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

GLICOBIOLOGIA ESTRUTURAL DA MODULAÇÃO DA CASCATA DE COAGULAÇÃO SANGUÍNEA POR HEPARINA

Laércio Pol Fachin

Porto Alegre – Brasil
Julho de 2013

GLICOBIOLOGIA ESTRUTURAL DA MODULAÇÃO DA CASCATA DE COAGULAÇÃO SANGUÍNEA POR HEPARINA

Laércio Pol Fachin

Tese de doutorado elaborada no Grupo de Bioinformática Estrutural do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul sob co-orientação e orientação, respectivamente, dos professores doutores:

Roberto Dias Lins Neto

Hugo Verli

Porto Alegre – Brasil

Julho de 2013

GLICOBIOLOGIA ESTRUTURAL DA MODULAÇÃO DA CASCATA DE COAGULAÇÃO SANGUÍNEA POR HEPARINA

Laércio Pol Fachin

Tese 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 Doutor em Biologia Celular e Molecular.

Banca Examinadora:

_____

Hugo Verli (Centro de Biotecnologia – UFRGS, Brasil) (Presidente)

_____

Célia Regina Carlini (Centro de Biotecnologia – UFRGS, Brasil)

_____

Fernando Ferreira Chiesa (Faculdad de Química – UDeLaR, Uruguai)

_____

José Raúl Grigera (IFLYSIB CONICET – UNLP, Argentina)

_____

Diego Bonatto (Centro de Biotecnologia – UFRGS, Brasil) (Suplente)

Esta tese 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 Doutor em Biologia Celular e Molecular, junto ao Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul.

FICHA CATALOGRÁFICA

POL-FACHIN, Laércio.

GLICOBIOLOGIA ESTRUTURAL DA MODULAÇÃO DA CASCATA DE COAGULAÇÃO SANGUÍNEA POR HEPARINA */ Laércio Pol Fachin. – 2013*

Orientador: Hugo Verli

Coorientador: Roberto Dias Lins Neto

Tese (Doutorado) – Universidade Federal do Rio Grande do Sul, Centro de Biotecnologia do Estado do Rio Grande do Sul, Programa de Pós-Graduação em Biologia Celular e Molecular, Porto Alegre, BR-RS, 2013.

1. Dinâmica Molecular 2. Heparina 3. Glicoproteínas 4. Coagulação

I. Verli, Hugo – orient.

II. Lins, Roberto Dias – coorient.

III. Título.

AGRADECIMENTOS

Minha imensa gratidão ao Prof. Hugo Verli, por ter me acolhido junto ao Grupo de Bioinformática Estrutural e por ter guiado meus passos nesses primeiros anos de trabalho na *arte* da Modelagem Molecular. Agradeço por toda confiança em mim depositada ao longo desses oito anos, pelos ensinamentos e conselhos, sempre visando ao bem do aluno, do ambiente de trabalho e do crescimento profissional e científico do grupo.

Meus sinceros agradecimentos ao Prof. Roberto Lins, pela acolhida sempre atenciosa em Recife, tenha sido ela de cunho acadêmico, científico, turístico ou gastronômico, e pelos aconselhamentos sempre lúcidos e serenos.

Aos professores Arthur Germano Fett-Neto e Hubert Karl Stassen, membros da Comissão de Acompanhamento, pela prestatividade, pelos conselhos e auxílio desde a minha defesa de Dissertação de Mestrado e ao longo deste período.

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

Aos colegas e amigos do Grupo de Bioinformática Estrutural, pela convivência sempre harmoniosa e por todo diálogo e companheirismo, tanto no laboratório quanto nos eventos fora dele, dos quais destaco, e levo gravada na memória, toda descontração e diversão nos congressos da SBBq.

Agradecimentos especiais à Cláudia Lemelle Fernandes e Conrado Pedebos, com quem também pude dividir momentos de preparação de manuscritos e tensão no aguardo das opiniões dos revisores e corpos editoriais.

À Maria Carolina de Araujo Melo, por toda força e apoio na reta final dos trabalhos e preparativos desta Tese.

Aos alunos do Laboratório de Biomateriais da Universidade Federal de Pernambuco, por toda a hospitalidade quando das minhas visitas à Recife.

Ao Luciano e Sílvia, secretários do PPGBCM, por toda a prestatividade, atenção e confiabilidade no pedido de informações, no atendimento e solução de problemas.

Aos membros da banca, por terem aceitado o convite.

A minha família, que sempre me apoiou em todos os momentos ao longo desta jornada. Em especial, a minha irmã Thais, pelo grande exemplo de esforço e dedicação aos estudos. E, principalmente, aos meus pais, Maria Fátima e Luiz - pelos conselhos, sempre sábios e sensatos, pelo suporte emocional e racional em todas as escolhas e decisões, e pelo constante e inabalável incentivo para a realização deste trabalho.

"*Se trabalharmos sobre o mármore, um dia ele acabará. Se trabalharmos sobre o metal, um dia o tempo o consumirá. Se erguermos templos, um dia eles se tornarão pó. Mas se trabalharmos sobre almas jovens e imortais, se nós as imbuirmos com os princípios do amor à humanidade, daqui a cem anos pouco importará o quanto tenhamos acumulado no banco; que tipo de casa, palacete ou carro possuímos. Mas o mundo poderá ser diferente, talvez porque fomos importantes na vida dos jovens.*"

Adaptado de Frank Sherman Land.

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

SUMÁRIO

LISTA DE ABREVIATURAS

AT – *antithrombin*, ou antitrombina

DM – dinâmica molecular

EGF – *epidermal growth factor*, ou fator de crescimento epidermal

FES – *free energy surface*

FT – fator tecidual

Fuc – fucose

GAG – glicosaminoglicano

Gal – galactose

GalNAc – *N*-acetil-galactosamina

Glc – glicose

GlcA – ácido glicurônico

GlcN – glicosamina

GlcNAc – *N*-acetil-glicosamina

IdoA – ácido idurônico

ITP – *individual topology parameters*

Man – manose

MD – *molecular dynamics*

MM – mecânica molecular

NeuAc – ácido neuramínico, ou ácido siálico

NOE – *nuclear Overhauser effect*

PDB – *protein data bank*

PME – *particle-mesh ewald*

RE – retículo endoplasmático

RCL – *reactive center loop*

SPC – *simple point-charge*

TF – *tissue factor*

TFPI – *tissue factor pathway inhibitor*, ou inibidor da via do fator tecidual

RESP – *restrained electrostatic potential*

RMN – ressonância magnética nuclear

RTP – *residue topology parameters*

Xyl – xilose

RESUMO

Antitrombina (AT), uma proteína membro da família dos inibidores de serino proteases, é uma glicoproteína que co-existe em duas isoformas, α e β, que se diferenciam pelo conteúdo de glicosilação e pela afinidade por glicosaminoglicanos (GAG), um grupo de polissacarídeos polisulfatados, dentre as quais se destaca a heparina. AT é ativada quando ligada a GAGs, tornando-se assim capaz de inibir, com alta eficiência, proteases da cascata de coagulação como trombina e fXa. Essas enzimas formam complexos ternários com heparina e AT, sendo cada uma subsequentemente inibidas preferencialmente por um mecanismo de ação distinto: [1] baseado em mudanças conformacionais (fXa), ou pelo [2] mecanismo de ponte (trombina). Adicionalmente, já foi observado que heparina isoladamente pode modular a atividade catalítica de fIIa e fXa. Considerando a falta de dados estruturais a respeito dos efeitos da glicosilação sobre a estrutura e flexibilidade de AT, bem como sobre o reconhecimento heparina-AT, e que as bases moleculares da inibição alostérica de fIIa e fXa por GAGs não é bem compreendida, o presente trabalho visa caracterizar o reconhecimento molecular de heparina por essas proteínas, através de dinâmica molecular (DM). Os resultados obtidos indicam que a heparina interage de forma diferente nas glicoformas de AT devido a uma interferência da glicana ligada à Asn135. Da mesma forma, diferentes arranjos do GAG na superfície de trombina e fXa podem estar relacionadas às suas diferentes susceptibilidades aos mecanismos de ação de heparina, uma vez que sua orientação na superfície de fIIa, mas não fXa, permite uma interação adequada com AT em complexos ternários. Nesse contexto, foi observado, durante as simulações, que heparina inibiu alostericamente trombina e fXa, promovendo mudanças na conformação da tríade catalítica das proteases, e ativou AT, promovendo rearranjos intramoleculares entre seus elementos de estrutura secundária. Em ambos os casos, as vias de transmissão de sinal associadas a essas mudanças de atividade foram traçadas pela primeira vez nesse trabalho. De forma geral, os resultados obtidos conferem novas evidências de que a glicosilação tem um papel crucial na ativação diferencial de α- e β-AT por heparina, e que a orientação de GAGs na superfície de trombina e fXa é o que determina o mecanismo de sua modulação por heparina.

ABSTRACT

Antithrombin (AT), a member of the serpin protease inhibitors family, is a glycoprotein that co-exists in two isoforms, α and β, which differ in their amount of glycosylation and affinity for glycosaminoglycans (GAG), a group of sulphated polysaccharides, as heparin. AT is activated when bound to GAGs, becoming capable to inhibit coagulation cascade serine proteases like thrombin and fXa. Such enzymes form ternary complexes with heparin and AT, being subsequently inhibited by two distinct mechanisms of action: [1] the conformational change-based mechanism or the [2] bridge mechanism. In addition, heparin binding itself was also observed to modulate the catalytic activity of both fIIa and fXa. Considering the lack of structural information regarding the effect of glycosylation over the structure and flexibility of AT, as well as in heparin-AT recognition, and that the molecular basis of the allosteric inhibition of fIIa and fXa, promoted by such GAG, is not fully understood, the current work intends to characterize the molecular recognition of heparin by such proteins through molecular dynamics (MD) simulations. The obtained results indicate that heparin binds differently on AT glycoforms due to an interference of Asn135-linked glycan. As well, distinct arrangements of the polysaccharide on the surface of thrombin and fXa may be related to their diverse susceptibilities to heparin mechanisms of action, as heparin orientation observed on the surface of fIIa, but not fXa, allows for an adequate long chain heparin binding to AT in ternary complexes. In this context, heparin was observed to allosterically inhibit thrombin and fXa by promoting changes in the proteases catalytic triad conformation, but to activate AT by promoting intramolecular rearrangements on its secondary structure elements. In both cases, the signal transmission pathways associated with such activity changes were traced by the present work for the first time. Altogether, the obtained results provide new evidences that glycosylation play a central role on the differential activation of α- and β-AT by heparin, and that GAGs orientation on the surface of thrombin and fXa determine the mechanism of their modulation by heparin.

ÍNDICE DE FIGURAS

ÍNDICE DE TABELAS

1 Introdução

1.1 Hemostasia

A hemostasia consiste em um processo fisiológico que controla a fluidez do sangue, sendo capaz de rapidamente responder a perturbações no endotélio vascular de forma a evitar o extravasamento de sangue (Vine, 2009). Essa fluidez é obtida através do balanço entre forças pró-coagulantes, responsáveis pela formação do tampão que irá impedir a perda sanguínea em caso de lesões aos vasos, e as anticoagulantes, que prevalecem em circunstâncias normais (Dahlbäck, 2000), cuja função é prevenir que a coagulação ocorra em situações inapropriadas e de forma exagerada (Segers *et al.*, 2007). Quando da ocorrência de um dano físico ao vaso, as forças pró-coagulantes se sobrepõe em um processo que envolve a agregação plaquetária na região danificada do vaso, aliada à formação de complexos enzimáticos estruturados a partir da membrana celular dessas plaquetas e ao redor da lesão (Rasche, 2001). Segundo o mais recente modelo da cascata de coagulação, baseado em superfícies celulares (Hoffman & Monroe III, 2001), esse processo ocorre em três etapas: iniciação (Figura 1A, amarelo), amplificação (Figura 1B, verde) e propagação (Figura 1C, azul).

1.1.1 Mecanismos pró-coagulantes

Como consequência de danos a vasos sanguíneos, células extravasculares como fibroblastos e células de musculatura lisa são expostas ao sangue circulante. A partir desse contato, o processo de coagulação é disparado, através da interação entre (1) uma glicoproteína transmembrana conhecida como fator tecidual (FT, do inglês *tissue factor* – TF), presente na superfície dessas células (porém ausente na superfície externa de tecido vascular sadio), e (2) fator VII sanguíneo (Segers *et al.*, 2007). Com relação a esta última proteína, uma enzima da família das serino proteases, sua interação com FT pode ocorrer através de qualquer das duas formas circulantes no plasma, de enzima ativada ou de zimogênio (Segers *et al.*, 2007), que é convertido rapidamente a forma de enzima ativa após a interação (Wildgoose & Kisiel, 1989). Uma vez iniciada, a coagulação do sangue ocorre ao redor da superfície dessas células contendo fVIIa ligado a FT, no qual forma-se um ambiente protegido dos mecanismos anticoagulantes plasmáticos (Vine, 2009). Esse complexo, fVIIa-FT, possui atividade catalítica sobre os fatores IX e X, gerando fIXa,

que permanece em grande parte aderido à superfície celular, e fXa, que pode difundir para o plasma, onde é rapidamente inibido por mecanismos anticoagulantes ou permanecer aderido à célula (Hoffman & Monroe III, 2001). Nesse caso, fXa se combina com fVa, previamente ativado a partir de fV pelo próprio fXa (Vine, 2009), para converter pequenas quantidades de protrombina em trombina (fIIa).

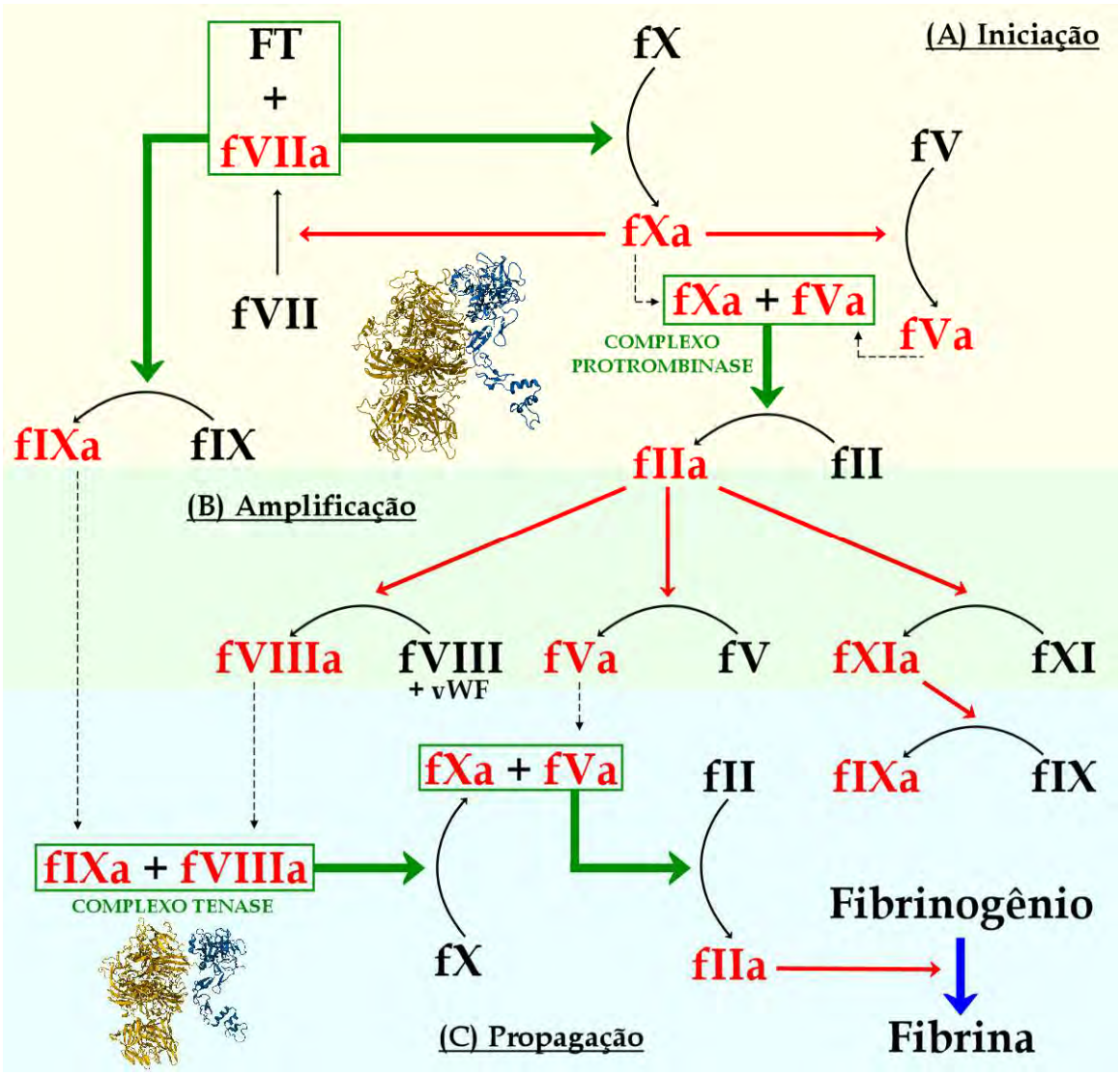


Figura 1: Representação esquemática do processo de coagulação sanguínea, envolvendo as fases de iniciação (A, amarelo), amplificação (B, verde) e propagação (C, azul). A atividade enzimática de proteínas e complexos proteicos está ilustrada por setas. As proteínas inativas estão apresentadas em preto, enquanto que as proteínas ativas e suas ações são mostradas em vermelho. A procedência das proteínas que compõe cada complexo enzimático está indicada por setas pontilhadas. Os complexos enzimáticos e suas ações são mostrados em verde.

Concomitantemente a isso, também em consequência dos danos ao vaso sanguíneo, fibras de colágeno são expostas ao sangue circulante (Segers *et al.*, 2007). A partir dessa exposição, um movimento de agregação plaquetária a essas fibras ocorre, mediado por fator *von Willebrand* circulante e receptores de colágeno presentes na superfície dessas plaquetas (Segers *et al.*, 2007). Esse processo ativa parcialmente essas células, da mesma forma que as aproxima dos sítios onde FT está localizado (Hoffman & Monroe III, 2001) e, consequentemente, da trombina gerada na etapa de iniciação da coagulação. Essa enzima apresenta um papel crucial na amplificação do processo de adesão, ativação e agregação plaquetária, da mesma forma que ativa os fatores da coagulação V, VIII e XI, gerando fVa, fVIIIa e fXIa (Hoffman & Monroe III, 2001; Vine, 2009).

