, A.; Pallai, P.; Recny, M. A.; Reinherz, E. L.: The CD58 (LFA-3) binding site is a localized and highly charged surface area on the AGFCC'C" face of the human CD2 adhesion domain. *Proc. Natl. Acad. Sci. U.S.A.*, **1993**, *90*, 11613-11617.
- Banner, D. W.; D'Arcy, A.; Chène, C.; Winkler, F. K.; Guha, A.; Konigsberg, W. H.; Nemerson, Y.; Kirchhofer, D.: The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature*, **1996**, *380*, 41-46.
- Barthe, P.; Pujade-Renaud, V.; Breton, F.; Gargani, D.; Thai, R.; Roumestand, C.; de Lamotte, F.: Structural Analysis of Cassiicolin, a Host-selective Protein Toxin from *Corynespora cassiicola*. *J. Mol. Biol.*, **2007**, *367*, 89-101.
- 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.
- Becker, C. F.; Guimarães, J. A.; Mourão, P. A. S.; Verli, H.: Conformation of sulfated galactan and sulfated fucan in aqueous solutions: Implications to their anticoagulant activities. *J. Mol. Graph. Model.*, **2007**, *26*, 391-399.
- Becker, C. F.: *Dissertação de Mestrado*. Universidade Federal do Rio Grande do Sul, Brasil, **2007**.
- 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.
- Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P.: The missing term in effective pair potentials. *J. Phys. Chem.*, **1987**, *91*, 6269-6271.
- Berman, H. M.; Henric, K.; Nakamura, H.; Markley, J.: Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.*, **2007**, *14*, 354-355.
- Begeot, M.; Hemming, F. J.; Dubois, P. M.; Combamous, Y.; Dubois, M. P.; Aubert, M. L.: Induction of pituitary lactotrope differentiation by luteinizing hormone  $\alpha$  subunit. *Science*, **1984**, *226*, 566-568.
- Biermann, C. H.; Marks, J. A.; Vilela-Silva, A. C. E. S.; Castro, M. O.; Mourão, P. A. S.: Carbohydrate-based species recognition in sea urchin fertilization: another avenue for speciation? *Evol. Dev.*, **2004**, *6*, 353-361.

- Bilan, M. I.; Vinogradova, E. V.; Shashkov, A. S.; Usov, A. I.: Structure of a highly pyruvylated galactan sulfate from the Pacific green alga *Codium yezoense* (Bryopsidales, Chlorophyta). *Carbohydr. Res.*, **2007**, *342*, 586-596.
- Blithe, D. L.; Iles, R. K.: The Role of Glycosylation in Regulating the Glycoprotein Hormone Free  $\alpha$ -Subunit and Free  $\beta$ -Subunit Combination in the Extraembryonic Coelomic Fluid of Early Pregnancy. *Endocrinology*, **1995**, *136*, 903-910.
- 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.
- Bodian, D. L.; Jones, E. Y.; Harlos, K.; Stuart, D. I.; Davis, S. J.: Crystal structure of the extracellular region of the human cell adhesion molecule CD2 at 2.5 Å resolution. *Structure*, **1994**, *2*, 755-766.
- 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.
- Bosques, C. J.; Tschampel, S. M.; Woods, R. J.; Imperiali, B.: Effects of Glycosylation on Peptide Conformation: A Synergistic Experimental and Computational Study. *J. Am. Chem. Soc.*, **2004**, *126*, 8421-8425.
- Brant, D. A.: Shapes and motions of polysaccharide chains. *Pure Appl. Chem.*, **1997**, *69*, 1885-1892.
- Butenas, S.; Mann, K. G.: Kinetics of human factor VII activation. *Biochemistry*, **1996**, *35*, 1904-1910.
- Case, D. A.; Cheatham, T. E. 3<sup>rd</sup>; 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.
- Choay, J.; Petitou, M.; Lormeau, J. C.; Sinayĭ, P.; Casu, B.; Gatti, G.: Structure-activity relationship in heparin: A synthetic pentasaccharide with high affinity for antithrombin III and eliciting high anti-factor Xa activity. *Biochem. Biophys. Res. Commun.*, **1983**, *116*, 492-499.
- Cipollo, J. F.; Awad, A. M.; Costello, C. E.; Hirschberg, C. B.: N-Glycans of *Caenorhabditis elegans* Are Specific to Developmental Stages. *J. Biol. Chem.*, **2005**, *280*, 26063-26072.

- 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.
- Corzana, F.; Motawia, M. S.; Du Penhoat, C. H.; Pérez, S.; Tschampel, S. M.; Woods, R. J.; Engelsens, S. B.: A Hydration Study of (1→4) and (1→6) Linked  $\alpha$ -Glucans by Comparative 10 ns Molecular Dynamics Simulations and 500-MHz NMR. *J. Comput. Chem.*, **2003**, *25*, 573-586.
- Corzana, F.; Busto, J. H.; Jiménez-Osés, G.; Asensio, J. L.; Jiménez-Barbero, J.; Peregrina, J. M.; Avenoza, A.: New Insights into  $\alpha$ -GalNAc–Ser Motif: Influence of Hydrogen Bonding versus Solvent Interactions on the Preferred Conformation. *J. Am. Chem. Soc.*, **2006**, *128*, 14640-14648.
- Couldrey, C.; Green, J. E.: Metastases: the glycan connection. *Breast Cancer Res.*, **2000**, *2*, 321-323.
- 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.
- Danielsson, A.; Raub, E.; Lindahl, U.; Björk, I.: Role of ternary complexes, in which heparin binds both antithrombin and proteinase, in the acceleration of the reactions between antithrombin and thrombin of factor Xa. *J. Biol. Chem.*, **1986**, *261*, 15467-15473.
- Darden, T.; York, D.; Pedersen, L.: Particle Mesh Ewald – an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **1993**, *98*, 10089-10092.
- de Beer, T.; Vliegthart, J. F.; Löffler, A.; Hofsteenge, J.: The Hexopyranosyl Residue That Is C-Glycosidically Linked to the Side Chain of Tryptophan-7 in Human RNase U<sub>s</sub> Is  $\alpha$ -Mannopyranose. *Biochemistry*, **1995**, *34*, 11785-11789.
- de Beer, T.; van Zuylen, C. W. E. M.; Leeftang, B. R.; Härd, K.; Boelens, R.; Kaptein, R.; Kamerling, J. P.; Vliegthart, J. F. G.: NMR studies of the free  $\alpha$  subunit of human chorionic gonadotropin: Structural influences of N-glycosylation and the  $\beta$  subunit on the conformation of the  $\alpha$  subunit. *Eur. J. Biochem.*, **1996**, *241*, 229-242.

- de Groot, B. L.; Grubmüller, H.: Water Permeation Across Biological Membranes: Mechanism and Dynamics of Aquaporin-1 and GlpF. *Science*, **2001**, *294*, 2353-2357.
- 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>
- Doucey, M. A.; Hess, D.; Cacan, R.; Hofsteenge, J.: Protein C-Mannosylation Is Enzyme-catalyzed and Uses Dolichyl-Phosphate-Mannose as Precursor. *Mol. Biol. Cell*, **1998**, *9*, 291-300.
- Dubois, M.; Gilles, K. A.; Hamilton, J. K.; Rebers, P. A.; Smith, F.: Colorimetric method for determination of sugars and related substances. *Anal. Chem.*, **1956**, *28*, 350-354.
- Dwek, R. A.: Glycobiology: Toward Understanding the Function of Sugars. *Chem. Rev.*, **1996**, *96*, 683-720.
- Erbel, P. J. A.; Karimi-Nejad, Y.; De Beer, T.; Boelens, R.; Kamerling, J. P.; Vliegthart, J. F. G.: Solution structure of the  $\alpha$ -subunit of human chorionic gonadotropin. *Eur. J. Biochem.*, **1999**, *260*, 490-498.
- 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.
- Eyal, E.; Gerzon, S.; Potapov, V.; Edelman, M.; Sobolev, V.: The Limit of Accuracy of Protein Modeling: Influence of Crystal Packing on Protein Structure. *J. Mol. Biol.*, **2005**, *351*, 431-442.
- Faham, S.; Hileman, R. E.; Fromm, J. R.; Linhardt, R. J.; Rees, D. C.: Heparin structure and interactions with basic fibroblast growth factor. *Science*, **1996**, *276*, 1116-1120.
- Farndale, R. W.; Buttle, D. J.; Barrett, A. J.: Improved quantitation and discrimination of sulphated glycosaminoglycans by use of dimethylmethylene blue. *Biochim. Biophys. Acta*, **1986**, *883*, 173-177.
- Farias, E. H.; Pomin, V. H.; Valente, A. P.; Nader, H. B.; Rocha, H. A.; Mourão, P. A. S.: A preponderantly 4-sulfated, 3-linked galactan from the green alga *Codium isthmocladum*. *Glycobiology*, **2008**, *18*, 250-259.

- Farkas, I.; Baranyi, L.; Ishikawa, Y.; Okada, N.; Bohata, C.; Budai, D.; Fukuda, A.; Imai, M.; Okada H.: CD59 blocks not only the insertion of C9 into MAC but inhibits ion channel formation by homologous C5b-8 as well as C5b-9. *J. Physiol.*, **2002**, *539*, 537-545.
- Fincher, G. B.: Exploring the evolution of (1,3;1,4)- $\beta$ -D-glucans in plant cell walls: comparative genomics can help! *Curr. Opin. Plant Biol.*, **2009**, *12*, 140-147.
- 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.
- Fukuda, M.: Leukosialin, a major O-glycan-containing sialoglycoprotein defining leukocyte differentiation and malignancy. *Glycobiology*, **1991**, *1*, 347-356.
- 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.
- Gama, C. I.; Hsieh-Wilson, L. C.: Chemical approaches to deciphering the glycosaminoglycan code. *Curr. Opin. Chem. Biol.*, **2005**, *9*, 609-619.
- Gavel, Y.; von Heijne, G.: Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Eng.*, **1990**, *3*, 433-442.
- Gerwig, G. J.; Kamerling, J. P.; Vliegthart, J. F. G.: Determination of the absolute configuration of monosaccharides in complex carbohydrates by capillary G.L.C. *Carbohydr. Res.*, **1979**, *77*, 10-17.
- Gupta, R.; Birch, H.; Rapacki, K.; Brunak, S.; Hansen, J. E.: O-GLYCBASE version 4.0: a revised database of O-glycosylated proteins. *Nucleic Acids Res.*, **1999**, *27*, 370-372.
- Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W.: A revised potential-energy surface for molecular mechanics studies of carbohydrates. *Carbohydr. Res.*, **1988**, *180*, 207-221.
- Hagen, F. S.; Gray, C. L.; O'Hara, P.; Grant, F. J.; Saari, G. C.; Woodbury, R. G.; Hart, C. E.; Insley, M.; Kisiel, W.; Kurachi, K.; Davie, E. W.: Characterization of a cDNA coding for human factor VII. *Proc. Natl. Acad. Sci. U. S. A.*, **1986**, *83*, 2412-2416.

- Handel, T. M.; Johnson, Z.; Crown, S. E.; Lau, E. K.; Proudfoot, A. E.: Regulation of protein function by glycosaminoglycans--as exemplified by chemokines. *Annu. Rev. Biochem.*, **2005**, *74*, 385-410.
- Hansen, A. P.; Petros, A. M.; Meadows, R. P.; Nettesheim, D. G.; Mazar, A. P.; Olejniczak, E. T.; Xu, R. X.; Pederson, T. M.; Henkin, J.; Fesik, S. W.: Solution structure of the amino-terminal fragment of urokinase-type plasminogen activator. *Biochemistry*, **1994**, *33*, 4847-4864.
- Hanson, S. R.; Culyba, E. K.; Hsu, T. L.; Wong, C. H.; Kelly, J. W.; Powers, E. T.: The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. *Proc. Natl. Acad. Sci. U. S. A.*, **2009**, *106*, 3131-3136.
- Hansson, K.; Stenflo, J.: Post-translational modifications in proteins involved in blood coagulation. *J. Thromb. Haemost.*, **2005**, *3*, 2633-2648.
- 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.
- Hashimoto, Y.; Toma, K.; Nishikido, J.; Yamamoto, K.; Haneda, K.; Inazu, T.; Valentine, K. G.; Opella, S. J.: Effects of Glycosylation on the Structure and Dynamics of Eel Calcitonin in Micelles and Lipid Bilayers Determined by Nuclear Magnetic Resonance Spectroscopy. *Biochemistry*, **1999**, *38*, 8377-8384.
- Helenius, A.; Aebi, M.: Intracellular Functions of N-Linked Glycans. *Science*, **2001**, *291*, 2364-2369.
- 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.
- 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.
- Hirai, E.: On the breeding seasons of invertebrates in the neighbourhood of the marine biological station of Asamushi. *Sci. Rep. Tohoku Univ. Fourth Ser. (Biol.)*, **1963**, *29*, 369-375.