A partir desses eventos, ocorre a propagação do processo de coagulação sanguínea, em que trombina é gerada em grande escala. Isso ocorre principalmente na superfície das plaquetas ativadas, que expressam sítios de ligação de alta afinidade pelos fatores fIXa, fXa e fXIa (Hoffman & Monroe III, 2001), envolvidos na composição e ativação das enzimas que formam o complexo protrombinase, responsável pela geração de fIIa. Especificamente, o complexo tenase (Figura 1, estrutura por Autin *et al.*, 2005), cuja função é ativar fX a fXa, é formado por fVIIIa, gerado por trombina na etapa de amplificação, e fIXa, gerado na etapa de iniciação ou convertido por fXIa presente na superfície celular das plaquetas (Hoffman & Monroe III, 2001). O fXa recém gerado a partir do complexo tenase irá se combinar com fVa, ativado por trombina na etapa de amplificação, formando o complexo protrombinase (Figura 1, estrutura por Lee *et al.*, 2008), que produz trombina em grandes quantidades (Hoffman & Monroe III, 2001).

Na última etapa da formação do tampão, trombina é responsável pela conversão de fibrinogênio, uma glicoproteína circulante no plasma sanguíneo, em fibrina, que irá se combinar entre si na forma de um agregado insolúvel (Segers *et al.*, 2007). Trombina, da mesma forma, é responsável pela formação de fXIII, um fator estabilizador de fibrina que possui atividade transglutaminase (Vine, 2009). Esse fator forma ligações entre fibras adjacentes de fibrina e incorpora outras proteínas ao agregado, como α_2-antiplasmina e fibronectina (Alzahrani & Ajjan, 2010). Essa ação, por sua vez, fortalece a agregação plaquetária, ao passo que a malha de fibrina se liga às plaquetas (Rasche, 2001). Após a formação do tampão, tem início uma recuperação dos tecidos componentes da parede do vaso danificado

o que, à medida que é concluído, permite a ação da plasmina, o produto da ativação de plasminogênio, em um processo conhecido como fibrinólise, em que os depósitos de fibrina são removidos das paredes vasculares (Segers *et al*., 2007).

1.1.2 Mecanismos anticoagulantes

Sob condições fisiológicas, nas quais não há perturbações ao sistema vascular, as plaquetas não se aderem aos vasos sanguíneos, não se agregam nem se ligam a fatores pró-coagulantes circulantes no plasma em proporções que levem à formação de coágulos (Rasche, 2001). Isso ocorre devido à ação de vários fatores endoteliais que contribuem para manter a permeabilidade vascular, evitando a formação desses coágulos, tais como inibidores e moduladores da agregação plaquetária, da coagulação sanguínea e da fibrinólise (Tabela 1), além dos próprios efeitos de diluição do sangue circulante (Rasche, 2001; Green, 2006).

Tabela 1: Elementos do endotélio que contribuem na permeabilidade vascular.

Ação Biológica	Componente do endotélio*
Inibição da ativação plaquetária	Adenosina difosfatase
	CD39 (ADPase/ATPase)
	Óxido nítrico
	Prostaciclina (PGI_2)
Inibição da coagulação sanguínea	Trombomodulina
	Heparan sulfato
Modulação da fibrinólise	Ativador de plasminogênio tecidual
	Ativador de plasminogênio tipo uroquinase

* Dados adaptados de Rasche, 2001 e Green, 2006.

Por exemplo, trombomodulina é uma glicoproteína transmembrana que atua como um receptor de trombina na superfície vascular. A interação entre essas duas moléculas diminui significativamente a atividade de trombina na ativação de plaquetas e na geração dos fatores Va e VIIIa durante a fase de amplificação da coagulação sanguínea (Li *et al.,* 2012). Adicionalmente, proteínas circulantes no plasma se somam a esses fatores endoteliais para manter a fluidez do sangue, dentre os quais se destacam o inibidor da via do fator tecidual (do inglês *tissue factor pathway inhibitor*, TFPI), proteína C ativada e a família das serpinas.

O TFPI consiste em um regulador da fase de iniciação da coagulação sanguínea, atuando através da inibição da atividade catalítica de fXa e sobre o complexo fVIIa-FT (Broze *et al.,* 1988; Hackeng *et al.,* 2009). É um inibidor de proteases da família Kunitz, apresentando 3 domínios de mesmo nome em sua estrutura, além de uma região N-terminal ácida e uma região C-terminal básica (Hackeng *et al.,* 2009). Seu mecanismo de ação principal deve-se à formação de um complexo quaternário envolvendo fXa, fVIIa e FT (Girardi *et al*., 1989), embora se saiba que proteína S, um cofator de proteína C ativada, possa também acelerar sua inibição específica sobre fXa (Hackeng *et al.,* 2009).

A proteína C ativada, por sua vez, é uma serino protease, tal como outras enzimas envolvidas na coagulação do sangue (isto é, fatores IIa, VIIa, IXa, Xa, XIa e XIIa). Sua atividade catalítica está na degradação seletiva de fVa e fVIIIa, o que inativa, por consequência, os complexos protrombinase e tenase, respectivamente (Dahlbäck & Villoutreix, 2005). Ela é gerada a partir de proteína C, através da ação do complexo trombomodulina-trombina (Li *et al.,* 2012) e, tal como TFPI, possui a proteína S como cofator, cujo mecanismo molecular de ação ainda não está bem compreendido (Malm *et al*., 2008).

Já as serpinas são inibidoras de serino proteases (do inglês *SERine Proteases INhibitor, serpin*) que possuem ampla distribuição na natureza e apresentam uma grande similaridade estrutural, a despeito de uma ampla diversidade funcional (Silverman *et al.,* 2001) e, por vezes, baixa identidade entre sequências de aminoácido (Gettins, 2002). Nesse sentido, embora representantes dessa classe de proteínas tenham sido originalmente caracterizados como inibidores de enzimas do tipo serino proteases, diversas serpinas são assim denominadas sem apresentarem atividade inibitória, dentre as quais estão angiotensinogênio, que atua na regulação da pressão sanguínea (Potempa *et al*., 1994) e as globulinas ligadoras de tiroxina e corticoesteróides, que agem no transporte de hormônios (Silverman *et al.,* 2001). Dentre aquelas com atividade inibitória, e que atuam sobre proteases da coagulação, destacam-se o cofator 2 de heparina e a antitrombina (AT), um dos alvos de estudo da presente tese. Ambas inibem seus alvos enzimáticos através de uma reação modulada por glicosaminoglicanos (GAG), como heparina (terapeuticamente) e heparan sulfato (fisiologicamente) (Gettins, 2002), cujo mecanismo de inibição, tais como de outras serpinas, requer uma drástica mudança conformacional após o contato com a protease alvo (Whisstock & Bottomley, 2008).

1.2 Inibição de proteases da coagulação por AT

No caso de AT, ocorrem dois tipos de alterações conformacionais na estrutura da serpina durante o processo de inibição das enzimas da coagulação: no primeiro, em que as proteases não necessariamente estão ligadas à AT, a heparina promove uma ativação alostérica na serpina, potencializando sua atividade inibitória (Whisstock *et al.*, 2000) e, no segundo, essas mudanças conformacionais na serpina levam as enzimas à inativação (Huntington *et al.*, 2000a). Destacadamente, um conjunto de folhas-β, chamado de folha-β A (Figura 2, laranja), participa nestas duas etapas de mudança conformacional (McCoy *et al.,* 2003).

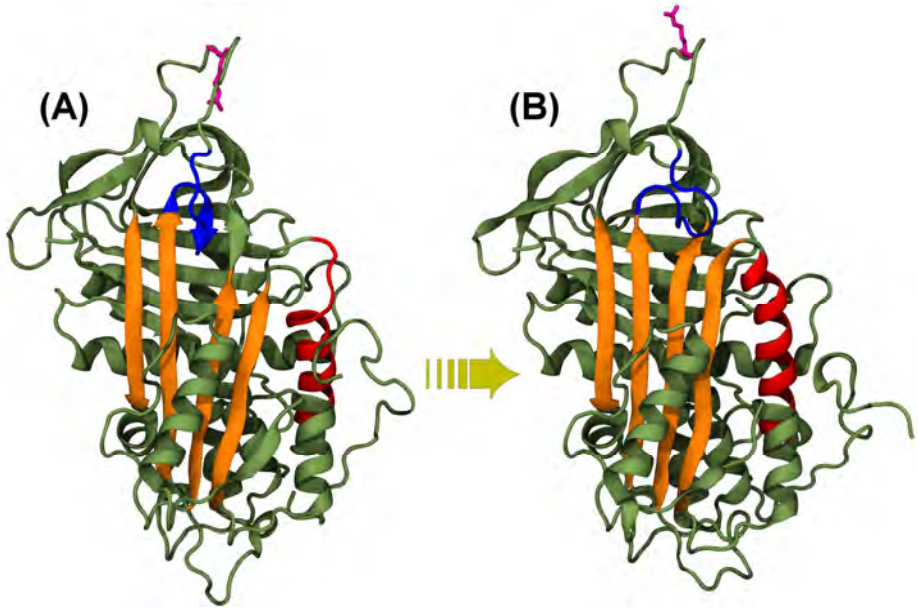


Figura 2: Ativação alostérica de (A) AT nativa à (B) AT ativada, mediada por GAG. As estruturas estão representadas em *cartoon*, enquanto que elementos importantes no processo de ativação de AT estão destacados em cor: hélice D (vermelho), β-folha A (laranja), região *hinge* do RCL (azul) e o aminoácido Arg393 (magenta).

No início dessa cascata de eventos, a estrutura de AT, chamada de *nativa*, apresenta três conjuntos de folhas-β, nove α-hélices (Figura 2A) e várias alças, algumas bastante longas, seja interligando esses elementos de estrutura secundária (contendo até aproximadamente 25 aminoácidos), ou localizada na região N-terminal da serpina, apresentando cerca de 50 resíduos (McCoy *et al.,* 2003). A partir da ligação do GAG à serpina, como heparina ou heparan sulfato, a AT é ativada alostericamente, ocasionando uma série de modificações na estrutura terciária da proteína, envolvendo o rearranjo de diversos elementos de estrutura secundária da

proteína (Langdown *et al.*, 2004). As três alterações principais envolvem [1] o aumento da extensão, a C-terminal, da chamada α-hélice D (Figura 2, vermelho), que passa de onze (His120 a Tyr131) a dezoito (Gln118 a Lys136) resíduos componentes (Belzar *et al.*, 2002), [2] a expulsão da região *hinge* (Figura 2, azul) de dentro da folha-β A (Langdown *et al.*, 2004), o que leva a [3] contração da folha-β A e sua maior organização. A região *hinge* consiste na porção N-terminal de uma alça específica, conhecida como alça reativa central (do inglês *reactive center loop*, RCL). Como consequência desses eventos, há um aumento da afinidade da AT pelo GAG (Jin *et al.*, 1997) e a exposição do principal resíduo de aminoácido relacionado à inativação das serino proteases, Arg393 (Figura 2, magenta).

No processo de inibição de enzimas da coagulação, GAGs podem exercer seu papel na potencialização da atividade inibitória de AT através de dois mecanismos de ação distintos (Gettins, 2002; Olson *et al.*, 2004). Em um deles, conhecido com mecanismo de ponte (Figura 3A), uma cadeia polissacarídica de GAG de no mínimo 18 resíduos se liga concomitantemente à serpina e à protease alvo. Essa interação entre polissacarídeo e as proteínas ocorre principalmente através de interações iônicas entre grupamentos sulfato (SO_3^-), sulfonamida ($NHSO_3^-$) e carboxilato (COO^-) do GAG, e resíduos de aminoácido de carga positiva, como Arg e Lys presentes em AT e na enzima. O outro mecanismo, conhecido como mecanismo de ativação conformacional ou baseado em mudanças conformacionais (Figura 3B), uma sequência de GAG de no mínimo 5 resíduos se liga exclusivamente à serpina (Gettins, 2002; Olson *et al.*, 2004).

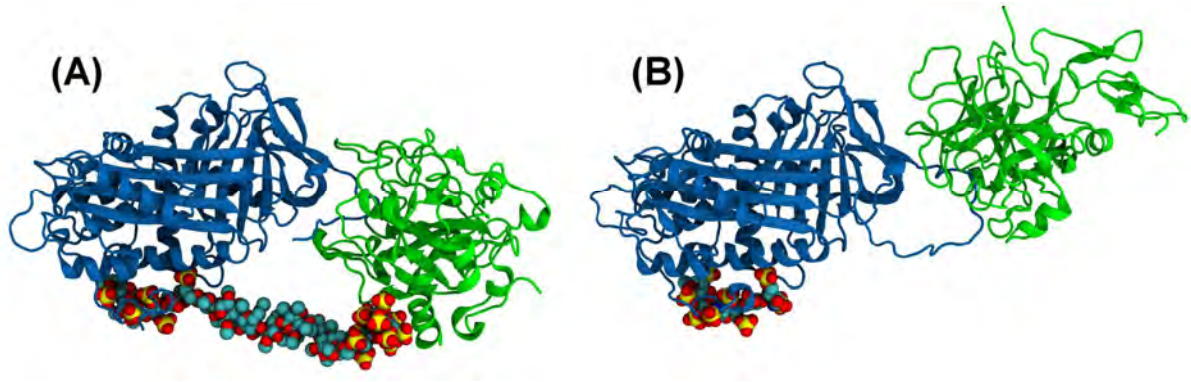


Figura 3: Mecanismos de ação da heparina (A-B) na inibição das proteases da coagulação. As estruturas protéicas estão representadas em *cartoon*, enquanto que heparina está ilustrada em esferas sólidas de *van der Waals* (VDW). AT está colorida de azul, e as proteases alvo estão mostradas em verde.

Com relação às principais proteases alvo da AT, trombina é somente inativada através do mecanismo de ponte (Laurent *et al.*, 1978; Oosta *et al.*, 1981; Lane *et al.*, 1984), o que indica que essa enzima não é sensível às mudanças conformacionais induzidas em AT por heparinas que contenham menos de 18 monossacarídeos, e que, de fato, trombina depende da interação com GAG para ser inibida. O fator Xa, por outro lado, é mais susceptível ao mecanismo de ativação conformacional (Choay *et al.*, 1983), embora esta enzima seja também inibida através do mecanismo de ponte na presença de uma concentração ótima (aproximadamente 50nM) de cálcio (Rezaie, 1998).

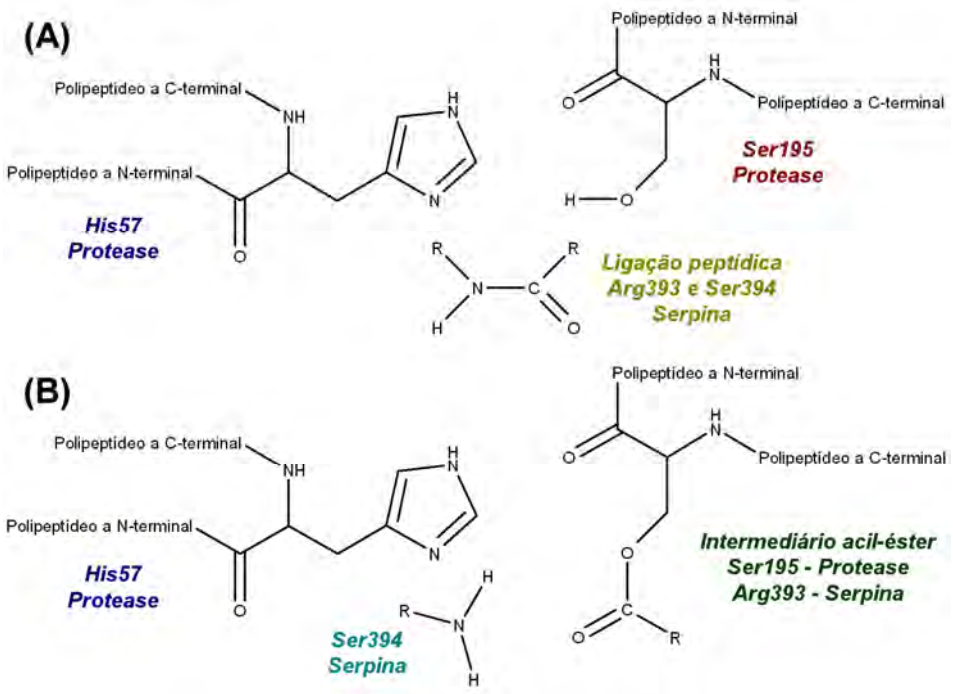


Figura 4: Representação esquemática do (A) estado inicial do reconhecimento do sítio ativo das proteases alvo com AT e do (B) intermediário acil-éster, formado pela ação enzimática incompleta das serino proteases.

No início do processo de inativação, essas proteases reconhecem uma sequência de aminoácidos da serpina, localizada no RCL (delimitada pelos resíduos Asn376 e Arg393), mas especificamente (Figura 4A) centralizada no resíduo de Arg393 (Gettins, 2002; Olson *et al.*, 2010). Fisiologicamente, a sequência consenso preferencial de reconhecimento de trombina demanda fortemente um resíduo de Pro anterior à Arg (Harris *et al.*, 2000), e pode ser sintetizado em Pro–Arg–Ser/Ala/Gly/Thr–X–Arg, onde X é qualquer aminoácido não ácido (Gallwitz *et al.*, 2012); a sequência consenso de reconhecimento de fator Xa é mais flexível, porém

pode ser resumida à Y–Z–Gly–Arg, onde Y representa aminoácidos alifáticos, tais como Ile, e Z é qualquer aminoácido menos Pro (Harris *et al.*, 2000). Em AT, a sequência ao redor do sítio de reconhecimento das proteases é Ile-Ala-Gly-Arg-Ser-Leu-Asn e, note-se, a presença do resíduo de Arg, centralizado na Arg393, é uma constante nas três sequências apresentadas. Após a interação entre a protease e AT (Figura 5A), a reação de clivagem dessa alça não ocorre de forma completa, uma vez que ela é interrompida em um estado acil-éster intermediário (Figura 4B), na qual o RCL já foi clivado, mantendo, porém, serpina e a enzima-alvo ligadas covalentemente através do resíduo catalítico de Ser195 da protease e da Arg393 de AT (Gettins, 2002; Olson *et al.*, 2010).

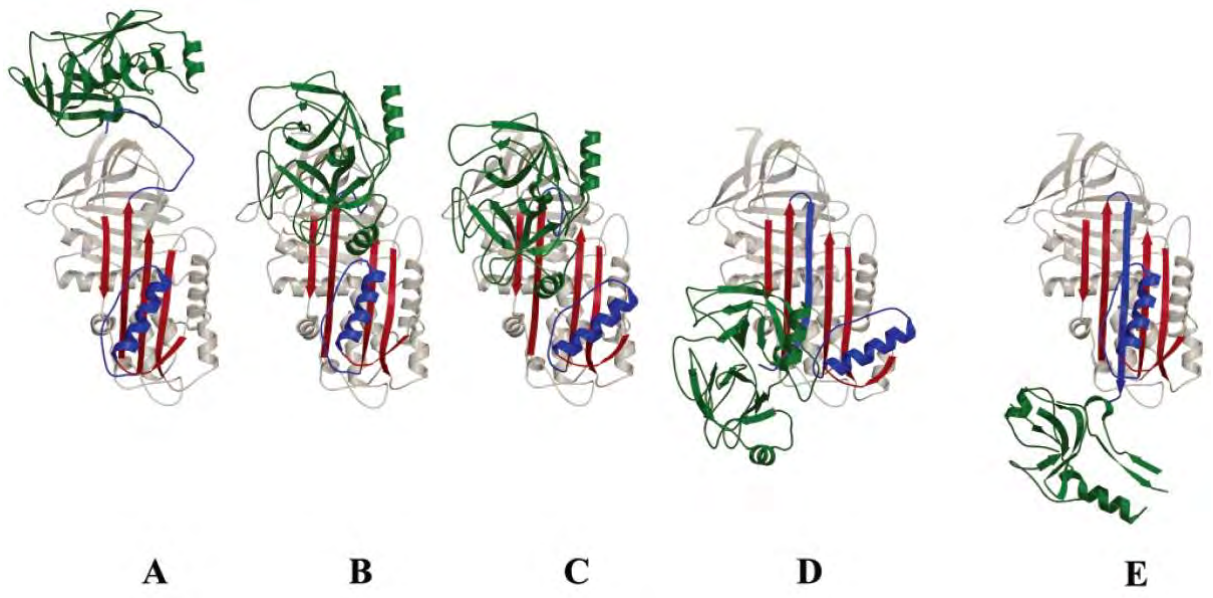


Figura 5: Representação esquemática da inibição de proteases da coagulação por AT. As estruturas estão representadas em *cartoon*, em que AT está colorida de cinza (com a folha-β A em vermelho e o RCL, que se insere em meio à folha-β A, em azul) e a protease alvo ilustrada em verde. Adaptado de Gettins, 2002.

Essa ligação covalente entre serpina e enzima é um atributo crucial para que AT exerça seu mecanismo de inibição sobre as proteases alvo. Nesse sentido, as serpinas são substratos suicidas, ou seja, após a inativação da enzima, nem AT nem a protease alvo possuirão atividade. E, nesse processo, o RCL apresenta um papel central (Gettins, 2002), uma vez que o trecho entre Asn376 e Arg393 se insere em meio à folha-β A (Figura 5D), movendo a enzima alvo de uma extremidade a outra

de AT (Figura 5E) e prendendo-a contra sua estrutura (Huntington *et al.*, 2000a). Dessa forma, AT promove uma grande mudança conformacional na forma global do complexo serpina-protease (Figura 5), levando a protease alvo a uma provável inibição por deformação estrutural (Huntington *et al.*, 2000a). A essa deformarção inclui-se uma quebra da organização da tríade catalítica, o que poderia estar relacionado à incapacidade da serino protease em seguir com seu mecanismo catalítico, após o intermediário acil-éster (Huntington *et al.*, 2000a). Adicionalmente, a inserção do RCL na folha-β A é um fator estabilizador do complexo AT-enzima (Gettins, 2002), porém a deformação da protease aumenta sua susceptibilidade à proteólise, o que pode também estar associado a sua maior depuração após a inibição por AT (Huntington *et al.*, 2000a).

1.3 O papel de carboidratos na coagulação

A hemostasia, incluindo a inibição das proteases da cascata de coagulação sanguínea por AT, tal como em qualquer outro evento biológico, envolve em grande parte a ação e a interação entre proteínas. No entanto, carboidratos estão profundamente associados a esses eventos e por vezes, inclusive, atuam como protagonistas. Na natureza, além do crucial envolvimento de glicose (Glc) na geração da energia a ser utilizada pelo organismo na forma de trifosfato de adenosina, carboidratos podem desempenhar papeis estruturais e metabólicos importantes, tais como compor a parede celular de bactérias, fungos e plantas (Fincher, 2009; Latgé, 2007) e participar na mediação do reconhecimento intercelular (Misevic & Burger, 1990a; Misevic & Burger, 1990b). A ampla variedade de papeis exercidos por carboidratos na natureza, alguns citados acima, parece estar associada a sua grande variedade estrutural e conformacional.

Ao contrário de ácidos nucleicos e proteínas, que são lineares, apresentam um único tipo de ligação e pouca variedade de unidades componentes, carboidratos podem ser ramificados, ligados a partir de dois estados anoméricos distintos (Figura 6A e 6B), que se diferenciam pela posição axial ou equatorial de seu substituinte na posição C1, e através de átomos diferentes (Petrescu *et al.*, 2006). Adicionalmente, estima-se que exista mais de 100 monossacarídeos na natureza (embora aproximadamente 10 em árvores sacarídicas de glicoproteínas), o que amplia ainda mais sua variedade estrutural (Imberty & Pérez, 2000). Não obstante, alguns monossacarídeos podem apresentar uma flexibilidade no arranjo espacial dos

átomos componentes do seu anel interno. A conformação desta família de biomoléculas é usualmente descrita de acordo com as recomendações da IUPAC (IUPAC-IUB, 1980; IUPAC-IUB, 1983). Segundo essas regras, um monossacarídeo piranosídico, ou seja, formado por um anel de seis átomos, é determinado por uma letra maiúscula, em itálico, designando a forma do anel, e por números, responsáveis pela distinção das possíveis variantes daquela conformação. As formas mais comuns, encontradas nos monossacarídeos componentes de glicoproteínas, denominam-se cadeiras $^{4}C_{1}$ e $^{1}C_{4}$ (Figura 6C e 6D), nos quais o formato do anel se assemelha à de uma cadeira, de fato, com os carbonos das posições C4 e C1 acima e abaixo do plano na conformação $^{4}C_{1}$, e em organização inversa na forma $^{1}C_{4}$.

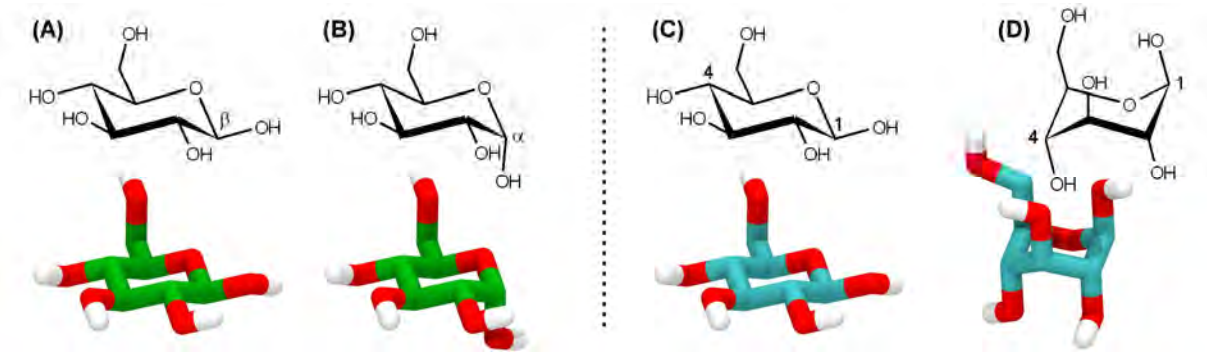


Figura 6: Variantes conformacionais e configuracionais associados a monossacarídeos, utilizando D-Glc como modelo: estados anoméricos α (A) ou β (B) e a conformação do anel piranosídico $^{4}C_{1}$ (C) e $^{1}C_{4}$ (D).

No contexto da coagulação sanguínea, a grande maioria dos aneis monossacarídicos segue o perfil conformacional descrito acima, à exceção do resíduo de IdoA em sequências específicas de heparina, como será discutido a seguir. Em termos funcionais, destaca-se o papel de carboidratos através da ação de GAGs (Choay *et al.*, 1983; Danielsson *et al.*, 1986) e da glicosilação de proteínas (Hansson & Stenflo, 2005).

1.3.1 GAGs

Os GAGs são uma classe de polissacarídeos lineares, na maioria das vezes sulfatados, envolvidos em uma série de processos fisiológicos tais como adesão, crescimento, sinalização e diferenciação celular (Jackson *et al.*, 1991), incluindo desenvolvimento neuronal, facilitação de oligomerização e prevenção à proteólise

(Handel *et al*., 2005). Esses carboidratos estão presentes na matriz extracelular de praticamente todas as células animais (Imberty *et al*., 2007; Gandhi & Mancera, 2008), majoritariamente na forma de proteoglicanas (Yung & Chan, 2006). Nessas moléculas, uma proteína está covalentemente ligada a cadeias lineares de GAG através de um resíduo de Ser (ou, em poucos casos, Asn) e um trissacarídeo composto por um resíduo de xilose (Xyl) e dois resíduos de galactose (Gal) (Gandhi & Mancera, 2008).

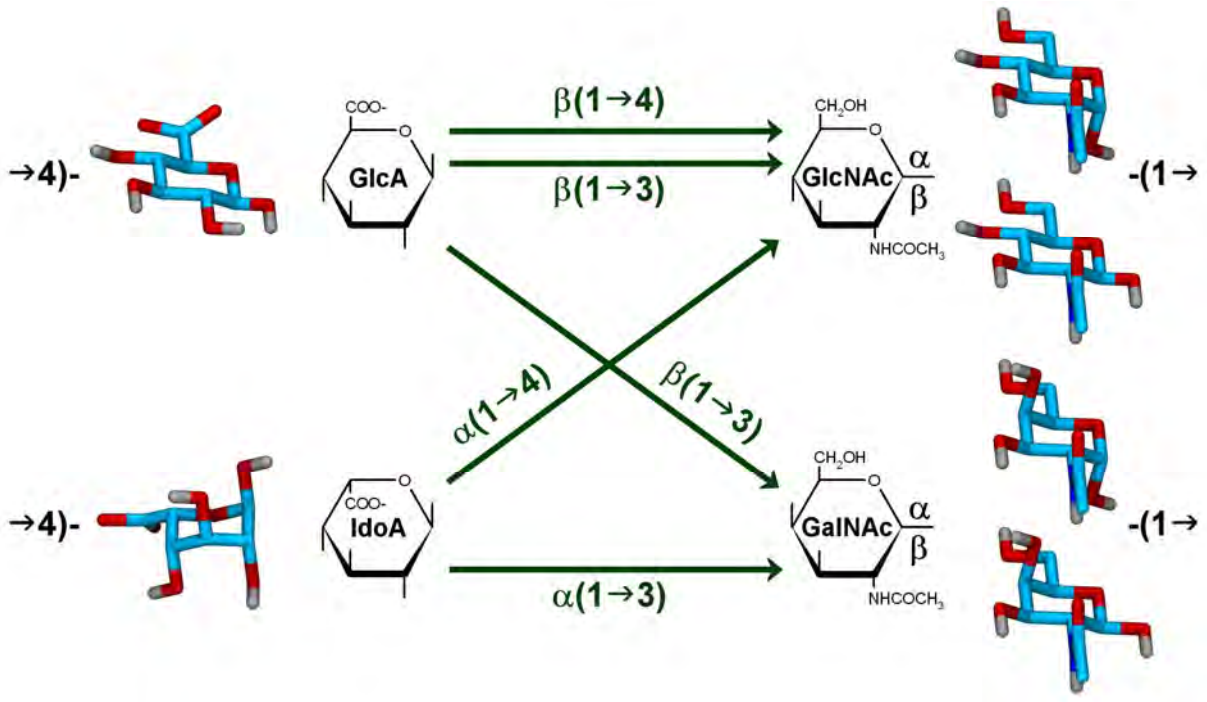


Figura 7: Representação esquemática da estrutura geral de GAGs. Baseado e adaptado de Sampaio *et al*., 2006.

Existem vários tipos de GAGs, dentre os quais destacam-se ácido hialurônico, condroitin sulfato, dermatan sulfato e heparina/heparan sulfato (Tabela 2) (Gandhi & Mancera, 2008). Essas moléculas são compostas por unidades dissacarídicas que se repetem ao longo da cadeia linear, cujos monossacarídeos componentes são uma hexosamina (N-acetil-glicosamina - GlcNAc, ou N-acetil-galactosamina, GalNAc) e um ácido urônico (ácido idurônico - IdoA, ou ácido glicurônico – GlcA) (Gandhi & Mancera, 2008). A união entre esses resíduos também varia, no que diz respeito à geometria dessa ligação, podendo ser dos tipos α ou β, e no que se refere aos átomos envolvidos, na forma das ligações 1→3 ou 1→4 (Figura 7). Existem outros tipos de GAGs, que diferenciam-se aos supracitados por apresentarem

características específicas: queratan sulfato, por exemplo, possui resíduos de Gal ao invés de um resíduo urônico (Gandhi & Mancera, 2008). Há, também, heparanosana e condrosana, que são formas não-sulfatadas de heparina e condroitin sulfato, respectivamente, e acharan sulfato, uma forma de GAG constituído por IdoA sulfatado e GlcNAc não-sulfatada (Sampaio *et al.*, 2006).

A classificação de GAGs pode ser feita de diferentes formas, sendo que a mais simples os divide entre não-sulfatados (ácido hialurônico) e sulfatados[1] (os demais). Nesses últimos, o padrão de sulfatação pode variar, ou seja, a quantidade de grupamentos sulfato e as posições onde eles estão localizados, de forma que um octassacarídeo pode conter mais de 1.000.000 de padrões de sulfatação diferentes (Sasisekharan & Venkataraman, 2000). A localização do grupamento sulfato em cada unidade dissacarídica é dada pelo átomo do anel do monossacarídeo envolvido na ligação e pela letra "S". Assim, por exemplo, um resíduo de IdoA sulfatado na posição do carbono 2 recebe o nome de IdoA2S. Outra forma de classificação se dá a partir da hexosamina componente do GAG, entre galactosaminoglicanos (condroitin sulfato e dermatan sulfato) e glicosaminoglicanos (ácido hialurônico e heparina/heparan sulfato).

Tabela 2: Composição dos diferentes tipos de GAGs.

GAG	Unidades dissacarídicas*	Sulfatação mais comum
Ácido hialurônico	→4)-β-GlcA-(1→4)-α-GlcNAc-(1→	—
Condroitin sulfato	→4)-β-GlcA-(1→3)-β-GalNAc-(1→	GalNAc4S
Dermatan sulfato	→4)-α-IdoA-(1→3)-β-GalNAc-(1→	GalNAc4S
Heparina / Heparan sulfato†	→4)-β-GlcA-(1→4)-α-GlcNAc-(1→ →4)-α-IdoA-(1→4)-α-GlcNAc-(1→	IdoA2S e GlcN6S
Queratan sulfato	→3)-β-Gal-(1→4)-β-GalNAc-(1→	GalNAc6S
Acharan sulfato	→4)-α-IdoA-(1→4)-α-GlcNAc-(1→	IdoA2S

* Dados adaptados de Gandhi & Mancera, 2008 e Sampaio *et al.*, 2006;

† Em heparina, resíduos de IdoA predominam; em heparan sulfato, GlcA prevalece.

[1] GAGs sulfatados não ocorrem em espécies do reino protista, plantas ou fungos. No reino animal, o aparecimento desse tipo de biomolécula (heparan sulfato e condroitin sulfato) ocorre apenas em espécies que possuem organização celular na forma de tecidos. A ocorrência de dermatan sulfato é um evento tardio na evolução, enquanto heparina ocorre de forma dispersa ao longo do reino animal (Sampaio *et al.*, 2006).

A primeira função terapêutica determinada para GAGs sulfatados está relacionada a sua ação anticoagulante, identificada através de heparina (Nader *et al.*, 2001). Fisiologicamente, esse polissacarídeo é encontrado em vesículas de armazenamento de mastócitos, principalmente em órgãos que possuem contato direto com o ambiente externo, tais como nos pulmões, no intestino e na pele (Sampaio *et al.*, 2006; Gandhi & Mancera, 2008), com função crucial em reações alérgicas e inflamatórias guiadas por esse tipo celular (Oschatz *et al.*, 2011). A heparina apresenta, em média, cerca de 2 grupamentos sulfato por unidade dissacarídica, sendo esse o maior conteúdo de sulfatação observado para esse tipo de molécula, mas que pode variar dependendo do tecido onde foi sintetizado e do seu organismo de origem (Nader *et al.*, 2001). Esse polissacarídeo vem sendo utilizado clinicamente há mais de 40 anos como agente anticoagulante e antitrombótico (Jacques, 1979), no tratamento de trombose, embolia e tromboflebite, atuando, como comentado anteriormente, através da potenciação da atividade de AT sobre proteases da cascata de coagulação (Choay *et al.*, 1983; Danielsson *et al.*, 1986).

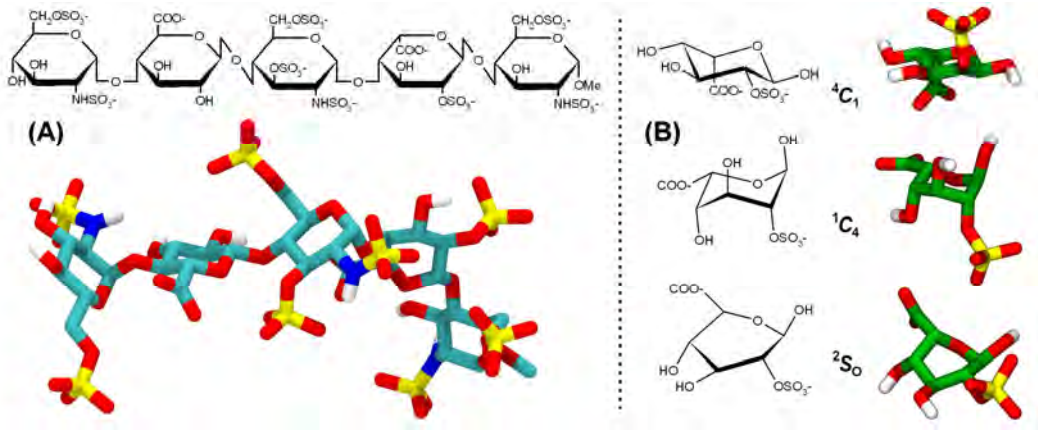


Figura 8: Em (A), o pentassacarídeo necessário para que a heparina tenha atividade potencializadora sobre AT é apresentado. Em (B), uma representação dos estados conformacionais adotados pelo resíduo IdoA2S é mostrada.