- Hirohashi, N.; Vacquier, V. D.: High Molecular Mass Egg Fucose Sulfate Polymer Is Required for Opening Both  $\text{Ca}^{2+}$  Channels Involved in Triggering the Sea Urchin Sperm Acrosome Reaction. *J. Biol. Chem.*, **2002a**, 277, 1182-1189.
- Hirohashi, N.; Vacquier, V. D.: Egg Sialoglycans Increase Intracellular pH and Potentiate the Acrosome Reaction of Sea Urchin Sperm. *J. Biol. Chem.*, **2002b**, 277, 8041-8047.
- Hirohashi, N.; Vacquier, V. D.: Egg fucose sulfate polymer, sialoglycan, and speract all trigger the sea urchin sperm acrosome reaction. *Biochem. Biophys. Res. Commun.*, **2002c**, 296, 833-839.
- Hirohashi, N.; Vilela-Silva, A. C. E. S.; Mourão, P. A. S.; Vacquier, V. D.: Structural requirements for species-specific induction of the sperm acrosome reaction by sea urchin egg sulfated fucan. *Biochem. Biophys. Res. Commun.*, **2002d**, 298, 403-407.
- Hirohashi, N.; Vacquier, V. D.: Store-operated calcium channels trigger exocytosis of the sea urchin sperm acrosomal vesicle. *Biochem. Biophys. Res. Commun.*, **2003**, 304, 285-292.
- Hofsteenge, J.; Müller, D. R.; de Beer, T.; Löffler, A.; Richter, W. J.; Vliegthart, J. F.: New type of linkage between a carbohydrate and a protein: C-glycosylation of a specific tryptophan residue in human RNase U<sub>s</sub>. *Biochemistry*, **1994**, 33, 13524-13530.
- Huang, X.; Barchi, J. J. Jr.; Lung, F. D.; Roller, P. P.; Nara, P. L.; Muschik, J.; Garrity, R. R.: Glycosylation affects both the three-dimensional structure and antibody binding properties of the HIV-1IIIB GP120 peptide RP135. *Biochemistry*, **1997**, 36, 10846-10856.
- Humphrey, W.; Dalke, A.; Schulten, K.: VMD – Visual Molecular Dynamics. *J. Mol. Graph.*, **1996**, 14, 33-38. <http://www.ks.uiuc.edu/Research/vmd/>.
- Hunt, L. T.; Dayhoff, M. O.: The Occurrence in proteins of the tripeptides Asn-X-Ser and Asn-X-Thr and of bound carbohydrate. *Biochem. Biophys. Res. Commun.*, **1970**, 39, 757-765.
- Hwang, M. J.; Ni, X.; Waldman, M.; Ewig, C. S.; Hagler, A. T.: Derivation of class II force fields. VI. Carbohydrate compounds and anomeric effects. *Biopolymers*, **1998**, 45, 435-468.

- 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.
- Imberty, A.; Pérez, S.: Stereochemistry of the N-glycosylation sites in glycoproteins. *Protein Eng.*, **1995**, *8*, 699-709.
- Imberty, A.; Pérez, S.: Structure, conformation, and dynamics of bioactive oligosaccharides: theoretical approaches and experimental validations. *Chem. Rev.*, **2000**, *100*, 4567–4588.
- Imberty, A.; Lortat-Jacob, H.; Pérez, S.: Structural view of glycosaminoglycan–protein interactions. *Carbohydr. Res.*, **2007**, *342*, 430-439.
- Imperiali, B.; Hendrickson, T. L.: Asparagine-linked glycosylation: specificity and function of oligosaccharyl transferase. *Bioorg. Med. Chem.*, **1995**, *3*, 1565-1578.
- Imperiali, B.: Protein Glycosylation: The Clash of the Titans. *Acc. Chem. Res.*, **1997**, *30*, 452-459.
- IUPAC-IUB Commission on Biochemical Nomenclature: Conformational nomenclature for five- and six-membered ring forms of monosaccharides and their derivatives (Recommendations 1980). *Eur. J. Biochem.*, **1980**, *111*, 295-298.
- IUPAC-IUB Commission on Biochemical Nomenclature: Symbols for specifying the conformation of polysaccharide chains. *Pure Appl. Chem.*, **1983**, *55*, 1269-1272.
- IUPAC-IUB Commission on Biochemical Nomenclature: Nomenclature for carbohydrates. *Pure Appl. Chem.*, **1996**, *68*, 1919-2008.
- Jaeken, J.; Carchon, H.: The carbohydrate-deficient glycoprotein syndromes: An overview. *J. Inher. Metab. Dis.*, **1993**, *16*, 813-820.
- Jacques, L. B.: Heparin: an old drug with a new paradigm. *Science*, **1979**, *206*, 528-533.
- Jain, S.; Sundaralingam, M.: Effect of Crystal Packing Environment on Conformation of the DNA Duplex. *J. Biol. Chem.*, **1989**, *264*, 12780-12784.
- 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.

- Julenius, K.: NetCGlyc 1.0: prediction of mammalian C-mannosylation sites. *Glycobiology*, **2007**, *17*, 868-876.
- Kabat, E. A.; Mayer, M. M.: *Experimental Immunochemistry*, 2<sup>nd</sup> Ed., Charles C. Thomas Publisher, Springfield, **1971**, 505-513.
- Kao, Y. H.; Lee, G. F.; Wang, Y.; Starovasnik, M. A.; Kelley, R. F.; Spellman, M. W.; Lerner, L.: The Effect of O-Fucosylation on the First EGF-like Domain from Human Blood Coagulation Factor VII. *Biochemistry*, **1999**, *38*, 7097-7110.
- Karplus, M.; Petsko, G. A.: Molecular dynamics simulations in biology. *Nature*, **1990**, *347*, 631-639.
- Kessler, M. J.; Reddy, M. S.; Shah, R. H.; Bahl, O. P.: Structures of N-glycosidic carbohydrate units of human chorionic gonadotropin. *J. Biol. Chem.*, **1979a**, *254*, 7901-7908.
- Kessler, M. J.; Mise, T.; Ghai, R. D.; Bahl, O. P.: Structure and location of the O-glycosidic carbohydrate units of human chorionic gonadotropin. *J. Biol. Chem.*, **1979b**, *254*, 7909-7914.
- Kimberley, F. C.; Sivasankar, B.; Morgan, P. B.: Alternative roles for CD59. *Mol. Immunol.*, **2007**, *44*, 73-81.
- Kircher, H. W.: Gas-liquid chromatography of methylated sugars. *Anal. Chem.*, **1960**, *32*, 1103-1106.
- Kirschner, K. N.; Woods, R. J.: Solvent interactions determine carbohydrate conformation. *Proc. Natl. Acad. Sci. U. S. A.*, **2001**, *98*, 10541-10545.
- Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J.: GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.*, **2008**, *29*, 622-655.
- Kitson, D. H.; Hagler, A. T.: Theoretical-studies of the structure and molecular-dynamics of a peptide crystal. *Biochemistry*, **1988**, *27*, 5246-5257.
- Kräutler, V.; Müller, M.; Hünenberger, P. H.: Conformation, dynamics, solvation and relative stabilities of selected beta-hexopyranoses in water: a molecular dynamics study with the GROMOS 45A4 force field. *Carbohydr. Res.*, **2007**, *342*, 2097-2124.
- Krieg, J.; Gläsner, W.; Vicentini, A.; Doucey, M. A.; Löffler, A.; Hess, D.; Hofsteenge, J.: C-Mannosylation of Human RNase 2 Is an Intracellular Process Performed by a Variety of Cultured Cells. *J. Biol. Chem.*, **1997**, *272*, 26687-26692.

- Krieg, J.; Hartmann, S.; Vicentini, A.; Gläsner, W.; Hess, D.; Hofsteenge, J.: Recognition Signal for C-Mannosylation of Trp-7 in RNase 2 Consists of Sequence Trp-x-x-Trp. *Mol. Biol. Cell*, **1998**, *9*, 301-309.
- 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.
- Lahaye, M.: Developments on gelling algal galactans, their structure and physico-chemistry. *J. Appl. Phycol.*, **2001**, *13*, 173,184.
- Lapthorn, A. J.; Harris, D. C.; Littlejohn, A.; Lustbader, J. W.; Canfield, R. E.; Machin, K. J.; Morgan, F. J.; Isaacs, N. W.: Crystal structure of human chorionic gonadotropin. *Nature*, **1994**, *369*, 455-461.
- Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M.: PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **1993**, *26*, 283-291.
- Latgé, J. P.: The cell wall: a carbohydrate armour for the fungal cell. *Mol. Microbiol.*, **2007**, *66*, 279-290.
- Leach, A. R.: Molecular Modelling Principles and Applications, 2<sup>nd</sup> Ed., **2001**, Longman, Cingapura.
- Leath, K. J.; Johnson, S.; Roversi, P.; Hughes, T. R.; Smith, R. A.; Mackenzie, L.; Morgan, B. P.; Lea, S. M.: High-resolution structures of bacterially expressed soluble human CD59. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, **2007**, *63*, 648-652.
- Li, S.; Liu, B.; Zeng, R.; Cai, Y.; Li, Y.: Predicting O-glycosylation sites in mammalian proteins by using SVMs. *Comput. Biol. Chem.*, **2006**, *30*, 203-208.
- Lins, R. D.; Hünenberger, P. H.: A New GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Comput. Chem.*, **2005**, *26*, 1400-1412.
- 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.

- Longhi, M. P.; Sivasankar, B.; Omidvar, N.; Morgan, B. P.; Gallimore, A.: Cutting edge: murine CD59a modulates antiviral CD4+ T cell activity in a complement-independent manner. *J. Immunol.*, **2005**, *175*, 7098-7102.
- Longhi, M. P.; Harris, C. L.; Morgan, B. P.; Gallimore, A.: Holding T cells in check--a new role for complement regulators? *Trends Immunol.*, **2006**, *27*, 102-108.
- Low, M. G.; Saltiel, A. R.: Structural and functional roles of glycosylphosphatidylinositol in membranes. *Science*, **1988**, *239*, 268-275.
- Low, M. G.: Glycosyl-phosphatidylinositol: a versatile anchor for cell surface proteins. *FASEB J.*, **1989**, *3*, 1600-1608.
- 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., 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.
- MacKerell, 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.
- MacKerell, A. D. Jr.: Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.*, **2004**, *25*, 1584-1604.
- Mandal, T. K.; Mukhopadhyay, C.: Effect of Glycosylation on Structure and Dynamics of MHC Class I Glycoprotein: A Molecular Dynamics Study. *Biopolymers*, **2001**, *59*, 11-23.
- Marshall, R. D.: Glycoproteins. *Annu. Rev. Biochem.*, **1972**, *41*, 673-702.

- Marshall, R. D.: The nature and metabolism of the carbohydrate-peptide linkage of glycoproteins. *Biochem. Soc. Symp.*, **1974**, *40*, 17-26.
- Martín-Pastor, M.; Espinosa, J. F.; Asensio, J. L.; Jiménez-Barbero, J.: A comparison of the geometry and of the energy results obtained by application of different molecular mechanics force fields to methyl  $\alpha$ -lactoside and the C-analogue of lactose. *Carbohydr. Res.*, **1996**, *298*, 15-49.
- Medvedová, L.; Farkas, R.: Hormonal control of protein glycosylation: role of steroids and related lipophilic ligands. *Endocr. Regul.*, **2004**, *39*, 65-79.
- Metzler, W. J.; Bajorath, J.; Fenderson, W.; Shaw, S. Y.; Constantine, K. L.; Naemura, J.; Leytze, G.; Peach, R. J.; Lavoie, T. B.; Mueller, L.; Linsley, P. S.: Solution structure of human CTLA-4 and delineation of a CD80/CD86 binding site conserved in CD28. *Nat. Struct. Biol.*, **1997**, *4*, 527-531.
- McCammon, J. A.; Gelin, B. R.; Karplus, M.: Dynamics of folded proteins. *Nature*, **1977**, *267*, 585-590.
- Melo, F. R.; Pereira, M. S.; Foguel, D.; Mourão, P. A. S.: Antithrombin-mediated Anticoagulant Activity of Sulfated Polysaccharides: Different Mechanism for Heparin and Sulfated Galactans. *J. Biol. Chem.*, **2004**, *279*, 20824-20835.
- Mikhailov, D.; Mayo, K. H.; Vlahov, I. R.; Toida, T.; Pervin, A.; Linhardt, R. J.: NMR solution conformation of heparin-derived tetrasaccharide. *Biochem. J.*, **1996**, *318*, 93-102.
- Mikhailov, D.; Linhardt, R. J.; Mayo, K. H.: NMR solution conformation of heparin-derived hexasaccharide. *Biochem. J.*, **1997**, *328*, 51-61.
- Misevic, G. N.; Burger, M. M.: The species-specific cell-binding site of the aggregation factor from the sponge *Microciona prolifera* is a highly repetitive novel glycan containing glucuronic acid, fucose, and mannose. *J. Biol. Chem.*, **1990a**, *265*, 20577-20584.
- Misevic, G. N.; Burger, M. M.: Involvement of a highly polyvalent glycan in the cell-binding of the aggregation factor from the marine sponge *Microciona prolifera*. *J. Cell Biochem.*, **1990b**, *43*, 307-314.
- Mitra, N.; Sinha, S.; Ramya, T. N. C.; Surolia, A.: N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.*, **2006**, *31*, 156-163.

- Moingeon, P.; Chang, H. C.; Sayre, P. H.; Clayton, L. K.; Alcover, A.; Gardner, P.; Reinherz, E. L.: The structural biology of CD2. *Immunol. Rev.*, **1989**, *111*, 111-144.
- Mourão, P. A. S.; Perlin, A. S.: Structural features of sulfated glycans from the tunic of *Styela plicata* (Chordata-Tunicata): a unique occurrence of L-galactose in sulfated polysaccharides. *Eur. J. Biochem.*, **1987**, *15*, 431-436.
- Mourão, P. A. S.: A carbohydrate-based mechanism of species recognition in sea urchin fertilization. *Braz. J. Med. Biol. Res.*, **2007**, *40*, 5-17.
- Mukhopadhyay, C.: Molecular Dynamics Simulation of Glycoprotein-Glycans of Immunoglobulin G and Immunoglobulin M. *Biopolymers*, **1998**, *45*, 177-190.
- Mulloy, B.; Forster, M. J.; Jones, C.; Davies, D. B.: N.m.r. and molecular-modelling studies of the solution conformation of heparin. *Biochem. J.*, **1993**, *293*, 849-858.
- Mulloy, B.; Ribeiro, A. -C.; Alves, A. -P.; Vieira, R. P.; Mourão, P. A. S.: Sulfated fucans from echinoderms have a regular tetrasaccharide repeating unit defined by specific patterns of sulfation at the O-2 and O-4 positions. *J. Biol. Chem.*, **1994**, *269*, 22113-22123.
- 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.
- Nader, H. B.; Pinhal, M. A. S.; Baú, E. C.; Castro, R. A. B.; Medeiros, G. F.; Chavante, S. F.; Leite, E. L.; Trindade, E. S.; Shinjo, S. K.; Rocha, H. A. O.; Tersariol, I. L. S.; Mendes, A. & Dietrich, C. P.: Development of new heparin-like compounds and other antithrombotic drugs and their interaction with vascular endothelial cells. *Braz. J. Med. Biol. Res.*, **2001**, *34*, 699-709.
- 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.
- Nakagaki, T.; Foster, D. C.; Berkner, K. L.; Kisiel, W.: Initiation of the extrinsic pathway of blood coagulation: evidence for the tissue factor dependent autoactivation of human coagulation factor VII. *Biochemistry*, **1991**, *30*, 10819-10824.

- 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.
- Nguyen, D. H.; Colvin, M. E.; Yeh, Y.; Feeney, R. E.; Fink, W. H.: The Dynamics, Structure, and Conformational Free Energy of Proline-Containing Antifreeze Glycoprotein. *Biophys. J.*, **2002**, *82*, 2892-2905.
- 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.
- Patankar, M. S.; Oehninger, S.; Barnett, T.; Williams, R. L.; Clark, G. F.: A revised structure for fucoidan may explain some of its biological-activities. *J. Biol. Chem.*, **1993**, *268*, 21770-21776.
- Paulick, M. G.; Bertozzi, C. R.: The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. *Biochemistry*, **2008**, *47*, 6991-7000.
- Pereira, M. S.; Mulloy, B.; Mourão, P. A. S.: Structure and anticoagulant activity of sulfated fucans. Comparison between the regular, repetitive, and linear fucans from echinoderms with the more heterogeneous and branched polymers from brown algae. *J. Biol. Chem.*, **1999**, *274*, 7656-7667.
- Pereira, M. S.; Vilela-Silva, A. C. E. S.; Valente, A. P.; Mourão, P. A. S.: A 2-sulfated, 3-linked  $\alpha$ -L-galactan is an anticoagulant polysaccharide. *Carbohydr. Res.*, **2002a**, *337*, 2231-2238.
- Pereira, M. S.; Melo, F. R.; Mourão, P. A. S.: Is there a correlation between structure and anticoagulant action of sulfated galactans and sulfated fucans? *Glycobiology*, **2002b**, *12*, 573-580.
- 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.; Imbert, A.; Engelsens, 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. Struct. Biol.*, **2005**, *15*, 517-524.
- Peterson, A.; Seed, B.: Monoclonal antibody and ligand binding sites of the T cell erythrocyte receptor (CD2). *Nature*, **1987**, *329*, 842-846.
- 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.; Milac, A. L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R.: Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology*, **2004**, *14*, 103-114.
- 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.
- Pickford, A. R.; Smith, S. P.; Staunton, D.; Boyd, J.; Campbell, I. D.: The hairpin structure of the  ${}^6\text{F1}{}^1\text{F2}{}^2\text{F2}$  fragment from human fibronectin enhances gelatin binding. *EMBO J.*, **2001**, *20*, 1519-1529.
- 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.
- Pomin, V. H.; Pereira, M. S.; Valente, A. P.; Tollefsen, D. M.; Pavão, M. S.; Mourão, P. A. S.: Selective cleavage and anticoagulant activity of a sulfated fucan: stereospecific removal of a 2-sulfate ester from the polysaccharide by mild acid hydrolysis, preparation of oligosaccharides, and heparin cofactor II-dependent anticoagulant activity. *Glycobiology*, **2005**, *15*, 369-381.
- Pomin, V. H.; Valente, A. P.; Pereira, M. S.; Mourão, P. A. S.: Mild acid hydrolysis of sulfated fucans: a selective 2-desulfation reaction and an alternative approach for preparing tailored sulfated oligosaccharides. *Glycobiology*, **2005**, *15*, 1376-1385.
- 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.

- Ravenscroft, N.; Gamian, A.; Romanowska, E.: 3-Deoxy-octulosonic-acid-containing hexasaccharide fragment of unusual core type isolated from *Hafnia alvei* 2 lipopolysaccharide. *Eur. J. Biochem.*, **1995**, *227*, 889-896.
- Recny, M. A.; Luther, M. A.; Knoppers, M. H.; Neidhardt, E. A.; Klandekar, S. S.; Concino, M. F.; Schimke, P. A.; Francis, M. A.; Moebius, U.; Reinhold, B. B.; Reinhold, V. N.; Reinherz, E. L.: N-Glycosylation Is Required for Human CD2 Immuno-adhesion Functions. *J. Biol. Chem.*, **1992**, *267*, 22428-22434.
- Renouf, D. V.; Hounsell, E. F.: Conformational studies of the backbone (poly-N-acetyllactosamine) and the core region sequences of O-linked carbohydrate chains. *Int. J. Biol. Macromol.*, **1993**, *15*, 37-42.
- Ress, D. A.; Scott, W. E.: Polysaccharide conformation. Part VI. *J. Chem. Soc. B*, **1971**, 469-479.
- Rigby, A. C.; Lucas-Meunier, E.; Kalume, D. E.; Czerwiec, E.; Hambe, B.; Dahlqvist, I.; Fossier, P.; Baux, G.; Roepstorff, P.; Baleja, J. D.; Furie, B. C.; Furie, B.; Stenflo, J.: A conotoxin from *Conus textile* with unusual posttranslational modifications reduces presynaptic Ca<sup>2+</sup> influx. *Proc. Natl. Acad. Sci. U. S. A.*, **1999**, *96*, 5758-5763.
- 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
- Rubinstein, A.; Kinarsky, L.; Sherman, S.: Molecular Dynamics Simulations of the O-glycosylated 21-residue MUC1 Peptides. *Int. J. Mol. Sci.*, **2004**, *5*, 119-128.
- Sáez, F. J.; Madrid, J. F.; Alonso, E.; Hernández, F.: Glycan composition of follicle (Sertoli) cells of the amphibian *Pleurodeles waltl*. A lectin histochemical study. *J. Anat.*, **2001**, *198*, 673-681.
- Saito, H.; Yamagata, T.; Suzuki, S.: Enzymatic methods for the determination of small quantities of isomeric chondroitin sulfates. *J. Biol. Chem.*, **1968**, *243*, 1536-1542.
- Sampaio Nogueira, C. E.; Ruggiero, J. R.; Sist, P.; Cescutti, P.; Urbani, R.; Rizzo, R.: Conformational features of cepacian: the exopolysaccharide produced by clinical strains of *Burkholderia cepacia*. *Carbohydr. Res.*, **2005**, *340*, 1025-1037.
- Sathyanarayana, B. K.; Rao, V. S. R.: Conformational studies of -glucans. *Biopolymers*, **1972**, *11*, 1379-1394.

- Sathyanarayana, B. K.; Stevens, E. S.: Theoretical study of the conformations of pustulan [(1----6)-beta-D-glucan]. *J. Biomol. Struct. Dyn.*, **1983**, *1*, 947-959.
- Schaftenaar, G. MOLDEN. CAOS/CAMM Center, University of Nijmegen, Toernooiveld 1, 6525 ED NIJMEGEN, The Netherlands, 1997.
- 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.
- Schmidt, M. A.; Riley, L. W.; Benz, I.: Sweet new world: glycoproteins in bacterial pathogens. *Trends Microbiol.*, **2003**, *11*, 554-561.
- 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.
- Sears, P.; Wong, C. -H.: Enzyme action in glycoprotein synthesis. *Cell. Mol. Life Sci.*, **1998**, *54*, 223-252.
- Shefter, E.; Trueblood, K. N.: The crystal and molecular structure of D(+)-barium uridine-5'-phosphate. *Acta Crystallogr.*, **1965**, *18*, 1067-1077.
- Shental-Bechor, D.; Levy, Y.: Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, *105*, 8256-8261.
- Silva, M. E.; Dietrich, C. P.: Structure of heparin. Characterization of the products formed from heparin by the action of a heparinase and a heparitinase from *Flavobacterium heparinum*. *J. Biol. Chem.*, **1975**, *250*, 6841-6846.
- Spiro, R. G.: Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, **2002**, *12*, 43R-56R.
- Sticht, H.; Pickford, A. R.; Potts, J. R.; Campbell, I. D.: Solution Structure of the Glycosylated Second Type 2 Module of Fibronectin. *J. Mol. Biol.*, **1998**, *276*, 177-187.
- Su, Y. H.; Vacquier, V. D.: Cyclic GMP-specific Phosphodiesterase-5 Regulates Motility of Sea Urchin Spermatozoa. *Mol. Biol. Cell*, **2006**, *17*, 114-121.

- 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.
- Takahashi, M.; Yokoe, S.; Asahi, M.; Lee, S. H.; Li, W.; Osumi, D.; Miyoshi, E.; Taniguchi, N.: *N*-glycan of ErbB family plays a crucial role in dimer formation and tumor promotion. *Biochim. Biophys. Acta*, **2008**, *1780*, 520-524.
- Talwar, G. P.: Human chorionic gonadotropin and ovarian and placental steroidogenesis. *J. Steroid Biochem.*, **1979**, *11*, 27-34.
- Thanka Christlet, T. H.; Veluraja, K.: Database Analysis of O-Glycosylation Sites in Proteins. *Biophys. J.*, **2001**, *80*, 952-960.
- Turnbull, J. E.; Field, R. A.: Emerging glycomics technologies. *Nat. Chem. Biol.*, **2007**, *3*, 74-77.
- Usov, A. I.; Bilan, M. I.; Shashkov, A. S.: Structure of a sulfated xylogalactan from the calcareous red alga *Corallina pilulifera* P. et R. (Rhodophyta, Corallinaceae). *Carbohydr. Res.*, **1997**, *303*, 93-102.
- Vacquier, V. D.; Moy, G. W.: Isolation of bindin: the protein responsible for adhesion of sperm to sea urchin eggs. *Proc. Natl. Acad. Sci. U.S.A.*, **1977**, *74*, 2456-2460.
- Vakonakis, I.; Langenhan, T.; Prömel, S.; Russ, A.; Campbell, I. D.: Solution Structure and Sugar-Binding Mechanism of Mouse Latrophilin-1 RBL: a 7TM Receptor-Attached Lectin-Like Domain. *Structure*, **2008**, *16*, 944-953.
- 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.; Berendsen, H. J. C.: Computer simulations of molecular dynamics: methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.*, **1990**, *29*, 992-1023.
- 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, Zúriche, Suíça, 1996.
- van Gunsteren, W. F.; Mark, A. E.: Validation of molecular dynamics simulation. *J. Chem. Phys.*, **1998**, *108*, 6109-6116.