Estruturalmente, a heparina é um polissacarídeo de alto peso molecular composto por resíduos de GlcA não sulfatado, IdoA majoritariamente sulfatado na posição do C2 e glicosamina (GlcN) com um padrão de sulfatação variável, desde não sulfatado até contendo três grupamentos sulfato (GlcNS,3,6S) (Silva & Dietrich. 1975). Para que a heparina potencialize rapidamente a atividade de AT, faz-se

necessária, independentemente do mecanismo de ação e do tamanho do polissacarídeo, que ocorra a interação de um pentassacarídeo mínimo (Figura 8A) com a estrutura da AT. Esse processo parece ser influenciado pela flexibilidade do resíduo IdoA2S, que, dependendo do ambiente em que estiver, apresenta um equilíbrio conformacional não usual (Figura 8B) entre as formas de cadeira $^{4}C_{1}$ (que ocorre majoritariamente quando livre em solução), cadeira $^{1}C_{4}$ e a *skew-boat* $^{2}S_{O}$ (Ferro *et al.*, 1986).

1.3.2 Glicosilação

A glicosilação é uma das modificações co- e/ou pós-traducionais de proteínas mais frequentes na natureza, ocorrendo em praticamente todos os organismos, desde eubactérias até eucariotos (Spiro, 2002). Até o momento, encontram-se descritos pelo menos seis tipos de ligações entre carboidratos e proteínas, envolvendo treze monossacarídeos e nove aminoácidos diferentes (Spiro, 2002; Stepper *et al.*, 2011): C-glicosilação, fosfoglicosilação, glipiação, S-glicosilação, O-glicosilação e N-glicosilação. Dentre essas, a N- e a O-glicosilação são os tipos de glicosilação mais comumente encontrados na natureza, e os únicos descritos em proteínas da coagulação sanguínea até hoje (Tabela 3).

Tabela 3: Glicosilação em proteínas envolvidas na coagulação sanguínea.

Proteína*	Número de sítios†	
	N-glicosilação	O-glicosilação
Protrombina (Fator II)	3	0
Fator V	37	0
Fator VII	2	2
Fator VIII	24	0
Fator IX	2	6
Fator X	2	3
Fator *von Willebrand*	12	10
Proteína C	4	0
Proteína S	3	0
AT	4	0

* Dados adaptados de Hansson & Stenflo, 2005.

† Informações obtidas a partir do banco de dados online do SWISS-PROT.

A O-ligação glicosídica é mais diversa e versátil que a N-glicosilação, considerando-se a elevada quantidade de aminoácidos e monossacarídeos envolvidos (Spiro, 2002). No entanto, em proteínas da coagulação, apenas resíduos de Ser e Thr foram relacionados a O-glicosilação e resíduos de Asn à N-glicosilação (Hansson & Stenflo, 2005). Até o momento, poucas sequências consenso conservadas foram identificadas para O-glicosilação (Spiro, 2002), embora algumas delas envolvam os sítios de O-glicosilação de fXa (Wilson *et al.*, 1991) e domínios EGF (do inglês *epidermal growth factor*), que ocorrem frequentemente em proteínas da coagulação (Hansson & Stenflo, 2005). Adicionalmente, o desenvolvimento de bases de dados e algoritmos de predição relacionados a sítios de O-glicosilação (Gupta *et al.*, 1999; Li *et al.*, 2006) ampliaram a visão sobre esse assunto, possibilitando a identificação de algumas propriedades inerentes a essas regiões, tais como a preferência por resíduos de Pro e a aversão por aminoácidos aromáticos nas proximidades do sítio de glicosilação (Thanka Christlet & Veluraja, 2001). A síntese de O-ligações glicosídicas é atribuída a várias enzimas diferentes, de forma que a ligação proteína-carboidrato ocorre através da transferência de monossacarídeos isolados – GalNAc, Xyl, manose (Man), fucose (Fuc) – pós-traducionalmente, ou seja, a polipeptídeos já enovelados (Imperiali & Hendrickson, 1995). Até hoje, vários tipos de O-ligações glicosídicas foram descritas, e a conformação de algumas delas, bem como os efeitos da O-glicosilação sobre a estrutura protéica em estudo, já foram avaliados por técnicas computacionais (Tabela 4). Adicionalmente, embora os resíduos de aminoácido e monossacardeo envolvidos já tenham sido determinados, o estado anomérico (α/β) de algumas outras conexões ainda não está identificado (Spiro, 2002), o que se constitui em uma ampla área de trabalhos a serem explorados, especialmente envolvendo furanoses e glicoproteínas de planta.

Por outro lado, a N-glicosilação possui uma sequência 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 Pro (Gavel & von Heijne, 1990; Marshall, 1974). As propriedades dessa sequência consenso e dos seus aminoácidos vizinhos são, assim como para a O-glicosilação, bem compreendidas, em que resíduos de Pro impedem a N-glicosilação (Gavel & von Heijne, 1990; Marshall, 1974), tal como supracitado, e a presença de resíduos aromáticos a

favorecem (Thanka Christlet & Veluraja, 2001). A síntese e modificação de oligossacarídeos durante a N-glicosilação são amplamente estudadas e compreendidas, nas quais a ação de determinadas enzimas é conhecida e estabelecida para etapas específicas do processo (Helenius & Aebi, 2001).

Tabela 4: O-ligações glicosídicas avaliadas por métodos computacionais

O-ligação glicosídica	Contexto do estudo	Referência
α-GalNAc-(1→O)-Ser	Glicopeptídeo	Huang *et al.*, 1997
α-GalNAc-(1→O)-Thr	Glicopeptídeo	Huang *et al.*, 1997
β-GlcNAc-(1→O)-Ser	Glicopeptídeo	Huang *et al.*, 1997
β-GlcNAc-(1→O)-Thr	Glicopeptídeo	Huang *et al.*, 1997
α-GalNAc-(1→O)-Thr	Glicopeptídeo	Nguyen *et al.*, 2002
α-GalNAc-(1→O)-Ser	AA-monossacarídeo	Corzana *et al.*, 2006a
β-Glc-(1→O)-Ser	AA-monossacarídeo	Corzana *et al.*, 2006b
β-Glc-(1→O)-Thr	AA-monossacarídeo	Corzana *et al.*, 2006b
β-GalNAc-(1→O)-Ser	AA-monossacarídeo	Corzana *et al.*, 2007
α-GalNAc-(1→O)-Thr	AA-monossacarídeo	Corzana *et al.*, 2007
β-GalNAc-(1→O)-Thr	AA-monossacarídeo	Corzana *et al.*, 2007
β-GlcNAc-(1→O)-Ser	AA-monossacarídeo	Fernández-Tejada *et al.*, 2009
β-GlcNAc-(1→O)-Thr	AA-monossacarídeo	Fernández-Tejada *et al.*, 2009
α-Fuc-(1→O)-Ser	Glicoproteína	Pol-Fachin *et al.*, 2009
α-GalNAc-(1→O)-Thr	Glicopeptídeo	Corzana *et al.*, 2010
α-GalNAc-(1→O)-MeSer	Glicopeptídeo	Corzana *et al.*, 2011
α-Gal-(1→O)-HyPro	Glicopeptídeo	Seção 9, trabalho VI
β-Ara-(1→O)-HyPro	Glicopeptídeo	Seção 9, trabalho VI
β-Ara-(1→O)-HyPro	Glicoproteína	Seção 9, trabalho III
β-Gal-(1→O)-HyPro	Glicoproteína	Seção 9, trabalho III
α-Gal-(1→O)-HyPro	AA-monossacarídeo	Naziga *et al.*, 2012
β-Gal-(1→O)-HyPro	AA-monossacarídeo	Naziga *et al.*, 2012
α-Fuc-(1→O)-Thr	Glicopeptídeo	Kaushik *et al.*, 2012

Essa cascata de eventos começa na face citosólica da membrana do retículo endoplasmático (RE), onde alguns dos monossacarídeos constituintes de um tetradecassacarídeo inicial são ligados a uma molécula de dolicol fosfato (Figura

9A). Quando sete resíduos (duas GlcNAc e cinco Man) estão ligados, este heptassacarídeo é revertido ao lúmen do RE (Figura 9B) e mais sete monossacarídeos são adicionados (Figura 9C). Em seguida, esse oligossacarídeo é transferido para o polipeptídeo nascente (Figura 9D) e pode sofrer ciclos de adição e remoção de resíduos de glicose (Figura 9E-F). Quando a proteína estabelece seu correto enovelamento (Figura 9G), ela é transferida para o complexo de Golgi, onde se inicia o processamento desse oligossacarídeo (Figura 9H), e onde tem origem a imensa diversidade observada nos glicoconjugados que atingem as superfícies celulares (Helenius & Aebi, 2001). No caso de oligossacarídeos ligados às proteínas estudadas na presente Tese, resíduos de Man são removidos da estrutura glicosídica (Figura 9H) e, após, tem início a adição de outros resíduos, tais como GlcNAc, Gal e ácido neuramínico (NeuAc), também conhecido como ácido siálico (Figura 9J-K-L). Alternativamente, algumas glicoproteínas conseguem evadir a transferência para o Golgi (Helenius & Aebi, 2001), e permanecem com um oligossacarídeo rico em manoses (Figura 9G).

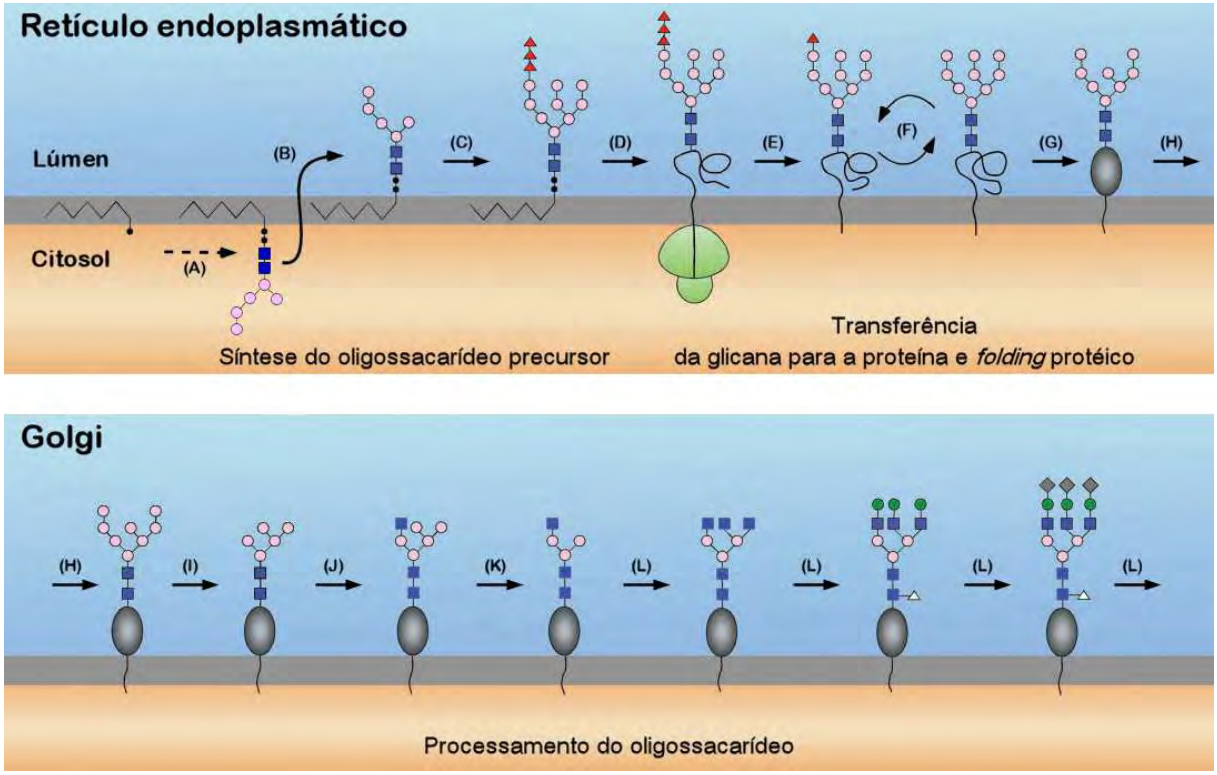


Figura 9: Processo de biosíntese de um oligossacarídeo N-ligado do tipo complexo. As formas azuis representam GlcNAc; as rosas, Man; as vermelhas, Glc; as verdes, Gal; as cinzas, NeuAc; e as brancas, Fuc. Adaptado de Helenius & Aebi, 2001.

Especificamente sobre as enzimas envolvidas na formação dos complexos ternários, objetos principais de estudo desta tese, fator X (ou seja, o zimogênio da enzima ativa fXa), possui dois sítios de N-glicosilação e três sítios de O-glicosilação (Tabela 3) (Hansson & Stenflo, 2005). Após seu processamento, gera-se fXaα, que apresenta um sítio de O-glicosilação ocupado na posição Thr258 (Hansson & Stenflo, 2005). No entanto, um processo de auto-clivagem remove parte da alça C-terminal onde esta glicana está localizada, gerando a forma final da enzima, fXaβ (Mertens & Bertina, 1980) que, portanto, não é glicosilada e não apresenta diferença de atividade em relação à fXaα (Pryzdial & Kessler, 1996). Protrombina (o zimogênio de trombina ativa) possui três de seus quatro sítios de N-glicosilação ocupados por oligossacarídeos (Tabela 3) (Hansson & Stenflo, 2005). Destes, um único sítio permanece presente na trombina ativa, localizado no resíduo de Asn60G, cuja remoção diminui a atividade catalítica da enzima sobre fibrinogênio para cerca de 90% daquela de proteína glicosilada (Rosenfeld & Danishefsky, 1984).

AT, por sua vez, apresenta quatro sítios de N-glicosilação ao longo de sua sequência de aminoácidos, três dos quais estão sempre ocupados, localizados nas posições Asn96, Asn155 e Asn192. O quarto sítio, localizado na Asn135, é o único inserido em um tripeptídeo de sequência Asn-X-Ser, ao contrário dos demais, presentes em sítios contendo sequência consenso Asn-X-Thr. Essa diferença faz com que uma porção minoritária da AT presente no plasma humano, que representa de 5-10% da AT total circulante, não apresente oligossacarídeos ligados à Asn135 (Picard *et al.*, 1995). À AT apresentando apenas três sítios de N-glicosilação ocupados, e menor peso molecular, dá-se o nome de β-AT, enquanto que àquela contendo quatro oligossacarídeos ligados dá-se o nome de α-AT (Peterson & Blackburn, 1985; Turko *et al.*, 1993). Essa diferença no conteúdo de glicosilação de α- e β-AT é relevante em dois sentidos: (1) na afinidade dessas glicoformas por heparina, uma vez que β-AT apresenta uma afinidade por heparina maior que α-AT em cerca de uma escala de grandeza (Turko *et al.*, 1993) e (2) na potencialização da atividade inibitória dessas glicoformas sobre as serino proteases alvo, uma vez que a presença de um oligossacarídeo ligado à Asn135 desestabilizaria o estado de AT ativada (Turk *et al.*, 1997).

1.4 Caracterização estrutural de biomoléculas

A fim de obter um mais amplo entendimento de qualquer evento biológico, tais como o processo de coagulação sanguínea, é essencial a caracterização da estrutura e conformação das biomoléculas envolvidas, seja isoladas ou formando os complexos moleculares observados fisiologicamente. Os resultados obtidos podem ser relacionados a propriedades que vão desde o nível atômico, como a importância da glicosilação para a estrutura e função proteica, até o nível celular, como a transdução de sinal mediada por interações proteína-proteína e proteína-ligantes. Por um lado, algumas técnicas geram informações acerca apenas da estrutura e forma global da molécula em estudo[2]. Por outro lado, existem metodologias capazes de oferecer dados em nível atômico a respeito da molécula em estudo, já tendo sido aplicadas a proteínas, glicoproteínas e polissacarídeos envolvidos na coagulação, dentre as quais destacam-se a cristalografia de raios-X e a ressonância magnética nuclear (RMN) em solução.

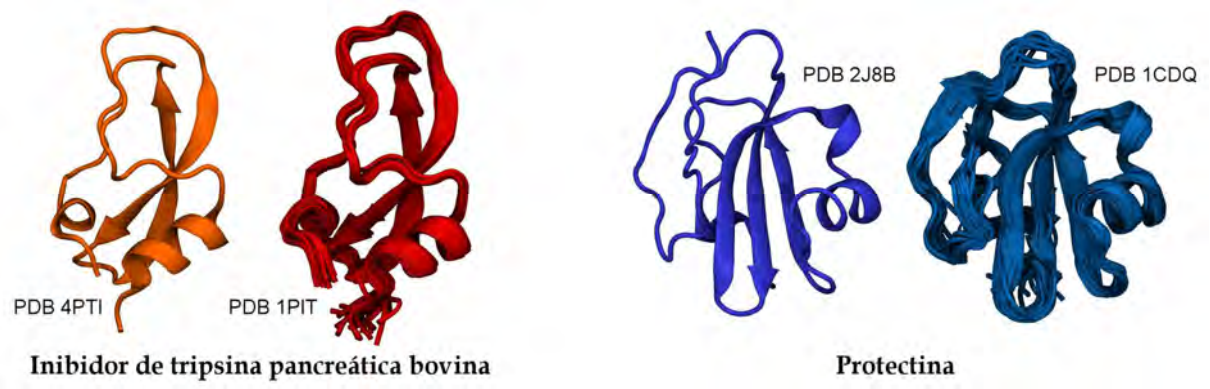


Figura 10: Estrutura 3D de duas proteínas, inibidor de tripsina pancreática bovina e protectina, obtidas por cristalografia de raios-X e pelo método de RMN.