- Varki, A.: Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, **1993**, *3*, 97-130.
- 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**.
- Vieira, R. P.; Mulloy, B.; Mourão, P. A. S.: Structure of a fucose-branched chondroitin sulfate from sea cucumber: evidence for the presence of 3-O-sulfo- $\beta$ -D-glucuronosyl residues. *J. Biol. Chem.*, **1991**, *266*, 13530-13536.
- Vilela-Silva, A. C. E. S.; Alves, A. P.; Valente, A. P.; Vacquier, V. D.; Mourão, P. A. S.: Structure of the sulfated  $\alpha$ -L-fucan from the egg jelly coat of the sea urchin *Strongylocentrotus franciscanus*: patterns of preferential 2-O- and 4-O-sulfation determine sperm cell recognition. *Glycobiology*, **1999**, *9*, 927-933.
- Vilela-Silva, A. C. E. S.; Castro, M. O.; Valente, A. P.; Biermann, C. H.; Mourão, P. A. S.: Sulfated Fucans from the Egg Jellies of the Closely Related Sea Urchins *Strongylocentrotus droebachiensis* and *Strongylocentrotus pallidus* Ensure Species-specific Fertilization. *J. Biol. Chem.*, **2002**, *277*, 379-387.
- Vilela-Silva, A. C. E. S.; Hirohashi, N.; Mourão, P. A. S.: The structure of sulfated polysaccharides ensures a carbohydrate-based mechanism for species recognition during sea urchin fertilization. *Int. J. Dev. Biol.*, **2008**, *52*, 551-559.
- Virudachalam, R.; Rao, V. S. R.: Theoretical Studies on Peptidoglycans. II. Conformations of the Disaccharide-Peptide Subunit and the Three-Dimensional Structure of Peptidoglycan. *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.
- Withka, J. M.; Wyss, D. F.; Wagner, G.; Arulanandam, A. R.; Reinherz, E. L.; Recny, M. A.: Structure of the glycosylated adhesion domain of human T lymphocyte glycoprotein CD2. *Structure*, **1993**, *1*, 69-81.

- Wong, N. K.; Renouf, D. V.; Lehmann, S.; Hounsell, E. F.: Glycosylation of prions and its effects on protein conformation relevant to amino acid mutations. *J. Mol. Graph. Model.*, **2000**, *18*, 126-134.
- Wood, M. J.; Sampoli Benitez, B. A.; Komives, E. A.: Solution structure of the smallest cofactor-active fragment of thrombomodulin. *Nat. Struct. Biol.*, **2000**, *7*, 200-204.
- Woods, R. J.: Three-dimensional structures of oligosaccharides. *Curr. Opin. Struct. Biol.*, **1995**, *5*, 591-598.
- Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B.: Molecular Mechanical and Molecular Dynamical Simulations of Glycoproteins and Oligosaccharides. 1. GLYCAM-93 Parameter Development. *J. Phys. Chem.*, **1995**, *99*, 3832-3846.
- 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.
- Wu, H.; Lustbader, J. W.; Liu, Y.; Canfield, R. E.; Hendrickson, W. A.: Structure of human chorionic gonadotropin at 2.6 Å resolution from MAD analysis of the selenomethionyl protein. *Structure*, **1994**, *2*, 545-558.
- Wyss, D. F.; Withka, J. M.; Knoppers, M. H.; Sterne, K. A.; Recny, M. A.; Wagner, G.: <sup>1</sup>H resonance assignments and secondary structure of the 13.6 kDa glycosylated adhesion domain of human CD2. *Biochemistry*, **1993**, *32*, 10995-11006.
- 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.
- 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.
- Zuegg, J.; Gready, J. E.: Molecular dynamics simulation of human prion protein including both N-linked oligosaccharides and the GPI anchor. *Glycobiology*, **2000**, *10*, 959-974.

## 9 Anexos

Durante a realização desta dissertação, alguns projetos iniciados durante meu período como aluno de Iniciação Científica foram concluídos, da mesma forma que uma colaboração adicional foi realizada com o grupo do Prof. Eliezer J. Barreiro (Laboratório de Avaliação e Síntese de Substâncias Bioativas - LASSBio, Faculdade de Farmácia - UFRJ). Destes resultados, surgiram trabalhos já publicados ou submetidos, incluídos a seguir.

## 9.1 Trabalho I

*Carbohydr. Res.*, 2008, 343, 1435-1445

### **Depiction of the forces participating in the 2-O-sulfo- $\alpha$ -L-iduronic acid conformational preference in heparin sequences in aqueous solutions**

**Laercio Pol-Fachin**, Hugo Verli

#### **RESUMO**

2-*O*-Sulfo- $\alpha$ -L-iduronic acid (IdoA2S) is one of the main components of heparin, an anticoagulant and antithrombotic polysaccharide able to potentiate the inhibitory effect of antithrombin over plasma serine proteases. This monosaccharide unit adopts an equilibrium between chair ( ${}^1C_4$ ) and skew-boat ( ${}^2S_0$ ) forms as a function of heparin sequence size and composition. Although the prevalence of the  ${}^1C_4$  chair conformation in monosaccharides is understood, the reasons for the increase in  ${}^2S_0$  contribution in the whole polysaccharide chain are still uncertain. In this context, 0.2  $\mu$ s molecular dynamics simulations of IdoA2S-containing oligosaccharides indicated that stabilization due to intramolecular hydrogen bonds around IdoA2S is highly correlated ( $p \leq 0.001$ ) with the expected conformational equilibrium for this residue in solution. This behavior explains the known effect of different heparin compositions, at the monosaccharide level, on IdoA2S conformation in biological solutions.

## 9.2 Trabalho II

*J. Mol. Graph. Model.*, 2009, 28, 446-454

### **Characterization of the conformational ensemble from bioactive N-acylhydrazone derivatives**

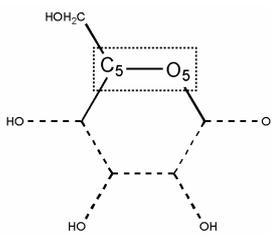
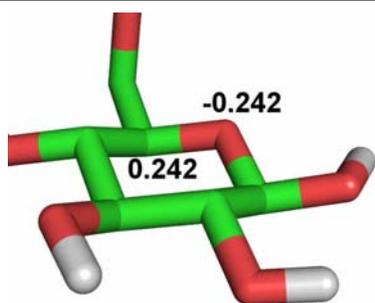
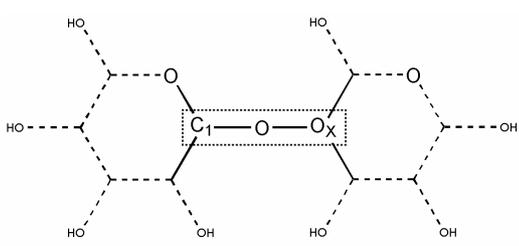
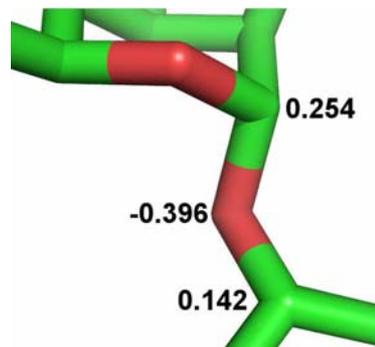
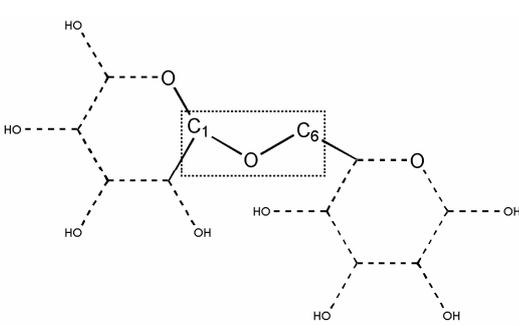
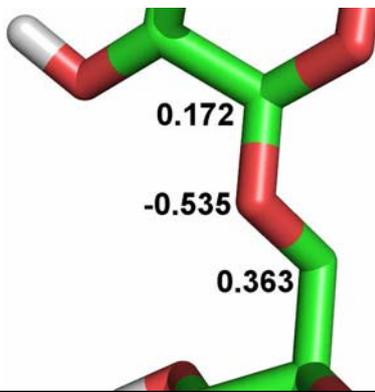
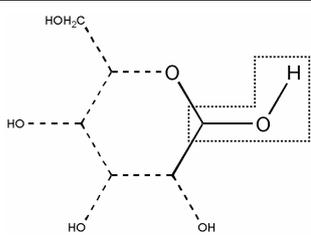
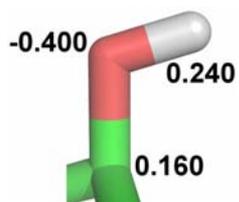
**Laercio Pol-Fachin**, Carlos Alberto Manssour Fraga, Eliezer J. Barreiro, Hugo Verli

#### **RESUMO**

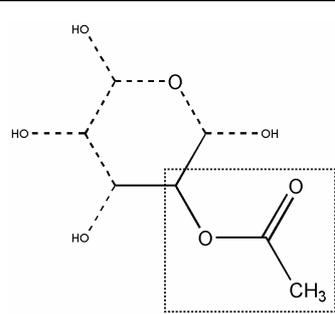
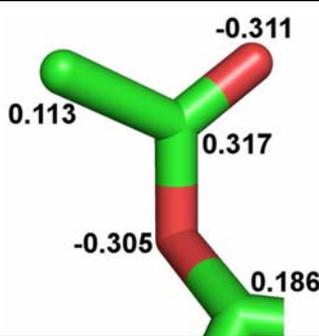
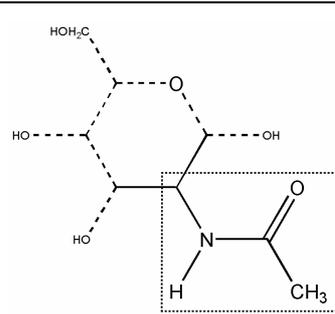
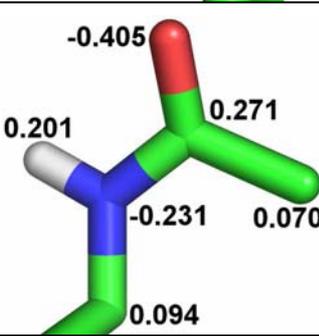
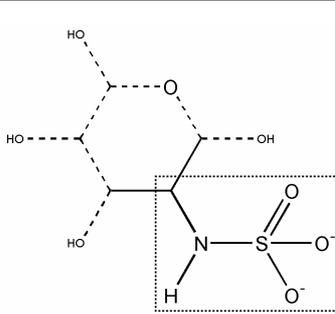
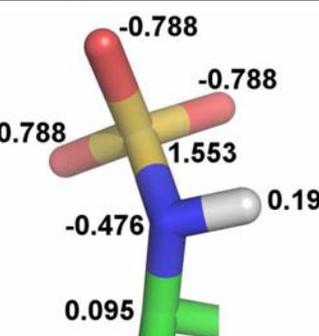
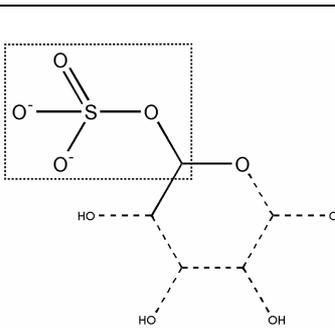
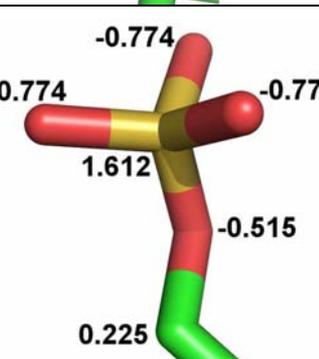
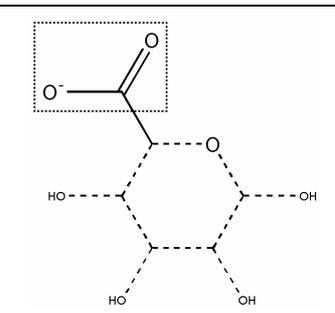
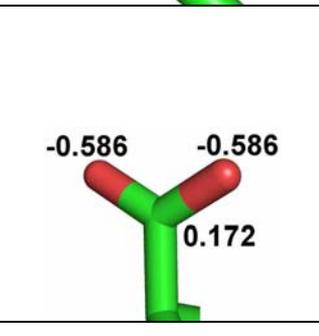
The search for bioactive conformations from prototypes is mostly referenced on crystallographic ligand-receptor complexes, in which the molecule conformation is already caged inside its binding site. However, the complexation process is a thermodynamic event depending on both complexed and uncomplexed states. As ligand affinity originates from such equilibrium, the development of novel computational models capable of supplying data on ligand dynamics in biological solutions is potentially applicable in more efficient methods for prediction of compounds binding and affinity. In this context, the current work employs a series of molecular dynamics simulations on three *N*-acylhydrazone derivatives, already shown to present promising cardioinotropic and vasodilatory activities, in order to obtain a precise characterization of each compound conformational ensemble in aqueous solutions, instead of a single minimum energy conformation. Consequently, we were able to observe the influence of each functional group of the studied molecules on the conformation of the entire compounds and thus on the exposure of functional groups that might potentially bind to target receptors. Additionally, the differences between the molecules conformational behavior were characterized, supporting a spatial and temporal image of each ligand, which may be potentially correlated to their biological activities. So in the context of conformational selection, such strategy may represent a useful methodology to contribute in the choice of ligands conformations for both QSAR and docking calculations.