A cristalografia de raios-X é a técnica que oferece a mais ampla descrição de uma molécula em nível atômico (Petrescu *et al.*, 2006). Sua aplicação na caracterização da estrutura e conformação de proteínas é bastante frequente, de

[2] Por exemplo, o dicroismo circular é um método capaz de avaliar o conteúdo de estrutura secundária e enovelamento proteico, sendo empregado para determinar, sem resolução atomística, se uma mutação altera a conformação e estabilidade de uma proteína (Greenfield, 2006). Da mesma forma, o *small-angle X-ray scattering* é uma técnica que utiliza espalhamento de raios-X a baixo ângulo na análise estrutural de proteínas, disponibilizando dados acerca da forma e tamanho da proteína ou complexo proteico em estudo (Fischer *et al.*, 2010), também sem resolução atomística.

forma que no maior banco de dados de estruturas de proteínas disponíveis na *web*, o Protein Data Bank (PDB), o número de estruturas depositadas contendo proteínas, estudadas por cristalografia, se aproxima de 80 mil (acessado em 24 de abril de 2013). O método de RMN, por sua vez, oferece um conjunto de restrições espaciais geralmente envolvendo átomos de hidrogênio, através de técnicas como NOESY, que representam uma estrutura média para a molécula em estudo. Essas restrições, denominadas sinais de NOE (do inglês nuclear Overhauser effect), são medidas em uma escala de tempo que varia entre 50 ms e 1 s (Woods, 1998; Wormald *et al.*, 2002), o que não afeta a determinação conformacional de moléculas mais rígidas, como proteínas enoveladas. Para proteínas, o método de RMN é bastante empregado, assim como a cristalografia de raios-X, com cerca de 9 mil estruturas depositadas no PDB (acessado em 24 de abril de 2013). No entanto, devido à dificuldade na interpretação dos mapas de contatos derivados de RMN com o aumento do tamanho da proteína em estudo, essas técnicas são geralmente empregadas a proteínas de até ~250 aminoácidos (30-40 kDa). De uma forma geral, o estudo de dada proteína por cristalografia de raios-X possibilita a obtenção de um único modelo para sua estrutura 3D, enquanto que, quando avaliada pelo método de RMN, uma série de modelos podem ser obtidos, pela diferente interpretação dos sinais de NOE derivados do experimento (Figura 10). Ainda assim, a flexibilidade da molécula em estudo pode ser avaliada em ambas as técnicas, isto é, através dos dados de B-*factor*, obtidos pela cristalografia de raios-X, cujos valores podem ser utilizados para indicar a mobilidade de átomos ou aminoácidos, e pela flexibilidade observada pela sobreposição dos diferentes modelos derivados dos experimentos de RMN.

No estudo de polissacarídeos[3], por outro lado, a flexibilidade conformacional dessas moléculas, estejam elas livres ou ligadas à superfície da proteína, covalentemente ou não, geralmente dificulta a formação de cristais (Bohne-Lang & von der Lieth, 2002; Petrescu *et al.*, 2006), assim como a elevada coordenação de

[3] Não são incomuns erros na determinação do estado anomérico, na anotação de conformações distorcidas de monossacarídeos (Lütteke *et al.*, 2004), ou na orientação relativa de dois resíduos unidos através de uma ligação glicosídica (Lütteke, 2009). Na verdade, avaliações ao banco de dados do PDB estimam que cerca de um terço das estruturas de carboidratos apresentam algum tipo de inconsistência (Lütteke *et al.*, 2004; Crispin *et al.*, 2007). Um exemplo que ocorre frequentemente em glicoproteínas envolve o resíduo de 2-(acetilamino)-2-deoxi-glicopiranose, ou N-acetil-glicosamina (GlcNAc). Nesse sentido, à época, foram encontrados aproximadamente duzentos casos em que foi incorretamente assinalado o resíduo α-GlcNAc, ao invés do correto, β-GlcNAc (Berman et al., 2007).

moléculas de água à carboidratos e a falta de interações lipofílicas ou dipolares entre seus monossacarídeos componentes (Woods, 1998). Adicionalmente, no caso de glicoproteínas, quando os cristais são obtidos, o mapa de densidade eletrônica é prejudicado pela alta mobilidade dessas glicanas, compromentendo sua correta determinação conformacional e configuracional (Petrescu *et al.*, 1999). No caso da técnica de RMN, uma determinação conformacional precisa utilizando essa metodologia só é possível se forem obtidos três ou mais sinais de NOE interresíduos em uma ligação glicosídica (Wormald *et al.*, 2002) ou entre proteína e carboidrato. Novamente, devido a elevada flexibilidade conformacional dessas moléculas, que usualmente coexistem em dois ou mais subestados conformacionais, a obtenção de sinais de NOE é prejudicada, uma vez que essas restrições podem representar propriedades conformacionais médias dessas moléculas (Woods, 1998; Wormald *et al.*, 2002), o que não necessariamente corresponde aos subestados conformacionais existentes em solução.

As proteínas, glicoproteínas e polissacarídeos de interesse para a presente Tese foram amplamente estudados pelas técnicas experimentais supracitadas. Especificamente, há cerca de 30 estruturas de AT, 200 de trombina e 80 de fXa depositadas no PDB, bem como duas estruturas de heparina derivadas de estudos de RMN (Mulloy *et al.*, 1993) e dados conformacionais de cristalografia disponíveis na literatura para esse polissacarídeo (Khan *et al.*, 2010). No entanto, a árvore de glicosilação completa de AT e trombina não puderam ser completamente determinadas a partir dos mapas de densidade eletrônica obtidos até então. Da mesma forma, embora grande parte dos subestados conformacionais adotados pela estrutura proteica de AT, trombina e fXa sejam conhecidos, a dinâmica de transição entre esses estados, a flexibilidade dos diferentes elementos componentes dessas moléculas (como elementos de estrutura secundária), a influência da glicosilação e de ligantes sobre essas estruturas ainda não é bem conhecida. Como uma metodologia alternativa, métodos computacionais e, em especial, dinâmica molecular (DM), já foram empregadas no estudo da interação AT-heparina (Verli & Guimarães, 2005) e de outras glicoproteínas, reproduzindo sua conformação e plasticidade (Pol-Fachin *et al.*, 2009), a partir de um protocolo desenvolvido durante meu Mestrado (Pol-Fachin, 2009).

1.5 Estudo de biomoléculas por simulações computacionais

O uso de simulações de DM, que consistem na avaliação do movimento dos átomos em uma molécula através de algoritmos computacionais, iniciou-se há mais de 35 anos para sistemas biológicos, com o estudo de um sistema proteico contendo aproximadamente 900 átomos, por um tempo de 8,8 picosegundos, envolvendo um inibidor de tripsina pancreática bovina (McCammon *et al.*, 1977). Atualmente, consiste em uma ferramenta amplamente difundida e empregada no estudo da estrutura, conformação, função e dinâmica de biomoléculas em geral, em sistemas que ultrapassam 2,5 milhões de átomos (Sanbonmatsu & Tung, 2006) por tempos de simulação na escala de milisegundos (Shaw *et al.*, 2010).

A movimentação dos átomos em uma molécula, tal como descrito acima, é obtida por DM através da sucessiva integração da equação de movimento de Newton ($d^2r_i(t)/dt^2 = F_i / m_i$) sobre todos os átomos do sistema em estudo (Leach, 2001). Essa integração, que é a característica fundamental dos cálculos de DM, é realizada de forma que uma força F_i acarreta uma aceleração $d^2r_i(t)/dt^2$ sobre um determinado átomo *i* de massa m_i e, em consequência, causa uma mudança de sua posição no espaço cartesiano num intervalo de tempo Δt (Leach, 2001). A sucessão desses intervalos de tempos viabiliza uma sequência de diferentes posições dos átomos em função do tempo, ao que denomina-se *trajetória*. A partir da equação de Newton, no entanto, não é possível determinar a intensidade e a direção da força F_i, ao passo que uma segunda equação ($F_i = -\delta V(r_i,...r_n)/\delta r_i$) faz parte dos cálculos de DM, colocando a força F_i como uma função das coordenadas cartesianas dos átomos ($r_i,...r_n$) e da energia potencial do sistema V, que representa a energia de cada molécula componente do sistema (Leach, 2001).

Afora essas equações, outro conjunto de funções e parâmetros se somam para determinar a trajetória adotada pelos átomos do sistema e a energia potencial de cada molécula, ao que se denomina *campo de força*. Diversos campos de força estão disponíveis para simulações de DM, dentre os quais destacam-se, por seu amplo uso na literatura, AMBER (Case *et al.*, 2005), CHARMM (MacKerell *et al.*, 1998), OPLS (Jorgensen *et al.*, 1988) e GROMOS (Scott *et al.*, 1999). Há um consenso na forma funcional básica desses campos de força, ou seja, nos termos de energia potencial que descrevem os sistemas em estudo. Esses termos definem tanto interações covalentes entre os átomos do sistema quanto interações não-

covalentes, entre átomos de moléculas diferentes ou entre átomos separados entre si por 2-3 ligações covalentes (van Gunsteren *et al.*, 2006). Os termos que definem a distância entre um dado par de átomos ligados e o ângulo formado entre três átomos seguem a forma funcional descrita na Figura 11, onde b é distância e θ é o ângulo do momento que está sendo calculado, b_o é a distância e θ_o é o ângulo de referência, e K_b ou K_θ são as constantes de força empregadas para manter esses valores em um equilíbrio ao redor dos valores de referência. Para diedros próprios, uma função cosseno é empregada (Figura 11), onde χ se refere ao valor do diedro e K_χ ao valor que determina a altura das barreiras de energia. Uma vez que a rotação de um diedro é considerada periódica, e pode apresentar múltiplos mínimos de energia, não é estabelecido um único valor de referência. Assim, o perfil rotacional dos diedros é determinado pelos parâmetros n, que se refere à periodicidade ou multiplicidade, e reflete o número de mínimos de energia, e δ, que diz respeito à mudança de fase, e à localização do máximo de energia ao longo do perfil da rotação do diedro[4].

Todos os campos de força citados acima são capazes de lidar com proteínas, ou seja, levar uma estrutura conhecida a uma evolução temporal, e descrevem sua estrutura, conformação e dinâmica de forma semelhante (Price & Brooks III, 2002). No caso de lipídeos, a maior parte dos estudos envolve os campos de força CHARMM e GROMOS, porém o último oferece um ganho computacional que pode chegar a nove vezes (Oostenbrink *et al.*, 2004) devido à sua natureza *united-atom*, ou seja, por descreverem os átomos de hidrogênio apolares unidos aos seus respectivos carbonos. Para ácidos nucleicos, o OPLS não possui parâmetros suficientes para descrever a ligação fosfodiéster e as porções sacarídicas e, embora também haja parâmetros GROMOS disponíveis, há uma mais ampla utilização e reprodução por AMBER e CHARMM de dados conformacionais experimentais para DNA e RNA (Ricci *et al.*, 2010). A parametrização de carboidratos, por sua vez, está

[4] Para diedros impróprios, que visam manter a quiralidade ou planaridade de um grupamento ao redor de um átomo com ligação covalente a outros três, duas equações são empregadas. CHARMM e GROMOS aplicam uma equação que se assemelha àquela empregada para distâncias e ângulos, o que é particularmente importante para GROMOS, uma vez que, por não descrever explicitamente hidrogênios apolares, é necessária uma equação mais cuidadosa para preservar a correta orientação de grupamentos ao redor de carbonos quirais. Por outro lado, os campos de força AMBER e OPLS aplicam um termo de função cosseno, tal como aquele empregado para diedros próprios, mas somente para grupamentos planares, tais como o grupamento NH de ligações peptídicas.

imersa em desafios devido a sua elevada complexidade estrutural e conformacional, de forma que uma sucessão de novos parâmetros tem surgido nos últimos anos, cuja utilização na literatura é também bastante variada.

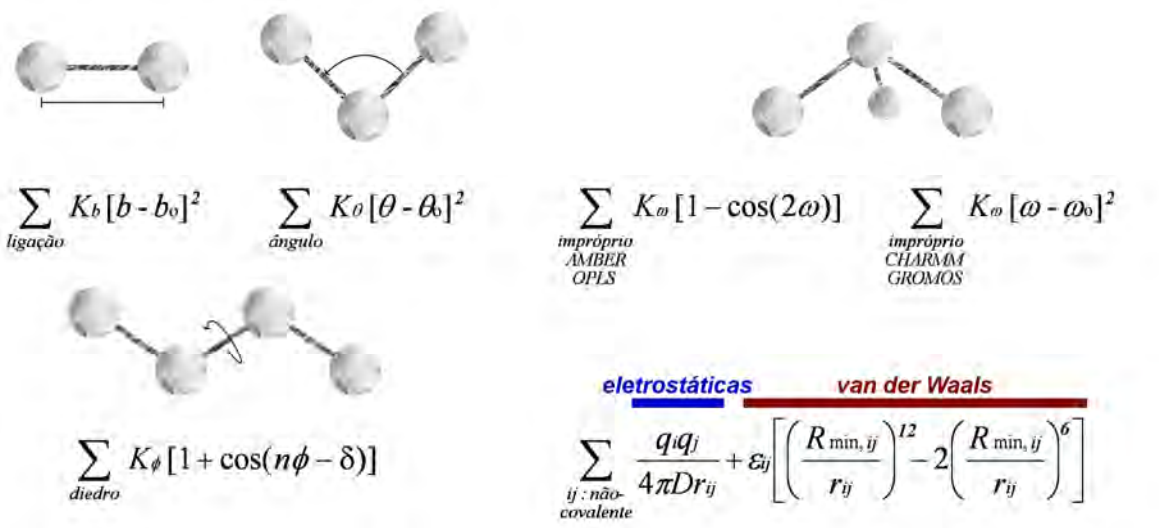


Figura 11: Termos de energia que compõe campos de força. Estão representadas as equações que descrevem o estiramento de ligações químicas, ângulos de ligação, diedros próprios, diedros impróprios e interações não-covalentes, considerando juntamente as interações eletrostáticas, ou Coulômbicas, e as interações de *van der Waals*, também identificadas como potencial de *Lennard-Jones*.

Nesse sentido, o Grupo de Bioinformática Estrutural vem se dedicando ao uso e contínuo aprimoramento e validação do campo de força GROMOS 43A1 para carboidratos, que teve início com simulações por DM de compostos sulfatados (Becker *et al.*, 2007; Castro *et al.*, 2009) e, em especial, heparina (Verli & Guimarães, 2004; Verli & Guimarães, 2005; Pol-Fachin & Verli, 2008). Mais recentemente, esses parâmetros foram expandidos à glicoproteínas, de forma que sua capacidade em reproduzir adequadamente as propriedades conformacionais dessas biomoléculas (Pol-Fachin *et al.*, 2009) e de predizer a conformação de suas estruturas sacarídicas (Fernandes *et al.*, 2010) foi demonstrada. Destacadamente, e baseados nesses estudos, propôs-se que a conformação de carboidratos em geral pode ser obtida a partir dos estados de maior abundância em solução de seus dissacarídeos constituintes (Fernandes *et al.*, 2010; seção 9.1 desta Tese). Da mesma forma, o Grupo de Biomateriais vem se dedicando ao uso e desenvolvimento

do campo de força GROMOS 45A4 (Lins & Hünenberger, 2005) para carboidratos do tipo aldohexopiranose.

Enquanto apenas algumas propriedades de sistemas biomoleculares estão acessíveis a métodos experimentais, simulações computacionais lidam não apenas com dados médios, mas com grandezas espaço- e tempo-dependentes (van Gunsteren *et al.*, 2006). Nesse sentido, a utilização dos parâmetros supracitados, visando ao entendimento de problemas biológicos, tais como a inibição de trombina e fXa por AT e heparina, representa um complemento a essas metodologias, e podem ser empregadas na predição das propriedades não observáveis por técnicas experimentais. Adicionalmente, o contínuo aprimoramento de campos de força, tais como GROMOS 43A1 e 45A4, baseados em ferramentas disponíveis gratuitamente, consiste em uma promissora estratégia para a descrição e predição conformacional de sistemas biológicos de interesse, envolvendo carboidratos e glicoconjugados, através de simulações de DM.

2 Objetivos

A partir do exposto, o presente trabalho se insere no contexto da caracterização estrutural, conformacional e funcional de proteínas e, principalmente, glicoproteínas envolvidas na cascata de coagulação sanguínea. Deve-se considerar, também, a ausência de dados em nível atômico e tempo-dependentes acerca da interação de AT, trombina e fXa com heparina, bem como sobre a dinâmica dos complexos ternários formados por essas moléculas e a influência da glicosilação de AT e trombina nesse processo. Nesse sentido, o presente trabalho visa avaliar as bases moleculares da modulação da cascata de coagulação por heparina através de simulações por DM. Adicionalmente, tendo em vista a necessidade do contínuo aprimoramento dos protocolos de simulação e parâmetros empregados nesses estudos, o presente trabalho visa refinar os parâmetros empregados atualmente para carboidratos. Espera-se, a partir desses resultados, melhor compreender os mecanismos de ação da heparina a nível molecular, bem como subsidiar futuros aprimoramentos e expansões do campo de força GROMOS 45A4 para uma mais ampla diversidade de carboidratos. Dessa forma, as seguintes metas foram estabelecidas:

- Revisar as metodologias empregadas atualmente para a construção, avaliação e análise da conformação e dinâmica de glicoproteínas;
- Caracterizar o comportamento conformacional das duas glicoformas circulantes de AT, α e β, tanto livres quanto complexadas à heparina;
- Caracterizar o comportamento conformacional das proteases da cascata de coagulação trombina (glicosilada e não glicosilada) e fXa, tanto livres quanto complexadas à heparina;
- Aprimorar o conjunto de parâmetros para carboidratos do GROMOS 45A4.

3 Metodologia

3.1 Programas utilizados

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

Os programas utilizados incluem:

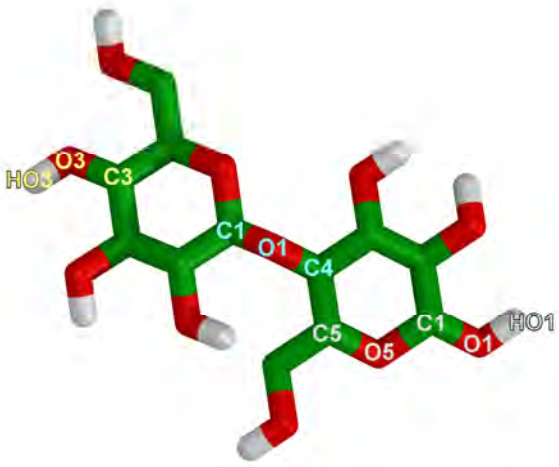
- Ferramentas de visualização de moléculas: VMD v1.9.1 (Humphrey *et al.*, 1996), PyMol v1.30 (DeLano, 2002) e MOLDEN (Schaftenaar & Noordik, 2000);
- Programa para cálculo *ab initio*: GAMESS (Schmidt *et al.*, 1993) e GAUSSIAN 98 (Frisch *et al.*, 2002);
- Programa para construção de mapas de contorno e simulações de dinâmica molecular: GROMACS, nas versões 3.3.3 (van der Spoel *et al.*, 2005), 4.5.1 e 4.5.4 (Hess *et al.*, 2009);
- Programa para simulações de dinâmica molecular por metadinâmica: GROMACS, versão 4.5.1 (Hess *et al.*, 2009) modificado pela plataforma PLUMED (Bonomi *et al.*, 2009);
- Programa para geração de topologias: PRODRG *Beta* v2.5 (Schuettelkopf & van Aalten, 2004);
- Programas para análise de estrutura secundária: PROCHECK (Laskowski *et al.*, 1993) e DSSP (Kabsch & Sander, 1983).

3.2 Cálculos utilizando métodos ab initio

A obtenção de cargas atômicas consiste em uma das primeiras etapas da parametrização de novas moléculas em um campo de força de mecânica molecular (MM). Nesse contexto, nosso grupo de pesquisas vem aplicando há cerca de dez anos, com sucesso, um procedimento baseado em cargas atômicas de Löwdin, obtidas por cálculos Hartree Fock, na descrição conformacional de carboidratos em geral, mas em especial de polissacarídeos sulfatados (Verli & Guimarães, 2004; Becker *et al.*, 2005; Becker, 2007; Pol-Fachin & Verli, 2008; Castro *et al.*, 2009). Por outro lado, as cargas atômicas mais aceitas e empregadas na parametrização de

compostos em geral são obtidas pelo método de potencial eletrostático restrito (do inglês *Restrained ElectroStatic Potential* – RESP).

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 e grupamentos em suas formas isoladas. Dessa forma, tais estruturas foram construídas utilizando o programa MOLDEN (Schaftenaar, 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, e HF/6-31G*, para RESP (Figura 12).



Átomo	Carga atômica	
	Löwdin	RESP
	Corpo do Anel	
C5	0.242	0.376
O5	-0.242	-0.480
C1	0.160	0.232
O1	-0.400	-0.538
HO1	0.240	0.410
	Ligação glicosídica	
C1	0.254	0.232
O1	-0.396	-0.360
C4	0.142	0.232
	Hidroxila	
C3	0.160	0.232
O3	-0.400	-0.642
HO3	0.240	0.410

Figura 12: Comparação entre as cargas atômicas de Löwdin, utilizadas em estudos com GROMOS 43A1, e RESP, utilizadas aos parâmetros do GROMOS 53A6GLYC.

3.3 Simulações por DM

3.3.1 Protocolo de simulação

O protocolo geral de simulação foi baseado em procedimentos previamente descritos (de Groot & Grubmüller, 2001), e detalhado em cada um dos trabalhos presentes na sessão de Resultados desta Tese. A temperatura dos sistemas foi ajustada para a temperatura fisiológica, de 310 K, para os sistemas contendo moléculas envolvidas na cascata de coagulação sanguínea, mimetizando o ambiente em que elas estão inseridas, e para a temperatura ambiente, de 298 K,

para a validação dos parâmetros do novo campo de força para carboidratos GROMOS 53A6GLYC, uma vez que as propriedades disponíveis para comparação foram determinadas nesta temperatura. Em comum a estes estudos, cada molécula foi solvatada utilizando condições periódicas de contorno, empregando o modelo de água *simple-point charge* (SPC) (Berendsen *et al.*, 1987), e o método LINCS (Hess *et al.*, 1997) foi utilizado a fim de fixar o comprimento de ligações covalentes, de forma a permitir um passo de integração de 2 fs. Contra-íons (cloreto, sódio ou cálcio) foram adicionados, conforme a necessidade, de forma a neutralizar as cargas dos sistemas estudados.

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 para os sistemas relacionados à coagulação sanguínea, conforme protocolo já estabelecido em estudos anteriores do grupo, inclusive para sistemas contendo complexos proteína-heparina (Verli & Guimarães, 2005). Na validação do novo campo de força GROMOS 53A6$_{GLYC}$, as interações eletrostáticas foram calculadas utilizando o método *reaction-field* (Tironi *et al.*, 1995), utilizando um raio de corte de 12 Å, conforme protocolo previamente estabelecido para pequenos carboidratos (Lins & Hünenberger, 2005). A temperatura e a pressão dos sistemas foram mantidas constantes através do acoplamento do soluto, íons e solvente a banhos externos, respeitando os protocolos previamente citados para cada tipo de sistema, a 298 ou 310 K, e a 1 bar, respectivamente. Para os complexos proteína-heparina, barostato e termostato de Berendsen foram empregados (Berendsen *et al.*, 1984), utilizando constantes de acoplamento de τ = 0,1 ps para temperatura e τ = 0,5 ps para pressão. Na validação do GROMOS53A6$_{GLYC}$, barostato de Parrinello–Rahman (Parrinello & Rahman, 1981; Nosé & Klein, 1983) e termostato de Nosé–Hoover (Nosé, 1984; Hoover, 1985) foram empregados, utilizando constantes de acoplamento de τ = 0,1 ps para temperatura e τ = 1,0 ps para pressão.

As simulações por DM podem ser divididas em três etapas: *termalização*, *equilibração* 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, utilizada somente nos estudos de sistemas envolvendo a cascata de coagulação sanguínea, 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. Em seguida, há a *equilibração*, em que variáveis termodinâmicas e estruturais (vide seção 3.3.2) do sistema em estudo devem assumir um equilíbrio em torno de um valor. Nos sistemas envolvendo coagulação sanguínea, assumiu-se a primeira metade da DM como fase de equilibração, de forma que as análises de *recolhimento de dados* foram feitas na segunda metade da simulação. Para a validação do GROMOS53A6$_{GLYC}$, na qual os sistemas eram menores, e que possuíam uma complexidade reduzida, uma etapa de 1 ns de equilibração com restrição de posição foi empregada. Após a termalização e/ou equilibração do sistema, a simulação prossegue pelo tempo estipulado. Nesse sentido, a duração das simulações realizadas foi variável, de acordo com o sistema em estudo (de 0,1 a 1 μs), totalizando aproximadamente 2 μs para biomoléculas envolvidas na coagulação sanguínea, e 40 μs para monossacarídeos e dissacarídeos estudados na parametrização do campo de força GROMOS 53A6$_{GLYC}$.

3.3.2 Análise da estabilidade e convergência dos sistemas

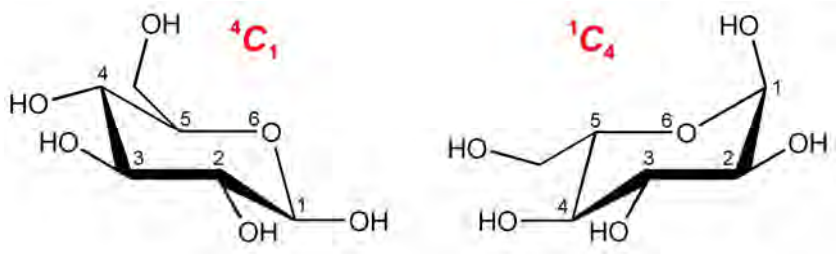
As primeiras análises a serem feitas em uma trajetória de DM visam averiguar sua estabilidade através de parâmetros termodinâmicos, tais como temperatura, pressão, densidade, volume e energia potencial do sistema. Os dados obtidos, medidos em função do tempo, devem estar equilibrados em torno de um valor (de um 1 bar, por exemplo, para pressão). Em seguida, a convergência do sistema para um estado de equilíbrio pode ser avaliada a partir do relaxamento da estrutura proteica, o que ocorre quando um platô pode ser verificado nos valores obtidos em duas análises: o *root mean square distance* (RMSD) em relação à estrutura inicial, que consiste em uma medida de distância de cada uma das estruturas calculadas ao longo da simulação em relação à estrutura de partida; e o raio de giro da molécula, que dá um indicativo do tamanho da molécula a partir da distância de cada um de seus átomos componentes em relação ao centro de massa.

3.3.3 Construção de topologias

A topologia é um arquivo contendo parâmetros para determinada molécula, através dos termos de energia descritos na sessão da Introdução 1.5 da presente Tese. Em resumo, esse arquivo contém termos para a manutenção da distância das ligações químicas e ângulos de ligação, bem como potenciais de energia para

diedros próprios e impróprios, parâmetros de interações não-covalentes e cargas atômicas. Nos estudos realizados durante a presente Tese, para o campo de força GROMOS96 43A1, utilizou-se uma ferramenta de construção de topologias denominada PRODRG (Schuettelkopf & van Aalten, 2004), que consiste em um servidor que funciona na *World Wide Web*, gerando um arquivo de topologia para os sistemas de interesse a partir de um arquivo PDB. O arquivo de saída deste servidor, no entanto, necessita de alguns refinamentos (Lemkul *et al.*, 2010). Para carboidratos, conforme previamente estabelecido (Pol-Fachin & Verli, 2008; Pol-Fachin *et al.*, 2009), incluem o acréscimo de cargas atômicas, calculadas por métodos *ab initio* (descrito no item 3.2), e a adição de diedros impróprios (Tabela 5), necessários para preservar os estados conformacionais (geralmente $^4C_1$ ou $^1C_4$) dos monossacarídeos estudados, de acordo com dados experimentais prévios. Além disso, o emprego destes parâmetros permite avaliar a influência de estados conformacionais específicos de um dado monossacarídeo, ou seja, explorando o papel destas conformações em um determinado fenômeno biológico (Verli & Guimarães, 2005; Pol-Fachin & Verli, 2008).

Tabela 5: Definição dos diedros impróprios utilizados para definir a conformação dos monossacarídeos contidos nos sistemas simulados.



Sequência de átomos definindo o diedro[a,b]	Ângulos (em graus)	
	$^4C_1$	$^1C_4$
5 – 2 – 4 – 1	2,0	-2,0
5 – 2 – 3 – 1	23,0	-23,0
5 – 2 – 3 – 6	-2,0	2,0

[a] Na figura, a glicose é utilizada como modelo. [b] Adaptado de Pol-Fachin, 2009.

3.4 Metadinâmica

Considerando que alguns eventos conformacionais, tais como enovelamento proteico, ocorrem em uma escala de tempo ainda inacessível aos tempos

computacionais disponíveis atualmente em simulações de DM convencionais (Laio & Gervasio, 2008), os métodos de DM com amostragem ampliada, tais como a metadinâmica, emergem como uma alternativa para observar esses eventos. Nesse sentido, os estudos por metadinâmica se baseiam na inclusão, durante os cálculos de DM, de uma variável, ou *collective variable* (Laio & Parrinello, 2002), que deve previamente identificada como capaz de descrever o processo de interesse (Laio & Gervasio, 2008). Sendo assim, o espaço conformacional adotado pelo sistema é determinado por essa variável, de forma que potenciais de energia são incluídos ao longo do tempo de simulação e o sistema fica impedido de retornar aos estados conformacionais adotados anteriormente (Laio & Parrinello, 2002). Para os trabalhos desenvolvidos no contexto desta Tese, as variáveis escolhidas incluem as coordenadas *puckering* (Cremer & Pople, 1975), que definem as conformações adotadas por monossacarídeos, e os diedros ϕ e ψ de ligações glicosídicas de dissacarídeos. Com isso, toda a superfície de energia livre (do inglês *free energy surface*, FES) do sistema é avaliada, indicando quais são os estados conformacionais de maior estabilidade. A descrição dos sistemas em estudo de forma mais minuciosa ou mais rápida (ou seja, com menor custo computacional) pode ser alterada através da variação dos valores de *height* (h) and *width* (σ). Visando uma descrição mais detalhada dos carboidratos avaliados no contexto da presente Tese, foram usados, respectivamente, valores de 0.1 e 0.5. O protocolo de simulação por metadinâmica utilizado durante a validação nos novos parâmetros do GROMOS 53A6GLYC foi baseado em estudos prévios para monossacarídeos (Autieri *et al.*, 2010) e detalhado na sessão de materiais e métodos do trabalho IV da presente Tese. Uma abordagem visual e dinâmica, que pode facilitar o entendimento desta técnica, pode ser visualizada no link http://goo.gl/jbLgB.

3.5 Construção de mapas de contorno para carboidratos

Quanto maior a complexidade estrutural da molécula, maior será a dificuldade e o custo computacional associado à obtenção de uma amostragem adequada para o espaço conformacional adotado pelo oligossacarídeo. Isso fica ainda mais evidente no contexto de carboidratos, uma vez que são moléculas altamente flexíveis, em que múltiplos subestados conformacionais podem co-existir. Como uma alternativa, essas estruturas podem ser construídas a partir de seus dissacarídeos constituintes (Naidoo *et al.*, 1997; Mandal & Mukhopadhyay, 2001; Fernandes *et al.*,

2010), cujas conformações podem ser obtidas por diferentes metodologias, incluindo a construção de mapas de contorno, que podem ser considerados análogos aos mapas de Ramachandran. Ao longo de dez anos de estudos envolvendo carboidratos em suas diversas formas, nosso grupo de pesquisas estudou entre 100 e 200 dissacarídeos diferentes (Becker *et al.*, 2007; Pol-Fachin & Verli, 2008; Castro *et al.*, 2009; Fernandes *et al.*, 2010; e artigos da presente tese), em uma primeira etapa, através de mapas de contorno, seguidos de um refinamento por simulações de DM em solução e, mais recentemente, por metadinâmica em solvente aquoso. Para a presente tese, duas técnicas diferentes foram empregadas na construção desses mapas, como descrito abaixo.

3.5.1 Por mecânica molecular

Nos estudos envolvendo a construção de mapas de contorno por mecânica molecular, ou DM convencional, utilizou-se 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). No caso de ligações 1→2, 1→3 e 1→4, os ângulos Φ e Ψ foram rotados de -180º a 150º, em passos de 30º, gerando um total de 144 confôrmeros para cada dissacarídeo. No caso de ligações 1→6, em que a ligação glicosídica é composta por três diedros: Φ, Ψ e ω, foram gerados 144 confôrmeros para cada par de ângulos, ou seja, 144 para o par Φ *versus* Ψ e 144 para o par Ψ *versus* ω. O protocolo empregado 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). Durante essas minimizações de energia e passos de DM, utilizam-se restrições (diedros impróprios) somente para os diedros da ligação glicosídica em estudo. Com isso, esses ângulos são fixados em cada uma das 144 possíveis combinações dos dois diedros em estudo, permitindo que o restante da molécula se acomode em relação à geometria da ligação glicosídica de interesse. Os valores de energia dessas 144 conformações são então empregados na construção de um mapa de contorno, considerando-se a estabilidade relativa à conformação mais estável de todas.

3.5.2 Por metadinâmica

Os estudos envolvendo a construção de mapas de contorno por metadinâmica objetivaram a validação de um novo campo de força, parametrizado ao longo da presente tese, intitulado GROMOS 53A6$_{GLYC}$. Nesses estudos, os dissacarídeos foram submetidos a uma minimização de energia utilizando o algoritmo de gradiente conjugado e, em seguida, a um passo de equilibração de DM de 1 ns, no qual o dissacarídeo é mantido preso por restrição de posição e as moléculas de água se acomodam ao seu redor. A metadinâmica é então realizada com os mesmos parâmetros de *height* (h) and *width* (σ) descritos acima (seção 3.4), avaliando a rotação dos diedros ϕ e ψ das ligações glicosídicas de interesse.

3.6 Validação das simulações de DM

A validação de estudos de simulação por DM está associada à comparação de propriedades observadas computacionalmente com propriedades experimentais, de forma que a precisão do método é avaliada a partir da verificação de quão precisamente variáveis previamente conhecidas são reproduzidas (van Gunsteren & Berendsen, 1990; Karplus & Petsko, 1990). No contexto da presente Tese, a validação dos resultados obtidos foi realizada através da comparação com (1) dados de RMN, tais como as geometrias de ligação glicosídica para os dissacarídeos componentes de glicoproteínas; (2) dados de cristalografia de raios-X, tais como na comparação do perfil de *B-factor* com o perfil de flexibilidade das proteínas, avaliado por *root mean square fluctuation* (RMSF), que avalia a flutuação de cada átomo, ou resíduo, na forma de um valor médio para a simulação; (3) dados de mutagênese sítio-dirigida, tais como na avaliação da importância de determinados resíduos de aminoácido na interação entre AT e heparina; (4) dados bioquímicos, tais como sobre a atividade catalítica das proteases estudas e a afinidade das diferentes glicoformas de AT à heparina; entre outros. A partir da validação das simulações, predições baseadas nos cálculos de DM podem ser feitas, cuja credibilidade também depende (1) da precisão do campo de força utilizado; e (2) do tamanho da amostragem do espaço conformacional, que necessita ser suficientemente ampla para descrever as propriedades do sistema (van Gunsteren & Mark, 1998).

4 Resultados

4.1 Preâmbulo

Os resultados obtidos, bem como documentos representativos do conjunto de assuntos abordados ao longo da presente Tese, serão apresentados a seguir na forma de trabalhos submetidos ou publicados. Uma breve descrição sobre cada um, e o contexto no qual eles se inserem na presente Tese, está apresentada abaixo.

- **Laércio Pol-Fachin**, Hugo Verli: Assessment of Glycoproteins Dynamics from Computer Simulations. *Mini Rev. Org. Chem.*, **2011**, *8*; 229-238.

Este artigo revisa os diferentes métodos e protocolos para o estudo de glicoproteínas, bem como trabalhos realizados nesse âmbito. No contexto da presente Tese, este artigo apresenta uma explicação aprofundada da metodologia de simulação por DM de glicoproteínas, e representa os trabalhos publicados envolvendo o estudo da estrutura, conformação e função dessas biomoléculas não envolvidas em coagulação sanguínea (vide seção 9 - Anexos):

- Pol-Fachin & Verli, *Glycobiology*, **2012**, *22*, 817-825;
- Virgens *et al.*, *J. Biomol. Struct. Dyn.*, **2013**, doi: 10.1080/07391102.2013.780982;
- Pol-Fachin & Verli, *Kerala: Transworld Research Network*, **2010**, 103-131;
- Velásquez *et al.*, *Science*, **2011**, *332*, 1401-1403.