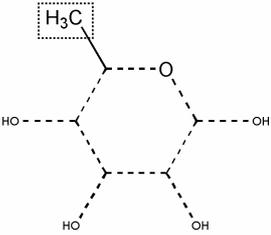
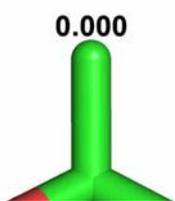
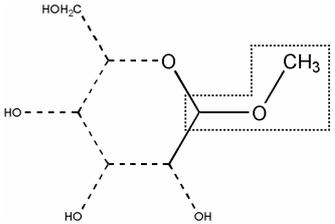
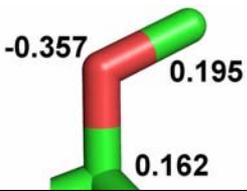
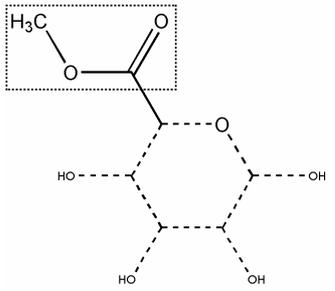
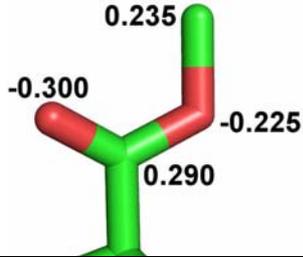
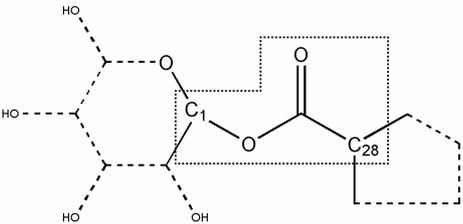
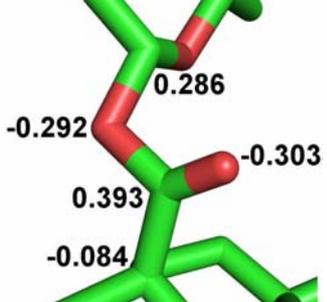
### 9.3 Cargas atômicas de Löwdin

As cargas atômicas parciais de Löwdin são divididas em grupamentos, ou grupos de carga, para que possam ser aplicadas para diferentes resíduos de carboidratos e, potencialmente, em outros campos de força.

Grupamento	Estrutura 2D	Estrutura 3D e cargas
C5-O5 do anel piranosídico		
Ligações glicosídicas 1→X*		
Ligação glicosídica 1→6		
Hidroxila		

\* X= 2, 3 ou 4

Grupamento	Estrutura 2D	Estrutura 3D e cargas
Acetila		
Acetoamido		
Sulfonamida		
Sulfato		
Carboxilato		

Grupamento	Estrutura 2D	Estrutura 3D e cargas
Metila		 <p>0.000</p>
Metoxila		 <p>-0.357 0.195 0.162</p>
Metil-Éster		 <p>0.235 -0.300 -0.225 0.290</p>
Ligação glicosídica entre triterpeno e monossacarídeo		 <p>-0.292 0.286 -0.303 0.393 -0.084</p>

#### 9.4 Topologias de monossacarídeos para GROMOS96 43a1

Esses arquivos contêm os parâmetros de ligação, ângulo de ligação e diedros para as moléculas em estudo. No caso específico do campo de força GROMOS96 43a1, utilizando o pacote GROMACS, essas topologias, juntamente com outras que descrevem outras classes de biomoléculas, fazem parte de um arquivo com extensão “rtp”, localizado em uma pasta de topologias compartilhadas do programa.

Abaixo, seguem as topologias para alguns dos monossacarídeos constituintes das glicoproteínas estudadas, assim como parâmetros para os aminoácidos a que esses oligossacarídeos se ligam.

#### *$\beta$ -GlcNAc-(1→N)-Asn*

```
[ ASN2 ]
[ atoms ]
  N      N      -0.28000    0
  H      H      0.28000    0
  CA     CH1     0.00000    1
  CB     CH2     0.00000    1
  CG     C       0.38000    2
  OD1    O      -0.38000    2
  ND2    NT     -0.26500    3
  HD2    H       0.15500    3
  C      C       0.380     4
  O      O      -0.380     4
[ bonds ]
  N      H      gb_2
  N      CA     gb_20
  CA     C      gb_26
  C      O      gb_4
  C      +N     gb_9
  CA     CB     gb_26
  CB     CG     gb_26
  CG     OD1    gb_4
  CG     ND2    gb_8
  ND2    HD2    gb_2
[ angles ]
; ai    aj    ak    gromos type
  -C    N     H     ga_31
  H     N     CA    ga_17
  -C    N     CA    ga_30
  N     CA    C     ga_12
  CA    C     +N    ga_18
  CA    C     O     ga_29
  O     C     +N    ga_32
  CA    C     O     ga_29
  O     C     +N    ga_32
  N     CA    CB    ga_12
  C     CA    CB    ga_12
  CA    CB    CG    ga_14
  CB    CG    OD1   ga_29
  CB    CG    ND2   ga_18
  OD1   CG    ND2   ga_32
  CG    ND2   HD2   ga_22
[ impropers ]
; ai    aj    ak    al    gromos type
  N     -C    CA    H     gi_1
  C     CA    +N    O     gi_1
  CA    N     C     CB    gi_2
  CG    OD1   ND2   CB    gi_1
[ dihedrals ]
; ai    aj    ak    al    gromos type
  -CA   -C    N     CA    gd_4
  -C    N     CA    C     gd_19
  N     CA    C     +N    gd_20
  N     CA    CB    CG    gd_17
  CA    CB    CG    ND2   gd_20
  CB    CG    ND2   HD2   gd_4
```

$\beta$ -GlcNAc-(1→4)- $\beta$ -GlcNAc-(1→N)-Asn

```

[ NAG1 ]
[ atoms ]
  C8  CH3    0.07000    0
  C7   C     0.27100    0
  O7   O    -0.40500    0
  N2   N    -0.23100    0
  H22  H     0.20100    0
  C2   CH1   0.09400    0
  C1   CH1   0.11000    1
  O5   OA   -0.24200    2
  C5   CH1   0.24200    2
  C6   CH2   0.36300    3
  O6   OA   -0.53500    3
  C4   CH1   0.14200    4
  O4   OA   -0.39600    4
  C3   CH1   0.16000    5
  O3   OA   -0.40000    5
  H32  H     0.24000    5
[ bonds ]
  C8  C7    gb_26
  C7  O7    gb_4
  C7  N2    gb_10
  N2  H22   gb_2
  N2  C2    gb_20
  C2  C1    gb_25
  C2  C3    gb_25
  C1  O5    gb_19
  O5  C5    gb_19
  C5  C6    gb_25
  C5  C4    gb_25
  C6  O6    gb_19
  C4  O4    gb_19
  C4  C3    gb_25
  O4  +C1   gb_19
  C3  O3    gb_19
  O3  H32   gb_1
[ angles ]
; ai  aj  ak  gromos type
  C8  C7  O7    ga_29
  C8  C7  N2    ga_18
  O7  C7  N2    ga_32
  C7  N2  H22   ga_31
  C7  N2  C2    ga_30
  H22 N2  C2    ga_24
  N2  C2  C1    ga_12
  N2  C2  C3    ga_12
  C1  C2  C3    ga_7
  C2  C1  O5    ga_8
  C1  O5  C5    ga_9
  O5  C5  C6    ga_8
  O5  C5  C4    ga_8
  C6  C5  C4    ga_7
  C5  C6  O6    ga_8
  C5  C4  O4    ga_8
  C5  C4  C3    ga_7
  O4  C4  C3    ga_8
  C4  O4  +C1   ga_9
  C2  C3  C4    ga_7
  C2  C3  O3    ga_8
  C4  C3  O3    ga_8
  C3  O3  H32   ga_11
[ impropers ]
; ai  aj  ak  al  gromos type
  C7  C8  O7    N2    gi_1
  N2  C7  H22   C2    gi_1
  C2  N2  C3    C1    gi_2
  C5  O5  C6    C4    gi_2
  C4  C5  C3    O4    gi_2
  C3  O3  C2    C4    gi_2
  C5  C2  C4    C1    gi_7
  C5  C2  C3    C1    gi_8
  C5  C2  C3    O5    gi_9
[ dihedrals ]
; ai  aj  ak  al  gromos type
  C8  C7  N2    C2    gd_4
  C3  C2  N2    C7    gd_14
  O5  C1  C2    N2    gd_8
  O3  C3  C2    N2    gd_8
  C2  C1  O5    C5    gd_14
  C4  C5  O5    C1    gd_14
  O6  C6  C5    O5    gd_8
  C3  C4  C5    O5    gd_7
  O3  C3  C4    C5    gd_7
  C2  C3  O3    H32   gd_12
  C3  C4  O4    +C1   gd_14
  C2  C3  C4    C5    gd_17
  C6  C5  C4    C3    gd_17
  C6  C5  C4    O4    gd_7
  N2  C2  C3    C4    gd_7
  C1  C2  C3    O3    gd_7
  O3  C3  C4    O4    gd_8
  O5  C5  C4    O4    gd_8
  C4  C5  C6    O6    gd_7
  C4  C5  C6    O6    gd_17
  O5  C1  C2    C3    gd_7
  C1  C2  C3    C4    gd_17
  C2  C3  C4    O4    gd_7