- **Laércio Pol-Fachin**, Camila Franco Becker, Jorge Almeida Guimarães, Hugo Verli: Effects of glycosylation on heparin binding and antithrombin activation by heparin. *Proteins*, **2011**, *79*; 2735-2745.

Este trabalho avalia a interação entre heparina e as duas glicoformas circulantes de AT no plasma humano (α e β), bem como as consequências dessa interação sobre a estrutura, conformação e função da serpina.

- **Laércio Pol-Fachin**, Hugo Verli: Structural glycobiology of heparin dynamics on the exosite 2 of coagulation cascade proteases: Implications for glycosaminoglycans antithrombotic activity. *Glycobiology*, **2014**, *24*; 97-105.

Este artigo, em continuação ao anterior, avalia a interação entre heparina e as proteases da coagulação sanguínea fIIa e fXa, relacionando os dados obtidos com sua atividade catalítica e com os mecanismos de ação da heparina na inibição dessas enzimas por AT.

- **Laércio Pol-Fachin**, Victor Holanda Rusu, Hugo Verli, Roberto Dias Lins Neto: GROMOS 53A6$_{GLYC}$, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Chem. Theory Comput.*, **2012**, *8*; 4681-4690.

Por fim, este trabalho propõe um novo conjunto de parâmetros para o estudo de carboidratos utilizando campos de força da série GROMOS, com possibilidade de utilização em conjunto com parâmetros pré-existentes para proteínas, lipídeos ou ácidos nucleicos. No contexto da presente Tese, este artigo representa a parte do trabalho realizado junto ao Laboratório de Biomateriais, do Departamento de Química Fundamental da Universidade Federal de Pernambuco, realizado sob orientação do prof. Roberto Dias Lins Neto.

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

4.2 Trabalho I

Este trabalho é o primeiro artigo da literatura que revisa os diferentes métodos de avaliação de glicoproteínas através de simulações computacionais. O emprego de técnicas experimentais, como cristalografia de raios-X e RMN, no estudo de glicoproteínas é prejudicado principalmente devido à elevada flexibilidade das cadeias oligossacarídicas dessas biomoléculas. Sendo assim, metodologias computacionais, quando propriamente validadas frente a esses dados experimentais, consistem em uma ferramenta promissora para o estudo estrutural e funcional de glicoproteínas, levando em consideração tanto a flexibilidade quanto o ambiente no qual essas moléculas estão inseridas.

Inicialmente, este trabalho discute as diferentes formas para obtenção de modelos iniciais adequados para a estrutura e conformação de glicoproteínas, levando em consideração a parte proteica, e especialmente a porção sacarídica dessas biomoléculas. Nesse sentido, as diferentes estratégias para o estudo conformacional de carboidratos são apresentadas, com foco naquelas envolvendo mecânica molecular. Em seguida, o refinamento dos modelos de glicoproteínas é discutido, considerando os métodos potencialmente empregáveis e os dados adequados para sua comparação e validação. Ainda, uma breve revisão da literatura é realizada, a respeito da metodologia utilizada e dos resultados obtidos, bem como perspectivas futuras.

Assessment of Glycoproteins Dynamics from Computer Simulations

Laércio Pol-Fachin, Hugo Verli

Mini Reviews in Organic Chemistry, **2011**, *8*; 229-238

Mini-Reviews in Organic Chemistry, 2011, *8*, 229-238

Assessment of Glycoproteins Dynamics from Computer Simulations

Laercio Pol-Fachin[a] and Hugo Verli[a,b,*]

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

[b]*Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Av Ipiranga 2752, Porto Alegre 90610-000, RS, Brazil*

Abstract: Glycoproteins are synthesized in most of the living organisms, being major components of the outer surface of mammalian cells and most of the secreted proteins in eukaryotes. Accordingly, a better comprehension of the biological processes in which they are involved requires the characterization of their structure and conformation. As such rationale faces several difficulties from both experimental and theoretical approaches, this review summarizes the current state of the art and methods employed to model and represent glycoproteins, including their carbohydrate moieties, through computer simulations.

Keywords: Conformational analysis, ensemble, glycan, glycoprotein, glycosylation, molecular dynamics, solvation.

1. INTRODUCTION

Glycoproteins are major components of the outer surface of mammalian cells and represent most of the secreted proteins in eukaryotes [1]. In fact, carbohydrate moieties may be covalently attached to polypeptide chains in most of the living organisms [2]. Such oligosaccharides are reported to be able to modify several properties of such molecules (Table **1**) [3-5], including solubility [6,7], folding and conformation [8-10], which, in turn, may influence glycoproteins biological roles. In this context, the comprehension, at the atomic level, of their interaction with solvent and target biomacromolecules passes through the characterization of their structure and conformation through experimental and/or theoretical techniques.

Table 1. Effects of Glycosylation on Proteins and their Surrounding Environment

Property	Effect[a]
Physico-chemical	Resistance to proteolysis
	Resistance to denaturation
	Increasing solution viscosity
	Lowering solution freezing point
	Increasing protein solubility
Enzymological	Shifting optimum pH
	Altering catalytic activity
Folding	Preventing aggregation
	Facilitating interaction with chaperones
	Nucleation of β turns
Biological activity	Altering protein–protein recognition
	Altering protein–carbohydrate recognition
	Altering transport and secretion tax
	Increasing / Decreasing multimerization

[a] Data from [3-5].

From the experimental point of view, the high flexibility of carbohydrate moieties may be a hindering factor for their crystallization, as well as their high degree of coordination to water molecules and the lack of strong lipophilic or dipolar inter-residue interactions [11]. In addition, the crystal environment may be capable to bend their structure, as a recent survey of Protein Data Bank (PDB) entries containing oligosaccharides suggests that about one-third of them contain significant errors in carbohydrate stereochemistry, nomenclature and consistency with the electron density [12,13]. As mentioned by Crispin and co-workers, some of the proposed models contain not only systematic errors in carbohydrate stereochemistry, but also hitherto unreported motifs in the primary structures of the glycans [13]. For example, there are approximately two hundred cases for which the PDB had to assign, by stereochemistry matching, the incorrect 2-(acetylamino)-2-deoxy-α-D-glucopyranose (NDG) rather than the correct 2-(acetylamino)-2-deoxy-β-D-glucopyranose (NAG) [14], which have been considered to differentially affect proteins structure and dynamics [15]. Unlike X-ray crystallography, that determine carbohydrates conformation at the crystal environment, NMR spectroscopy provides a set of spatial constraints representing solution averaged conformations. Unfortunately, NMR techniques may face difficulties in supplying an adequate number of NOE signals for carbohydrates three-dimensional (3D) characterization [11]. In addition, the derived solution averaged conformational states may not correspond to conformations populated in solution [16]. From the theoretical point of view, when properly validated and in conjunction with the above-mentioned experimental

*Address correspondence to this author at the Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil; Tel: +55-51-3308-7770; Fax: +55-51-3308-7309; E-mail: hverli@cbiot.ufrgs.br

1570-193X/11 $58.00+.00

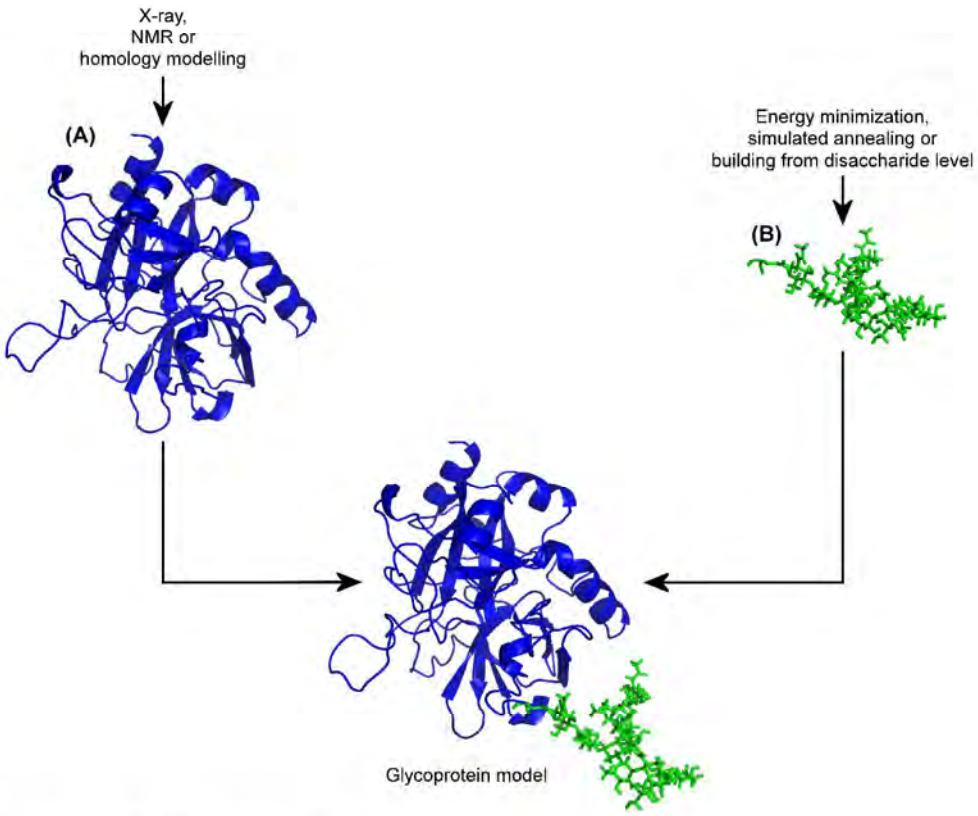


Fig. (1). Schematic representation of building a glycoprotein model.

Table 2. Available Resources for Building Glycoproteins and their Carbohydrate Moieties

Resource	Description	Ref
Online Glycoprotein Builder	Building oligosaccharides and their linking to proteins	[18]
Carbohydrate 3D Structure Predictor	Generation of 3D models for carbohydrates	[18]
Glydict	Prediction of 3D structures for N-oligosaccharides	[19]
Glyprot	Attachment of N-linked oligosaccharides to proteins	[20]

information, as well as relevant biochemical data, computer simulations emerge as a promising tool, capable to describe glycoproteins and mimic biological solutions.

2. GENERAL IDEA ON BUILDING MODELS FOR GLYCOPROTEINS

While the simulation of proteins may employ a starting geometry derived directly from PDB or from comparative modeling, computational studies on glycoproteins require additional steps as, in most cases, their complete forms are still not available in PDB. As a general feature, as illustrated in Fig. (**1**) (using thrombin structure under PDB ID 1DOJ as a model [17]), these steps include: (**I**) obtaining a 3D structure for the protein moiety; (**II**) obtaining a 3D structure for the carbohydrate moiety; and (**III**) attachment of the saccharidic and peptidic segments into a final model. In this context, while web-based resources [18-20] may be employed to build glycan chains and attach them to proteins (Table **2**), each step may be completed separately using distinct approaches and refinements and so supporting a fine tuning of the glycan conformation.

Regarding the first step, which embraces the obtaining of 3D structures for the protein moiety (Fig. (**1A**)), retrieving an X-ray or NMR structure from PDB data bank is generally the better option. Additionally, comparative modeling may be adequately employed in the absence of previous 3D experimental data to generate realistic models, given that the homologous sequence (target) shares with the experimentally established protein structure (template) significant sequence (30% or more) or structural similarity [21-23]. As well, in order to increase the accuracy of the produced models, the conformational space of such protein may be sampled, which consist in a good test and application for simulations methods as molecular dynamics (MD) [22].

In relation to the second step, comprising the achievement of 3D models for the carbohydrate moiety, theoretical approaches represent the most accessible source of information, seeing that some difficulties may be faced on obtaining meaningful experimental models for glycans (as above-discussed). Yet, X-ray structures, NMR models, *J* coupling constants, ring puckering or NOE signals should be employed, whenever available, to validate the obtained conformations.

Finally, in the third stage, a model for the studied glycoprotein is obtained by linking both protein and carbohydrate moieties in conformity to specific geometrical terms, which are well known for the *N*-glycosidic linkage [24,25], but mostly absent for other types of carbohydrate-amino acid linkages (see further in the text). However, as it will be discussed below, given the dynamical features of carbohydrates, the use of static, single models of glycoproteins for

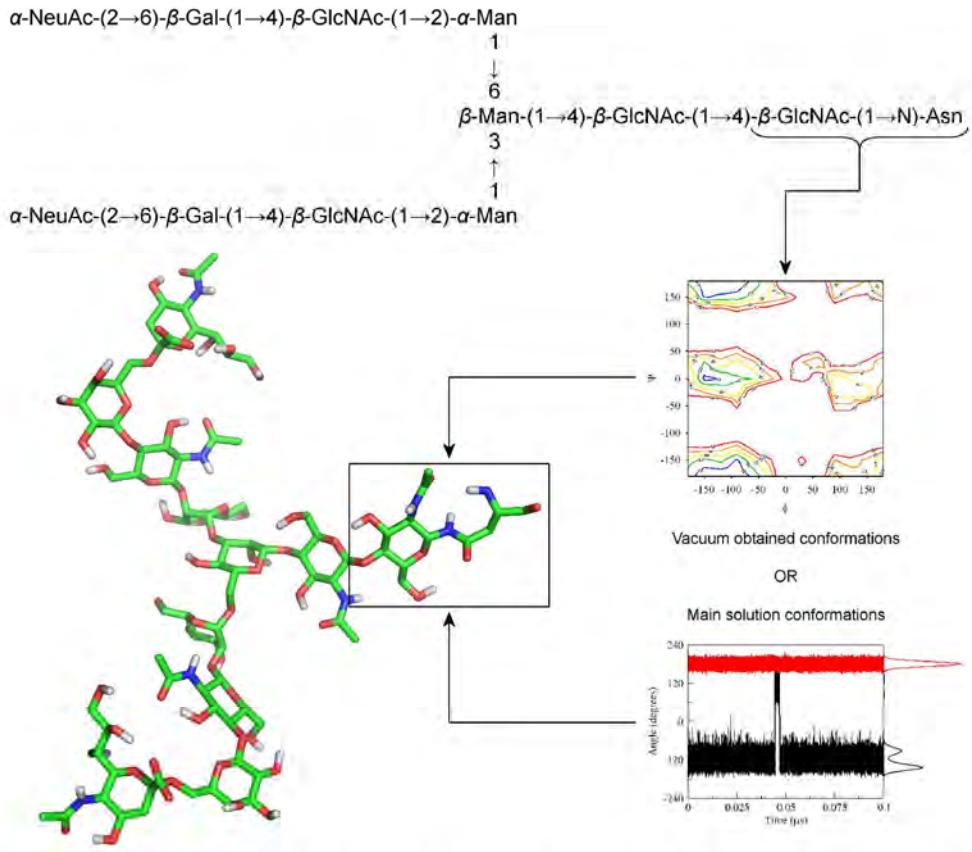


Fig. (2). Schematics of constructing a glycan chain from the disaccharide level. The conformational preferences of each glycosidic linkage composing a given oligosaccharide may be obtained from energy contour plots or solution simulations (β-GlcNAc-(1→N)-Asn employed as an example). Such conformational states from disaccharide units may be further combined in order to build the 3D model for the studied carbohydrate moiety.

interpretation of structural, functional and biological aspects of this class of biomolecules, should be carefully evaluated.

3. OBTAINING MODELS FOR GLYCANS

Carbohydrates are considered to have several orders of magnitude higher potential information content than any other biological macromolecule [26], mainly due to their great structural diversity, comprehending: the number of possible monosaccharide units, the formation of linear or branched structures, the two stereochemical possibilities on the linkages between saccharide units (α or β), the two potential isomeric forms (-D or -L) and covalent modifications in sugar residues, as methylation, sulfation, acetylation and phosphorylation [27]. In fact, while mammalian glycans rely on a group of approximately 10 common monosaccharide units, including *N*-acetyl-D-glucosamine (GlcNAc), D-mannose (Man), D-galactose (Gal), L-fucose (Fuc), neuraminic acid (NeuAc), *N*-acetyl-D-galactosamine (GalNAc) and D-glucose (Glc) [28,29], glycans assembled by other organisms may include an indeterminate number of such building blocks, employing a series of additional monosaccharide units, not found in mammalian organisms, such as pentoses [2,27,30]. All these properties make carbohydrates one of the most challenging classes of biomolecules for conformational characterization [11].

3.1. Strategies for Conformational Analysis of Carbohydrates

The methods employed in conformational analysis of compounds may generally fall into three distinct categories: (1) those that are random or stochastic, as MD, Monte Carlo and distance geometry based techniques; (2) those based on heuristics and artificial intelligence methods; and (3) those that are systematic [31]. While Monte Carlo approaches were shown to support a solid search for the conformational space of carbohydrates [32], strategies employed to date, specifically for modeling entire carbohydrate moieties of glycoproteins, include energy minimization [33] and simulated annealing [34]. However, the greater is the glycan moiety complexity, more difficult becomes the sampling of the molecule conformational space. Therefore, alternatively, such oligosaccharides may be built from the disaccharide level [35-37], in which minimum energy conformers from systematic analyses and/or MD derived conformations of isolated disaccharides may be used to build the complete oligosaccharide (Fig. (**2**)). This simplification appears to not impair the assessment of reliable glycoprotein models when compared to experimental data [37,38].