```

$\beta$ -Man-(1→4)- $\beta$ -GlcNAc-(1→4)- $\beta$ -GlcNAc

```

[ NAG2 ]
[ atoms ]
  C8  CH3    0.07000    0
  C7   C     0.27100    0
  O7   O    -0.40500    0
  N2   N    -0.23100    0
  H22  H     0.20100    0
  C2   CH1   0.09400    0
  C1   CH1   0.25400    1
  O5   OA   -0.24200    2
  C5   CH1   0.24200    2
  C6   CH2   0.16000    3
  O6   OA   -0.40000    3
  H63  H     0.24000    3
  C4   CH1   0.14200    4
  O4   OA   -0.39600    4
  C3   CH1   0.16000    5
  O3   OA   -0.40000    5
  H32  H     0.24000    5
[ bonds ]
  C8  C7      gb_26
  C7  O7      gb_4
  C7  N2      gb_10
  N2  H22     gb_2
  N2  C2      gb_20
  C2  C1      gb_25
  C2  C3      gb_25
  C1  O5      gb_19
  O5  C5      gb_19
  C5  C6      gb_25
  C5  C4      gb_25
  C6  O6      gb_19
  O6  H63     gb_1
  C4  O4      gb_19
  O4  +C1     gb_19
  C4  C3      gb_25
  C3  O3      gb_19
  O3  H32     gb_1
[ angles ]
; ai  aj  ak  gromos type
  C8  C7  O7      ga_29
  C8  C7  N2      ga_18
  O7  C7  N2      ga_32
  C7  N2  H22     ga_31
  C7  N2  C2      ga_30
  H22 N2  C2      ga_24
  N2  C2  C1      ga_12
  N2  C2  C3      ga_12
  C1  C2  C3      ga_7
  C2  C1  O5      ga_8
  C1  O5  C5      ga_9
 -O4  C1  O5      ga_8
 -O4  C1  C2      ga_8
  O5  C5  C6      ga_8
  O5  C5  C4      ga_8
  C6  C5  C4      ga_7
  C5  C6  O6      ga_8
  C6  O6  H63     ga_11
  C5  C4  O4      ga_8
  C5  C4  C3      ga_7
  C4  O4  +C1     ga_9
  O4  C4  C3      ga_8
  C2  C3  C4      ga_7
  C2  C3  O3      ga_8
  C4  C3  O3      ga_8
  C3  O3  H32     ga_11
[ impropers ]
; ai  aj  ak  al  gromos type
  C1  -O4  O5  C2  gi_2
  C7  C8  O7  N2  gi_1
  N2  C7  H22  C2  gi_1
  C2  N2  C3  C1  gi_2
  C5  O5  C6  C4  gi_2
  C4  C5  C3  O4  gi_2
  C3  O3  C2  C4  gi_2
  C5  C2  C4  C1  gi_7
  C5  C2  C3  C1  gi_8
  C5  C2  C3  O5  gi_9
[ dihedrals ]
; ai  aj  ak  al  gromos type
  C8  C7  N2  C2  gd_4
  C3  C2  N2  C7  gd_14
  O5  C1  C2  N2  gd_8
  O3  C3  C2  N2  gd_8
  C2  C1  O5  C5  gd_14
  C4  C5  O5  C1  gd_14
 -C4  -O4  C1  C2  gd_14
 -O4  C1  C2  C3  gd_17
 -O4  C1  C2  C3  gd_7
 -O4  C1  C2  O2  gd_8
  O6  C6  C5  O5  gd_8
  C3  C4  C5  O5  gd_7
  O3  C3  C4  C5  gd_7
  C2  C3  O3  H32  gd_12
  C3  C4  O4  +C1  gd_14
  H63 O6  C6  C5  gd_12
  C2  C3  C4  C5  gd_17
  C6  C5  C4  C3  gd_17
  C6  C5  C4  O4  gd_7
  N2  C2  C3  C4  gd_7
  C1  C2  C3  O3  gd_7
  O3  C3  C4  O4  gd_8
  O5  C5  C4  O4  gd_8
  C4  C5  C6  O6  gd_7
  C4  C5  C6  O6  gd_17
  O5  C1  C2  C3  gd_7
  C1  C2  C3  C4  gd_17
  C2  C3  C4  O4  gd_7

```



```
[ BMA1 ]
[ atoms ]
  C6  CH2    0.36300    0
  O6  OA    -0.53500    0
  C5  CH1    0.24200    1
  O5  OA    -0.24200    1
  C1  CH1    0.25400    2
  C2  CH1    0.16000    3
  O2  OA    -0.40000    3
  H22 H      0.24000    3
  C3  CH1    0.14200    4
  O3  OA    -0.39600    4
  C4  CH1    0.16000    5
  O4  OA    -0.40000    5
  H42 H      0.24000    5
[ bonds ]
  C6  C5      gb_25
  C5  O5      gb_19
  C5  C4      gb_25
  O5  C1      gb_19
  C1  C2      gb_25
  C2  O2      gb_19
  C2  C3      gb_25
  O2  H22     gb_1
  C3  O3      gb_19
  C3  C4      gb_25
  C4  O4      gb_19
  O4  H42     gb_1
  C6  O6      gb_19
  O6  +C1     gb_19
[ angles ]
; ai  aj  ak  gromos type
  C6  C5  O5      ga_8
  C6  C5  C4      ga_7
  O5  C5  C4      ga_8
  C5  O5  C1      ga_9
  O5  C1  C2      ga_8
  C1  C2  O2      ga_8
-04  C1  O5      ga_8
-04  C1  C2      ga_8
  C1  C2  C3      ga_7
  O2  C2  C3      ga_8
  C2  O2  H22     ga_11
  C2  C3  O3      ga_8
  C2  C3  C4      ga_7
  O3  C3  C4      ga_8
  C5  C4  C3      ga_7
  C5  C4  O4      ga_8
  C3  C4  O4      ga_8
  C4  O4  H42     ga_11
  C5  C6  O6      ga_8
  C6  O6  +C1     ga_9
[ impropers ]
; ai  aj  ak  al  gromos type
  C1  -O4  O5  C2  gi_2
  C5  C6  C4  O5  gi_2
  C2  C3  O2  C1  gi_2
  C3  C4  O3  C2  gi_2
  C4  O4  C5  C3  gi_2
  C5  C2  C4  C1  gi_7
  C5  C2  C3  C1  gi_8
  C5  C2  C3  O5  gi_9
[ dihedrals ]
; ai  aj  ak  al  gromos type
  C2  C1  O5  C5  gd_14
  C3  C2  C1  O5  gd_7
  C1  C2  O2  H22  gd_12
  C4  C3  C2  C1  gd_17
  O4  C4  C3  C2  gd_7
  C5  C4  O4  H42  gd_12
  C2  C3  C4  C5  gd_17
  C1  O5  C5  C4  gd_14
-04  -O4  C1  C2  gd_14
-04  C1  C2  C3  gd_17
-04  C1  C2  C3  gd_7
-04  C1  C2  O2  gd_8
  O5  C5  C4  C3  gd_7
  C6  C5  C4  C3  gd_17
  C6  C5  C4  O4  gd_7
  C5  C6  O6  +C1  gd_14
  O5  C1  C2  O2  gd_8
  O2  C2  C3  C4  gd_7
  O2  C2  C3  O3  gd_8
  C1  C2  C3  O3  gd_7
  O3  C3  C4  C5  gd_7
  O3  C3  C4  O4  gd_8
  O5  C5  C4  O4  gd_8
  C4  C5  C6  O6  gd_7
  C4  C5  C6  O6  gd_17
  O5  C5  C6  O6  gd_8
```

$\alpha$ -Man-(1→6)- $\alpha$ -Man-(1→6)- $\beta$ -Man

```

[ MAN1 ]
[ atoms ]
  C6  CH2    0.36300    0
  O6  OA     -0.53500    0
  C5  CH1    0.24200    1
  O5  OA     -0.24200    1
  C1  CH1    0.17200    2
  C2  CH1    0.16000    3
  O2  OA     -0.40000    3
  H22 H      0.24000    3
  C3  CH1    0.14200    4
  O3  OA     -0.39600    4
  C4  CH1    0.16000    5
  O4  OA     -0.40000    5
  H42 H      0.24000    5
[ bonds ]
  C6  C5      gb_25
  C5  O5      gb_19
  C5  C4      gb_25
  O5  C1      gb_19
  C1  C2      gb_25
  C2  O2      gb_19
  C2  C3      gb_25
  O2  H22     gb_1
  C3  O3      gb_19
  C3  C4      gb_25
  C4  O4      gb_19
  O4  H42     gb_1
  C6  O6      gb_19
  O6  +C1     gb_19
[ angles ]
; ai  aj  ak  gromos type
  C6  C5  O5      ga_8
  C6  C5  C4      ga_7
  O5  C5  C4      ga_8
  C5  O5  C1      ga_9
  O5  C1  C2      ga_8
  C1  C2  O2      ga_8
 -O6  C1  O5      ga_8
 -O6  C1  C2      ga_8
  C1  C2  C3      ga_7
  O2  C2  C3      ga_8
  C2  O2  H22     ga_11
  C2  C3  O3      ga_8
  C2  C3  C4      ga_7
  O3  C3  C4      ga_8
  C5  C4  C3      ga_7
  C5  C4  O4      ga_8
  C3  C4  O4      ga_8
  C4  O4  H42     ga_11
  C5  C6  O6      ga_8
  C6  O6  +C1     ga_9
[ impropers ]
; ai  aj  ak  al  gromos type
  C1  -O6  C2  O5      gi_2
  C5  C6  C4  O5      gi_2
  C2  C3  O2  C1      gi_2
  C3  C4  O3  C2      gi_2
  C4  O4  C5  C3      gi_2
  C5  C2  C4  C1      gi_7
  C5  C2  C3  C1      gi_8
  C5  C2  C3  O5      gi_9
[ dihedrals ]
; ai  aj  ak  al  gromos type
  C2  C1  O5  C5      gd_14
  C3  C2  C1  O5      gd_7
  C1  C2  O2  H22     gd_12
  C4  C3  C2  C1      gd_17
  O4  C4  C3  C2      gd_7
  C5  C4  O4  H42     gd_12
  C2  C3  C4  C5      gd_17
  C1  O5  C5  C4      gd_14
 -C6  -O6  C1  C2      gd_14
 -O6  C1  C2  C3      gd_17
 -O6  C1  C2  C3      gd_7
 -O6  C1  C2  O2      gd_8
  O5  C5  C4  C3      gd_7
  C6  C5  C4  C3      gd_17
  C6  C5  C4  O4      gd_7
  C5  C6  O6  +C1     gd_14
  O5  C1  C2  O2      gd_8
  O2  C2  C3  C4      gd_7
  O2  C2  C3  O3      gd_8
  C1  C2  C3  O3      gd_7
  O3  C3  C4  C5      gd_7
  O3  C3  C4  O4      gd_8
  O5  C5  C4  O4      gd_8
  C4  C5  C6  O6      gd_7
  C4  C5  C6  O6      gd_17
  O5  C5  C6  O6      gd_8

```

*$\alpha$ -Man-(1 $\rightarrow$ 2)- $\alpha$ -Man-(1 $\rightarrow$ 6)- $\alpha$ -Man*

```

[ MAN2 ]
[ atoms ]
  C6  CH2    0.16000    0
  O6  OA    -0.40000    0
  H63 H      0.24000    0
  C5  CH1    0.24200    1
  O5  OA    -0.24200    1
  C1  CH1    0.17200    2
  C2  CH1    0.14200    3
  O2  OA    -0.39600    3
  C3  CH1    0.16000    4
  O3  OA    -0.40000    4
  H32 H      0.24000    4
  C4  CH1    0.16000    5
  O4  OA    -0.40000    5
  H42 H      0.24000    5
[ bonds ]
  C6  C5      gb_25
  C5  O5      gb_19
  C5  C4      gb_25
  O5  C1      gb_19
  C1  C2      gb_25
  C2  O2      gb_19
  O2  +C1     gb_19
  C2  C3      gb_25
  C3  O3      gb_19
  C3  C4      gb_25
  O3  H32     gb_1
  C4  O4      gb_19
  O4  H42     gb_1
  C6  O6      gb_19
  O6  H63     gb_1
[ angles ]
; ai  aj  ak  gromos type
  C6  C5  O5      ga_8
  C6  C5  C4      ga_7
  O5  C5  C4      ga_8
  C5  O5  C1      ga_9
  O5  C1  C2      ga_8
  C1  C2  O2      ga_8
  C1  C2  C3      ga_7
-06  C1  O5      ga_8
-06  C1  C2      ga_8
  O2  C2  C3      ga_8
  C2  O2  +C1     ga_9
  C2  C3  O3      ga_8
  C2  C3  C4      ga_7
  O3  C3  C4      ga_8
  C3  O3  H32     ga_11
  C5  C4  C3      ga_7
  C5  C4  O4      ga_8
  C3  C4  O4      ga_8
  C4  O4  H42     ga_11
  C5  C6  O6      ga_8
  C6  O6  H63     ga_11
[ impropers ]
; ai  aj  ak  al  gromos type
  C1  -O6  C2  O5      gi_2
  C5  C6  C4  O5      gi_2
  C2  C3  O2  C1      gi_2
  C3  C4  O3  C2      gi_2
  C4  O4  C5  C3      gi_2
  C5  C2  C4  C1      gi_7
  C5  C2  C3  C1      gi_8
  C5  C2  C3  O5      gi_9
[ dihedrals ]
; ai  aj  ak  al  gromos type
  C2  C1  O5  C5      gd_14
  C3  C2  C1  O5      gd_7
  C4  C3  C2  C1      gd_17
  C2  C3  O3  H32     gd_12
  O4  C4  C3  C2      gd_7
  C5  C4  O4  H42     gd_12
  C2  C3  C4  C5      gd_17
  C1  O5  C5  C4      gd_14
  O5  C5  C4  C3      gd_7
-06  -O6  C1  C2      gd_14
-06  C1  C2  C3      gd_17
-06  C1  C2  C3      gd_7
-06  C1  C2  O2      gd_8
  C6  C5  C4  C3      gd_17
  C6  C5  C4  O4      gd_7
  O5  C1  C2  O2      gd_8
  O2  C2  C3  C4      gd_7
  O2  C2  C3  O3      gd_8
  C3  C2  O2  +C1     gd_14
  C1  C2  C3  O3      gd_7
  O3  C3  C4  C5      gd_7
  O3  C3  C4  O4      gd_8
  O5  C5  C4  O4      gd_8
  C4  C5  C6  O6      gd_7
  C4  C5  C6  O6      gd_17
  O5  C5  C6  O6      gd_8
  C5  C6  O6  H63     gd_12

```

$\alpha$ -Man-(1 $\rightarrow$ 2)- $\alpha$ -Man

```

[ MAN3 ]
[ atoms ]
  C6  CH2    0.16000    0
  O6  OA    -0.40000    0
  H63 H      0.24000    0
  C5  CH1    0.24200    1
  O5  OA    -0.24200    1
  C1  CH1    0.25400    2
  C2  CH1    0.16000    3
  O2  OA    -0.40000    3
  H22 H      0.24000    3
  C3  CH1    0.16000    4
  O3  OA    -0.40000    4
  H32 H      0.24000    4
  C4  CH1    0.16000    5
  O4  OA    -0.40000    5
  H42 H      0.24000    5
[ bonds ]
  C6  C5      gb_25
  C5  O5      gb_19
  C5  C4      gb_25
  O5  C1      gb_19
  C1  C2      gb_25
  C2  O2      gb_19
  C2  C3      gb_25
  O2  H22     gb_1
  C3  O3      gb_19
  C3  C4      gb_25
  O3  H32     gb_1
  C4  O4      gb_19
  O4  H42     gb_1
  C6  O6      gb_19
  O6  H63     gb_1
[ angles ]
; ai  aj  ak  gromos type
  C6  C5  O5      ga_8
  C6  C5  C4      ga_7
  O5  C5  C4      ga_8
  C5  O5  C1      ga_9
  O5  C1  C2      ga_8
  C1  C2  O2      ga_8
  C1  C2  C3      ga_7
  O2  C2  C3      ga_8
 -O2  C1  O5      ga_8
 -O2  C1  C2      ga_8
  C2  O2  H22     ga_11
  C2  C3  O3      ga_8
  C2  C3  C4      ga_7
  O3  C3  C4      ga_8
  C3  O3  H32     ga_11
  C4  O4  C5      ga_8
  C5  O6  H63     ga_11
[ impropers ]
; ai  aj  ak  al  gromos type
  C1  -O2  C2  O5      gi_2
  C5  C6  C4  O5      gi_2
  C2  C3  O2  C1      gi_2
  C3  C4  O3  C2      gi_2
  C4  O4  C5  C3      gi_2
  C5  C2  C4  C1      gi_7
  C5  C2  C3  C1      gi_8
  C5  C2  C3  O5      gi_9
[ dihedrals ]
; ai  aj  ak  al  gromos type
  C2  C1  O5  C5      gd_14
  C3  C2  C1  O5      gd_7
  C1  C2  O2  H22     gd_12
  C4  C3  C2  C1      gd_17
  C2  C3  O3  H32     gd_12
  O4  C4  C3  C2      gd_7
  C5  C4  O4  H42     gd_12
  C2  C3  C4  C5      gd_17
  C1  O5  C5  C4      gd_14
  O5  C5  C4  C3      gd_7
  C6  C5  C4  C3      gd_17
  C6  C5  C4  O4      gd_7
  O5  C1  C2  O2      gd_8
  O2  C2  C3  C4      gd_7
  O2  C2  C3  O3      gd_8
 -C2  -O2  C1  C2      gd_14
 -O2  C1  C2  C3      gd_17
 -O2  C1  C2  C3      gd_7
 -O2  C1  C2  O2      gd_8
  C1  C2  C3  O3      gd_7
  O3  C3  C4  C5      gd_7
  O3  C3  C4  O4      gd_8
  O5  C5  C4  O4      gd_8
  C4  C5  C6  O6      gd_7
  C4  C5  C6  O6      gd_17
  O5  C5  C6  O6      gd_8
  C5  C6  O6  H63     gd_12

```

$\beta$ -GlcNAc-(1→2)- $\alpha$ -Man-(1→6)- $\beta$ -Man

```

[ MAN6 ]
[ atoms ]
  C6  CH2    0.16000    0
  O6  OA    -0.40000    0
  H63 H      0.24000    0
  C5  CH1    0.24200    1
  O5  OA    -0.24200    1
  C1  CH1    0.17200    2
  C2  CH1    0.14200    3
  O2  OA    -0.39600    3
  C3  CH1    0.16000    4
  O3  OA    -0.40000    4
  H32 H      0.24000    4
  C4  CH1    0.16000    5
  O4  OA    -0.40000    5
  H42 H      0.24000    5
[ bonds ]
  C6  C5      gb_25
  C5  O5      gb_19
  C5  C4      gb_25
  O5  C1      gb_19
  C1  C2      gb_25
  C2  O2      gb_19
  C2  C3      gb_25
  C3  O3      gb_19
  O3  H32     gb_1
  C3  C4      gb_25
  C4  O4      gb_19
  O4  H42     gb_1
  C6  O6      gb_19
  O6  H63     gb_1
  O2  +C1     gb_19
[ angles ]
; ai  aj  ak  gromos type
  C6  C5  O5      ga_8
  C6  C5  C4      ga_7
  O5  C5  C4      ga_8
  C5  O5  C1      ga_9
  O5  C1  C2      ga_8
  C1  C2  O2      ga_8
-06  C1  O5      ga_8
-06  C1  C2      ga_8
  C1  C2  C3      ga_7
  O2  C2  C3      ga_8
  C2  C3  O3      ga_8
  C3  O3  H32     ga_11
  C2  C3  C4      ga_7
  O3  C3  C4      ga_8
  C5  C4  C3      ga_7
  C5  C4  O4      ga_8
  C3  C4  O4      ga_8
  C4  O4  H42     ga_11
  C5  C6  O6      ga_8
  C6  O6  H63     ga_11
  C2  O2  +C1     ga_9
[ impropers ]
; ai  aj  ak  al  gromos type
  C1  -O6  C2  O5      gi_2
  C5  C6  C4  O5      gi_2
  C2  C3  O2  C1      gi_2
  C3  C4  O3  C2      gi_2
  C4  O4  C5  C3      gi_2
  C5  C2  C4  C1      gi_7
  C5  C2  C3  C1      gi_8
  C5  C2  C3  O5      gi_9
[ dihedrals ]
; ai  aj  ak  al  gromos type
  C2  C1  O5  C5      gd_14
  C3  C2  C1  O5      gd_7
  C4  C3  C2  C1      gd_17
  O4  C4  C3  C2      gd_7
  C5  C4  O4  H42     gd_12
  C2  C3  C4  C5      gd_17
  C1  O5  C5  C4      gd_14
-C6  -O6  C1  C2      gd_14
-06  C1  C2  C3      gd_17
-06  C1  C2  C3      gd_7
-06  C1  C2  O2      gd_8
  O5  C5  C4  C3      gd_7
  C6  C5  C4  C3      gd_17
  C6  C5  C4  O4      gd_7
  C3  C2  O2  +C1     gd_14
  O5  C1  C2  O2      gd_8
  O2  C2  C3  C4      gd_7
  O2  C2  C3  O3      gd_8
  C1  C2  C3  O3      gd_7
  C2  C3  O3  H32     gd_12
  O3  C3  C4  C5      gd_7
  O3  C3  C4  O4      gd_8
  O5  C5  C4  O4      gd_8
  C4  C5  C6  O6      gd_7
  C4  C5  C6  O6      gd_17
  O5  C5  C6  O6      gd_8
  C5  C6  O6  H63     gd_12

```

$\beta$ -Gal-(1 $\rightarrow$ 4)- $\beta$ -GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -Man

```

[ NAG9 ]
[ atoms ]
  C8  CH3    0.07000    0
  C7   C     0.27100    0
  O7   O    -0.40500    0
  N2   N    -0.23100    0
  H22  H     0.20100    0
  C2   CH1   0.09400    0
  C1   CH1   0.25400    1
  O5   OA   -0.24200    2
  C5   CH1   0.24200    2
  C6   CH2   0.16000    3
  O6   OA   -0.40000    3
  H63  H     0.24000    3
  C4   CH1   0.14200    4
  O4   OA   -0.39600    4
  C3   CH1   0.16000    5
  O3   OA   -0.40000    5
  H32  H     0.24000    5

[ bonds ]
  C8  C7      gb_26
  C7  O7      gb_4
  C7  N2      gb_10
  N2  H22     gb_2
  N2  C2      gb_20
  C2  C1      gb_25
  C2  C3      gb_25
  C1  O5      gb_19
  O5  C5      gb_19
  C5  C6      gb_25
  C5  C4      gb_25
  C6  O6      gb_19
  O6  H63     gb_1
  C4  O4      gb_19
  O4  +C1     gb_19
  C4  C3      gb_25
  C3  O3      gb_19
  O3  H32     gb_1

[ angles ]
; ai  aj  ak  gromos type
  C8  C7  O7      ga_29
  C8  C7  N2      ga_18
  O7  C7  N2      ga_32
  C7  N2  H22     ga_31
  C7  N2  C2      ga_30
  H22 N2  C2      ga_24
  N2  C2  C1      ga_12
  N2  C2  C3      ga_12
  C1  C2  C3      ga_7
  C2  C1  O5      ga_8
  C1  O5  C5      ga_9
 -O2  C1  O5      ga_8
 -O2  C1  C2      ga_8
  O5  C5  C6      ga_8
  O5  C5  C4      ga_8

  C6  C5  C4      ga_7
  C5  C6  O6      ga_8
  C6  O6  H63     ga_11
  C5  C4  O4      ga_8
  C5  C4  C3      ga_7
  C4  O4  +C1     ga_9
  O4  C4  C3      ga_8
  C2  C3  C4      ga_7
  C2  C3  O3      ga_8
  C4  C3  O3      ga_8
  C3  O3  H32     ga_11

[ impropers ]
; ai  aj  ak  al  gromos type
  C1  -O2  O5  C2  gi_2
  C7  C8  O7  N2  gi_1
  N2  C7  H22  C2  gi_1
  C2  N2  C3  C1  gi_2
  C5  O5  C6  C4  gi_2
  C4  C5  C3  O4  gi_2
  C3  O3  C2  C4  gi_2
  C5  C2  C4  C1  gi_7
  C5  C2  C3  C1  gi_8
  C5  C2  C3  O5  gi_9

[ dihedrals ]
; ai  aj  ak  al  gromos type
  C8  C7  N2  C2  gd_4
  C3  C2  N2  C7  gd_14
  O5  C1  C2  N2  gd_8
  O3  C3  C2  N2  gd_8
  C2  C1  O5  C5  gd_14
  C4  C5  O5  C1  gd_14
 -C2 -O2  C1  C2  gd_14
 -O2  C1  C2  C3  gd_17
 -O2  C1  C2  C3  gd_7
 -O2  C1  C2  O2  gd_8
  O6  C6  C5  O5  gd_8
  C3  C4  C5  O5  gd_7
  O3  C3  C4  C5  gd_7
  C2  C3  O3  H32  gd_12
  C3  C4  O4  +C1  gd_14
  H63 O6  C6  C5  gd_12
  C2  C3  C4  C5  gd_17
  C6  C5  C4  C3  gd_17
  C6  C5  C4  O4  gd_7
  N2  C2  C3  C4  gd_7
  C1  C2  C3  O3  gd_7
  O3  C3  C4  O4  gd_8
  O5  C5  C4  O4  gd_8
  C4  C5  C6  O6  gd_7
  C4  C5  C6  O6  gd_17
  O5  C1  C2  C3  gd_7
  C1  C2  C3  C4  gd_17
  C2  C3  C4  O4  gd_7

```

**$\beta$ -Gal-(1 $\rightarrow$ 4)- $\beta$ -GlcNAc**

```

[ GAL1 ]
[ atoms ]
  C6  CH2    0.16000    0
  O6  OA    -0.40000    0
  H63 H      0.24000    0
  C5  CH1    0.24200    1
  O5  OA    -0.24200    1
  C1  CH1    0.25400    2
  C2  CH1    0.16000    3
  O2  OA    -0.40000    3
  H22 H      0.24000    3
  C3  CH1    0.16000    4
  O3  OA    -0.40000    4
  H32 H      0.24000    4
  C4  CH1    0.16000    5
  O4  OA    -0.40000    5
  H42 H      0.24000    5
[ bonds ]
  C6  C5    gb_25
  C5  O5    gb_19
  C5  C4    gb_25
  O5  C1    gb_19
  C1  C2    gb_25
  C2  O2    gb_19
  C2  C3    gb_25
  O2  H22   gb_1
  C3  O3    gb_19
  C3  C4    gb_25
  O3  H32   gb_1
  C4  O4    gb_19
  O4  H42   gb_1
  C6  O6    gb_19
  O6  H63   gb_1
[ angles ]
; ai  aj  ak  gromos type
  C6  C5  O5    ga_8
  C6  C5  C4    ga_7
  O5  C5  C4    ga_8
  C5  O5  C1    ga_9
  O5  C1  C2    ga_8
  C1  C2  O2    ga_8
  C1  C2  C3    ga_7
  O2  C2  C3    ga_8
 -04  C1  O5    ga_8
 -04  C1  C2    ga_8
  C2  O2  H22   ga_11
  C2  C3  O3    ga_8
  C2  C3  C4    ga_7
  O3  C3  C4    ga_8
  C3  O3  H32   ga_11
  C5  C4  O4    ga_8
  C4  O4  H42   ga_11
  C5  C6  O6    ga_8
  C6  O6  H63   ga_11
[ impropers ]
; ai  aj  ak  al  gromos type
  C1  -O4  O5    C2    gi_2
  C5  C6  C4    O5    gi_2
  C2  C3  C1    O2    gi_2
  C3  C4  O3    C2    gi_2
  C4  O4  C3    C5    gi_2
  C5  C2  C4    C1    gi_7
  C5  C2  C3    C1    gi_8
  C5  C2  C3    O5    gi_9
[ dihedrals ]
; ai  aj  ak  al  gromos type
  C2  C1  O5    C5    gd_14
  C3  C2  C1    O5    gd_7
  C1  C2  O2    H22   gd_12
  C4  C3  C2    C1    gd_17
  C2  C3  O3    H32   gd_12
  O4  C4  C3    C2    gd_7
  C5  C4  O4    H42   gd_12
  C2  C3  C4    C5    gd_17
  C1  O5  C5    C4    gd_14
  O5  C5  C4    C3    gd_7
  C6  C5  C4    C3    gd_17
  C6  C5  C4    O4    gd_7
  O5  C1  C2    O2    gd_8
  O2  C2  C3    C4    gd_7
  O2  C2  C3    O3    gd_8
 -C4  -O4  C1    C2    gd_14
 -O4  C1  C2    C3    gd_17
 -O4  C1  C2    C3    gd_7
 -O4  C1  C2    O2    gd_8
  C1  C2  C3    O3    gd_7
  O3  C3  C4    C5    gd_7
  O3  C3  C4    O4    gd_8
  O5  C5  C4    O4    gd_8
  C4  C5  C6    O6    gd_7
  C4  C5  C6    O6    gd_17
  O5  C5  C6    O6    gd_8
  C5  C6  O6    H63   gd_12

```

## Fuc-(1→O)-Ser

```

[ SER ]
[ atoms ]
  N      N      -0.28000    0      -C      N      CA      ga_30
  H      H      0.28000    0      N      CA      C      ga_12
  CA     CH1     0.00000    1      CA      C      +N     ga_18
  CB     CH2     0.17000    2      CA      C      O      ga_29
  OG     OA     -0.32000    2      O      C      +N     ga_32
  C      C      0.380      3      N      CA     CB     ga_12
  O      O     -0.380      3      C      CA     CB     ga_12
[ bonds ]
  N      H      gb_2
  N      CA     gb_20
  CA     C      gb_26
  C      O      gb_4
  C      +N     gb_9
  CA     CB     gb_26
  CB     OG     gb_17
[ angles ]
; ai    aj    ak    gromos type
  -C    N     H     ga_31
  H     N     CA    ga_17
[ impropers ]
; ai    aj    ak    al    gromos type
  N     -C    CA    H     gi_1
  C     CA    +N    O     gi_1
  CA    N     C     CB    gi_2
[ dihedrals ]
; ai    aj    ak    al    gromos type
  -CA   -C    N     CA    gd_4
  -C    N     CA    C     gd_19
  N     CA    C     +N    gd_20
  N     CA    CB    OG    gd_17

```

*Fuc*-(1→*O*)-*Ser*

```

[ FUC2 ]
[ atoms ]
  C6  CH3    0.00000    0
  C5  CH1    0.24200    1
  O5  OA    -0.24200    1
  C1  CH1    0.15000    2
  C2  CH1    0.16000    3
  O2  OA    -0.40000    3
H22  H      0.24000    3
  C3  CH1    0.16000    4
  O3  OA    -0.40000    4
H32  H      0.24000    4
  C4  CH1    0.16000    5
  O4  OA    -0.40000    5
H42  H      0.24000    5
[ bonds ]
  C6  C5      gb_25
  C5  O5      gb_19
  C5  C4      gb_25
  O5  C1      gb_19
  C1  C2      gb_25
  C2  O2      gb_19
  C2  C3      gb_25
  O2  H22     gb_1
  C3  O3      gb_19
  C3  C4      gb_25
  O3  H32     gb_1
  C4  O4      gb_19
  O4  H42     gb_1
[ angles ]
; ai  aj  ak  gromos type
  C6  C5  O5      ga_8
  C6  C5  C4      ga_7
  O5  C5  C4      ga_8
  C5  O5  C1      ga_9
  O5  C1  C2      ga_8
  C1  C2  O2      ga_8
  C1  C2  C3      ga_7
  O2  C2  C3      ga_8
  C2  O2  H22     ga_11
  C2  C3  O3      ga_8
  C2  C3  C4      ga_7
  O3  C3  C4      ga_8
  C3  O3  H32     ga_11
  C5  C4  C3      ga_7
  C5  C4  O4      ga_8
  C3  C4  O4      ga_8
  C4  O4  H42     ga_11
[ impropers ]
; ai  aj  ak  al  gromos type
  C5  C6  O5  C4      gi_2
  C2  C1  C3  O2      gi_2
  C3  C2  O3  C4      gi_2
  C4  O4  C5  C3      gi_2
  C5  C2  C4  C1      gi_4
  C5  C2  C3  C1      gi_5
  C5  C2  C3  O5      gi_6
[ dihedrals ]
; ai  aj  ak  al  gromos type
  O4  C4  C5  C6      gd_7
  C2  C1  O5  C5      gd_14
  C3  C2  C1  O5      gd_7
  C1  C2  O2  H22     gd_12
  C4  C3  C2  C1      gd_17
  C2  C3  O3  H32     gd_12
  O4  C4  C3  C2      gd_7
  C5  C4  O4  H42     gd_12
  C2  C3  C4  C5      gd_17
  C1  O5  C5  C4      gd_14
  O5  C5  C4  C3      gd_7
  C6  C5  C4  C3      gd_17
  O5  C1  C2  O2      gd_8
  O2  C2  C3  C4      gd_7
  O2  C2  C3  O3      gd_8
  C1  C2  C3  O3      gd_7
  O3  C3  C4  C5      gd_7
  O3  C3  C4  O4      gd_8
  O5  C5  C4  O4      gd_8

```

## Curriculum Vitae

### I. Formação acadêmica:

Graduação em Biomedicina pela Universidade Federal do Rio Grande do Sul, de 2004/1 a 2007/2. *Caracterização da geometria e flexibilidade da heparina através de simulações de dinâmica molecular*. Orientador: Prof. Hugo Verli.

### II. Trabalhos científicos apresentados em congressos

#### a. Nacionais:

Becker, C. F., **Pol-Fachin, L.**, Guimarães, J. A. & Verli, H. **2006** Conformational study and molecular dynamics simulations in aqueous solution of the glycan structure of antithrombin. *Programas e Resumos da XXXV Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, M-7. XXXV Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 1 a 4 de julho de 2006, Águas de Lindóia, SP, Brasil.

**Pol-Fachin, L.**, Verli, H. & Guimarães, J. A., **2006** Caracterização da geometria e flexibilidade da heparina através de simulações de dinâmica molecular. *Resumos do XVIII Salão de Iniciação Científica da Universidade Federal do Rio Grande do Sul*. Universidade Federal do Rio Grande do Sul, XVIII Salão de Iniciação Científica e XV Feira de Iniciação Científica da Universidade Federal do Rio Grande do Sul, 15 a 20 de outubro de 2006, Porto Alegre, RS, Brasil.

**Pol-Fachin, L.** & Verli, H. **2007** Caracterização das forças responsáveis pelo equilíbrio conformacional do resíduo IdoA na heparina. *Resumos do XIX Salão de Iniciação Científica da Universidade Federal do Rio Grande do Sul*. Universidade Federal do Rio Grande do Sul, XIX Salão de Iniciação Científica e XVI Feira de Iniciação Científica da Universidade Federal do Rio Grande do Sul, 21 a 26 de outubro de 2007, Porto Alegre, RS, Brasil.

Woicickoski, C., **Pol-Fachin, L.** & Verli, H. **2008** Análise conformacional de derivados N-acilidrazonas em solvente aquoso. *Resumos do XX Salão de Iniciação Científica da Universidade Federal do Rio Grande do Sul*. Universidade Federal do Rio Grande do Sul, XX Salão de Iniciação Científica e XVII Feira

de Iniciação Científica da Universidade Federal do Rio Grande do Sul, 20 a 25 de outubro de 2008, Porto Alegre, RS, Brasil.

Pedebos, C., **Pol-Fachin, L.** & Verli, H. **2009** Conformational characterization of saponins conformation in pyridine. *Programas e Resumos da XXXVIII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, M-34. XXXVIII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 16 a 19 de maio de 2009, Águas de Lindóia, SP, Brasil.

**Pol-Fachin, L.** & Verli, H. **2009** Effect of glycosylation over the structure and flexibility of proteins. *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.

#### **b. Internacionais:**

**Pol-Fachin, L.** & Verli, H. **2007** Characterization of the forces responsible for iduronic acid conformational equilibrium. *Programas e Resumos da XXXVI Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular e X Conferência da International Union of Biochemistry and Molecular Biology*, M-52. XXXVI Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular e X Conferência da International Union of Biochemistry and Molecular Biology, 21 a 25 de maio de 2007, Salvador, BA, Brasil.

Fernandes, C. L., **Pol-Fachin, L.** & Verli, H. **2008** Prediction of disaccharides conformational ensembles in solution from sequence. *Programas e Resumos da XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular e XI Congresso da Pan-American Association for Biochemistry and Molecular Biology*, M-1. XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular e XI Congresso da Pan-American Association for Biochemistry and Molecular Biology, 17 a 20 de maio de 2008, Águas de Lindóia, SP, Brasil.

**Pol-Fachin, L.**, Fernandes, C. L. & Verli, H. **2008** Prediction of glycoproteins conformational ensembles in solution. *Programas e Resumos da XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular e XI Congresso da Pan-American Association for Biochemistry and Molecular Biology*, M-35. XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica

e Biologia Molecular e XI Congresso da Pan-American Association for Biochemistry and Molecular Biology, 17 a 20 de maio de 2008, Águas de Lindóia, SP, Brasil.

Woicickoski, C., **Pol-Fachin, L.**, Fraga, C. A. M., Barreiro, E. J. & Verli, H. **2008** Description of the conformational ensemble of N-acylhydrazone derivatives in aqueous solutions. *Programas e Resumos da XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular e XI Congresso da Pan-American Association for Biochemistry and Molecular Biology*, V-7. XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular e XI Congresso da Pan-American Association for Biochemistry and Molecular Biology, 17 a 20 de maio de 2008, Águas de Lindóia, SP, Brasil.

### III. Publicações em periódicos especializados:

#### a. Internacionais:

**Pol-Fachin, L.** & Verli, H. **2008** Depiction of the forces participating in the 2-O-sulfo- $\alpha$ -L-iduronic acid conformational preference in heparin sequences in aqueous solutions *Carbohydr. Res.*, *343*, 1435-1445.

**Pol-Fachin, L.**, Fernandes, C. L. & Verli, H. **2009** GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations *Carbohydr. Res.*, *344*, 491-500.

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. **2009** 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.*, *284*, 18790-18800.

**Pol-Fachin, L.**, Fraga, C. A. M., Barreiro, E. J. & Verli, H. **2010** Characterization of the conformational ensemble from bioactive N-acylhydrazone derivatives *J. Mol. Graph. Mod.*, *28*, 446-454.

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. Orientações de Iniciação Científica:**

##### **a. Concluídas:**

Clóvis Woicickoski Júnior. *Análise conformacional de derivados N-acilidrazonas em solvente aquoso*. Centro de Biotecnologia, UFRGS, de março de 2008 a agosto de 2008.

##### **b. Em andamento:**

Conrado Pedebos. *Caracterização conformacional de saponinas em piridina*. Centro de Biotecnologia, UFRGS, a partir de julho de 2008.

#### **V. Bolsa recebida:**

Bolsista FAPERGS de setembro de 2005 a julho de 2006.

Bolsista CNPq – PIBIC de agosto de 2006 a fevereiro de 2008.

Bolsista CAPES do programa de mestrado em biologia molecular e celular pelo Centro de Biotecnologia da UFRGS a partir de março de 2008.