The determination of disaccharides conformational preferences are usually performed by describing their preferred conformations on potential energy surfaces as a function of their glycosidic torsional angles Φ and Ψ [26]. Several methods may be employed for calculating such maps. As it is usually considered that the most important energy variations are those related with the glycosidic dihedral angles [39], the hard sphere potential surfaces approach considers the constituent monosaccharides as rigid spheres, including the exocyclic groups. As a consequence, the glycosidic linkage geometry is determined by the spatial arrangement of such spheres [40]. Whereas simplified, such approach has being recently shown to support a reliable conformational description of complex systems in coarse grained MD simulations [41-43]. In spite of that, important variations in pyranoid ring geometries and orientations of pendent groups associated with Φ and Ψ rotation may be observed, which emphasize the need for a model to include bond length and angle degrees of freedom in some cases [40]. For instance, by al-

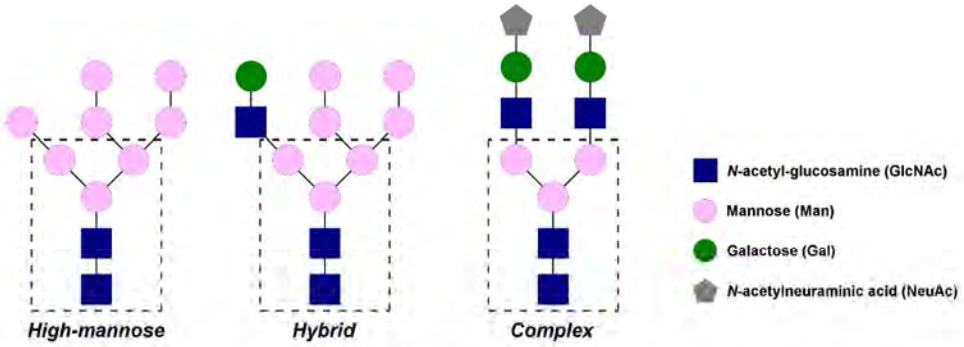


Fig. (3). The three main types of *N*-glycans observed in eukaryotes. The core pentasaccharide is shown inside the dashed lines.

lowing the atom coordinates to be minimized, a relaxed contour plot may be obtained, consisting in a fast approach for describing disaccharides conformational preferences. Such approach usually allows a lowering in both energy barriers between minima and energy of global and local minima, in comparison to maps obtained with hard sphere potential surfaces technique [40]. However, while in relaxed maps only rotations around glycosidic linkage dihedrals are systematically sampled, it should be considered that the orientation of secondary hydroxyl, hydroxymethyl and primary hydroxyl groups, as well as the different degrees of ring puckering, may influence the calculated final energy [39,44]. Therefore, when all of such angles are taken into account for searching the lowest energy of each point in the Φ-Ψ space, an adiabatic map is obtained.

Moreover, *ab initio* [45,46], density functional theory (DFT) [47,48] and molecular mechanics (MM) [49] have been employed to generate energy contour plots for disaccharides. While *ab initio* and DFT procedures may better describe the aspects determining glycosidic linkages geometry (as the *exo*-anomeric effect), and it may be difficult to properly describe such interactions in MM force fields [11], when such parameters are well validated, MM may provide adequate agreement with experimental data. As well, the many local minima possibly existing in *ab initio* or DFT obtained conformational maps for disaccharides may masquerade each other, reducing the contour plot quality [45]. Also, considering that MM is much less time demanding than quantum mechanical calculations, the former methodology is usually more applied to study the conformational preferences of disaccharides.

3.2. Force Fields for Carbohydrates

In the context of the carbohydrate moieties within glycoproteins, few groups of parameters have been employed for their conformational description, mainly by means of building energy contour plots in vacuum, as CHARMM [50], CVFF [51] and GROMOS96 [52,53]. Nevertheless, the largest amount of data about such disaccharides is based on MM3 [54,55], which is recognized to offer a highly detailed representation of carbohydrates conformational features in vacuum [56,57], whereas not used for representing glycoproteins (see further). The so obtained disaccharide conformations have been used in tools for 3D prediction of N-linked oligosaccharides (Glydict) [19] and for their attachment to proteins PDB structures (Glyprot) [20]. Such topic, including several other force field parameters for carbohydrates, is reviewed elsewhere [58,59].

Conformational search methods have indeed been employed to characterize the conformational behavior of carbohydrates [60], mainly through energy contour plots for disaccharides in vacuum. On the other hand, explicit solvent simulations have been considered as capable to better reproduce the conformational properties of oligosaccharides in comparison to calculation in its absence [32,61-64]. In spite of that, such vacuum derived conformations are not commonly associated with solution MD simulations, in order to achieve a proper conformational ensemble [60].

3.3. Solution Simulations of Carbohydrates

The explicit inclusion of solvent molecules has been described to reveal a distinct set of conformers when compared to calculations in its absence [65]. Specifically in the case of the carbohydrate moiety of glycoproteins, which present a high degree of branching, solvation appeared to be required to disclose conformations closer to those observed by experimental data [35,37]. In this context, while branching has been proposed to stiffen oligosaccharide main chains [11,64], the majority of conformational transitions in the carbohydrate moiety of glycoproteins have been observed to occur at their branching linkages [35]. Yet, the conformation and dynamics of these glycans [35,37] and other branched oligosaccharides [64], obtained from computer MD simulations with diverse force field parameters, have been observed to be in accordance to previous experimental data, mainly NMR, which further support the validity of solution simulations on these systems.

In fact, the role of solvation on the conformational preferences of glycosidic linkages composing the pentasaccharidic central core of *N*-glycans (Fig. **3**), where some branching points are located, is reinforced by several studies [66,67]. Regarding the β-D-Man-(1→4)-β-D-GlcNAc-(1→4)-β-D-GlcNAc unit, a more entropically favored conformation could be observed after solvation, in comparison to the vacuum structure, by means of a glycosidic linkage geometry modification and further interaction with water molecules [67]. In addition, residency times for water molecules were observed to be highly prevalent around the central core pentasaccharide, influencing the flexibility and overall topology of an oligomannose *N*-glycan [66]. These data reinforce the importance in submitting carbohydrate structural models to simulation techniques, if possible under explicit solvent conditions, as a strategy to obtain reliable solution models for such class of molecules.

4. SETTLING AND REFINING MODELS FOR GLYCOPROTEINS

After suitable models for both protein and carbohydrate moieties are obtained, the glycosylated amino acid residue and the first monosaccharide of the glycan must be correctly attached, as the geometry of such glycosidic linkages determines the exposure of the glycan chains on the protein surface. To date, about five different glycosylation types are identified, comprising ~40 glycosidic linkages occurring in nature [2], distributed in essentially all living organisms (Table **3**). In spite of such diversity, experimental data on the conformational preferences of such linkages are mostly restricted to the *N*- glycosyl bond (β-GlcNAc-(1→N)-Asn) [24,25], while only ~10 *N*- and *O*-glycosylation motifs have been hitherto conformationally studied by molecular modeling techniques [38,68-72]. In the absence of an adequate picture of the conformational features associated to such additional types of monosaccharide-amino acid linkages, as C-glycosylation and P-glycosylation, the

Table 3. Phylogenetic Distribution of Glycosidic Linkages Between Monosaccharides and Amino Acids

Glycosylation Type	Distribution[a]		
	Eukaryotes	*Archaea*	Bacteria
N-glycosylation	+	+	+
O-glycosylation	+	+	+
Glypiation	+	+	–
C-glycosylation	+	–	–
P-glycosylation	+	–	–

[a] Based on data from reference [2];

Table 4. Glycoproteins MD Simulations Studies

Glycoprotein	Number of Models	Simulated Time (Each)	Refs
Lectin[a]	1	0.3 ns	[35]
gp120[b]	2	0.1 ns	[74]
MHC Class I	4	0.5 ns	[36]
Human prion protein	1	~ 2 ns	[33]
Antifreeze glycoprotein 8	10	2 ns	[75]
Human coagulation fVII	1	~ 3 ns	[76]
Human mucin[b]	3	1 ns	[77]
Hemagglutinin[a]	2	10 ns	[15]
EGF-*like* domain	1	50 ns	[38]
CD59[c]	3	50 ns	[38]
Human CD2 domain	2	50 ns	[38]
α-subunit of hCG[d]	2	50 ns	[38]
Ovine ciclooxygenase-1	1	50 ns	[37]
Murine ciclooxygenase-2	1	50 ns	[37]

[a] From *Erythrina corallodendron*;
[b] Glycopeptides derived from such molecules;
[c] Human complement regulatory protein CD59;
[d] Human chorionic gonadotropin.

derived glycoprotein models may probably reveal biologically irrelevant protein-carbohydrate interactions, thus failing to properly describe one of its main structural aspects.

Subsequent to the obtaining of a glycoprotein model, refinement techniques may be employed to enhance its biological meaning. While some studies make use of energy minimization, for instance, to minimize bad steric interactions [73], solution simulations are often employed for this purpose [15,33,35-38,74-77], also supporting an investigation of the conformational space associated to such macromolecules. As MD simulations are a usual choice, it carries the challenge of covering sufficiently long time scales, necessary to adequately represent biological phenomena and/or experimental data [78]. For instance, reproduction of atomic force microscopy (AFM) data based on steered molecular dynamics (SMD) is not readily achievable due to the gap between the time scales of computer simulations (up to ~1 μs) and AFM measurements (~1 s) [79]. In spite of that, the recent advances regarding both hardware and software have supported a progressive increase in timescales accessible through MD simulations. The computational cost associated with carbohydrate simulations, for instance, had supported conformational samplings in the microsecond timescale, providing detailed information on such compounds dynamics, partially inaccessible to experiment, at the monosaccharide [80], disaccharide [81] and oligosaccharide [82] levels. Conversely, in the context of glycoproteins, the difficulties in achieving higher time scales are proportional to the systems size, being usually in the order of tens of nanosecond [37, 38].

4.1. Conformational Ensemble and Multiple Models

Conformational fluctuations are considered essential to the functions of proteins [78] and, consequently, of glycoproteins. In this context, simulation methods, as MD simulations, are capable to provide data on macromolecules conformational ensemble at the atomic level [83]. In order to obtain a proper conformational sampling during MD simulations, the energy barriers associated to the system degrees of freedom should be overcome in order to allow its configurational space to be explored [84]. While this sampling may be achieved by means of long time simulations, conversely, multiple simulations may be performed for the same system, starting from diverse initial conformations, which may support divergent [85] and convergent conformational descriptions [86] at the achieved time scales. Accordingly, multiple starting geometries had been employed when simulating glycoproteins, usually being obtained from multiple NMR models, multiple starting geometries for a given glycan moiety or multiple glycan compositions for a same glycoprotein, as demonstrated in Table **4** [15,33,35-38,74-77].

The use of multiple starting models is particularly important when simulating glycoproteins as a strategy to accurately describe their glycan moieties flexible pattern [11,26]. As illustrated in Fig. (**4**), a single glycoprotein 3D structure is not capable to properly represent the conformational ensemble adopted by any of three glycan chains determined by distinct NMR studies [87-89]. Similarly, a single glycoprotein static model is not representative of its plasticity, requiring an adequate conformational ensemble description, as obtained by means of simulation based methods. In fact, the

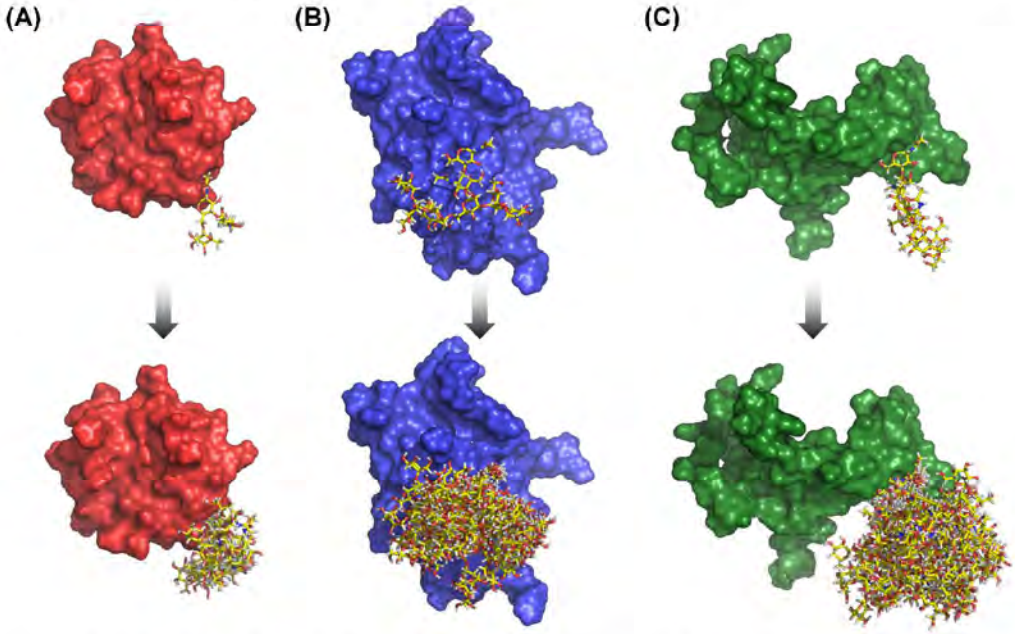


Fig. (4). Comparison between single and multiple models for representing the carbohydrate moiety of glycoproteins.

inclusion of molecular motion in glycoproteins study, as well as protein-carbohydrate complexes, may not be only important to evaluate glycosidic linkages modifications in the carbohydrate moiety, but also alterations in monosaccharide conformations. In this context, while chair conformations are considered the most populated in aqueous solutions [90], different ring puckering conformers have been described as relevant for some systems, both when free in solution [80,81] or bound to a target molecule [91].

4.2. Force Fields for Glycoproteins

As a general feature, in order to allow the study of glycoproteins, both carbohydrate and protein moieties should be described by the same, properly validated, force field [11]. In this context, most MD simulation programs are well-suited for modeling the protein part, while a lower number of them have tools for modeling of oligosaccharides moiety [26,58]. On the other hand, the linkage between them is considered, at times, to not be correctly taken into account [68], which indicates that a proper parameterization of the linkages between amino acids and monosaccharides is an important challenge for future refinements in force fields for glycoproteins. For example, while five different types of connections have been so far observed to occur in living organisms (Table **3**), studies on their conformational preferences have been reported for only two of them, that is, N- and O-glycosidic linkages. Accordingly, advancements in the comprehension of their molecular behavior, by means of new and better force field parameters, may further contribute in increasing our knowledge on glycoproteins biological functions.

Furthermore, to properly mimic glycoproteins biological environment, water molecules and counter ions are usually considered during simulations. In this context, taking into account that some of the parameters most employed to study isolated disaccharides (as the MM3 force field) were developed for the gas phase, their applicability for describing such molecules in solution is unclear [57]. Moreover, the behavior of such force fields when applied to simulate proteins in solution is also a matter of debate [56]. Nevertheless, other sets of parameters have been customized for studying glycoproteins, including AMBER [92], AMBER – GLYCAM [93,94], CHARMM [50], CVFF [51] and GROMOS96 [38]. Unfortunately, no comparative study between them has been carried out, as there has been for carbohydrates [57,58]. Still, such studies had been able to add insights into glycosylation effects over polypeptide chains.

4.3. Mutual Influence Between Protein and Carbohydrate Moieties

The comprehension of carbohydrates role on the biological function of glycoproteins pass through the comprehension on how protein and carbohydrate moieties interact and mutually influence each other in a single biomacromolecule. Such analyses, frequently based on root mean square deviation (RMSD) measurements, are mainly focused on the carbohydrate influence on the protein flexibility [33,38,77]. Among several studies comprising computer simulations of glycoproteins [33,38,74,75,95], conformational stabilization is frequently characterized as an important influence of glycosylation over polypeptide chains, mainly by means of mobility restriction. Such effect, when occurring close to the glycosylation site, considering *N*-linked glycans, is mostly attributed to the central core pentasaccharide (Fig. **3**) [10]. As well, it is proposed that such effect is raised when the number of oligosaccharides bounded to the protein is increased [9]. Considering that an N-linked oligosaccharide typically show from two to three branches, each presenting up to four monosaccharide residues, glycans may be considered to cover a vast part of the protein surface [1]. As a consequence, such oligosaccharides may influence glycoproteins conformation and dynamics in different regions of the polypeptide sequence, both through direct intra-molecular interactions and mediation of solvent molecules.

Considering the evaluation of the protein influence over the glycan chain, although RMSD is also used to measure oligosaccharides mobility [35,36], analysis on glycosidic linkage geometries may offer a more complete picture of carbohydrates conformational preferences. Nevertheless, the protein scaffold is recognized to influence some properties of their linked oligosaccharides, mainly by reducing the conformational flexibility of glycans [35,36]. Such restriction has been observed to occur due to a reduction on the number of possible conformers assumed by the glycosidic linkages that compose these glycan chains [35], possibly as a result of interactions between the polypeptide chain and their attached carbohydrate moieties [96].

4.4. Validating Simulations

Regarding the protein moiety of glycoproteins, the huge amount of data available on their structure and dynamics usually supports the process of validation. For instance, X-ray structures provide the most complete description of a structure, at the atomic level [97], and can be easily obtained from PDB. However, it is recognized that the crystal environment may produce packing effects, capable to influence protein conformation [98-101]. Moreover, the majority of our knowledge on proteins comes from time- and/or ensemble-averaged experiments [83]. In this context, when such data, as solution NMR, are used for comparison and validation, computer simulations must also consider ensemble-average properties, thus treating the obtained results in a manner that is analogous to what happens experimentally [83]. Besides, different protein force fields are recognized to behave comparably during MD simulations in relation to several structural and dynamical properties, such as solvent-accessible surface area, radius of gyration, deviation from their respective experimental structures and secondary structure (in this case, in the absence of glycan chains) [102].

While data related to the structure and dynamics of proteins are abundant, conformational information on carbohydrate moieties of glycoproteins, to be used for validation, are quite incipient. This occurs mainly due to the resistance of oligosaccharides to crystallization, the frequently low number of inter-residue NOEs and the difficulties associated with interpreting NOEs in terms of conformation [11]. When available, NOE signals should be employed as a measure of internuclear distances for validation. Nevertheless, some of such glycans were well described by NMR methods [87-89], providing multiple models and a realistic set of solution-averaged conformations to be compared to computer simulations. However, the glycosidic linkages composing glycoproteins carbohydrate moieties have been barely analyzed in most of the studies to date, which impairs a better depiction of such glycans conformational preferences.

Additional sources of experimental information on carbohydrates conformation may be obtained through *J* coupling constants [103]. In this context, the glycosidic linkage angles may be determined through inter-glycosidic $^{3}J_{HCOC}$, $^{3}J_{CCOC}$ and $^{2}J_{COC}$ couplings by employing Karplus-type correlation curves [104-106]. As well, vicinal $^{3}J_{HH}$ may be useful when analyzing ring puckering coordinates [107,108], that is, for determining endocyclic torsion angles required for describing ring conformations (for instance, in terms of chairs, boats or skew-boats), mainly by means of Cremer-Pople puckering parameters [109].

5. FUTURE PROSPECTS

While it has been proposed that more than half of known proteins can be potentially glycosylated [110,111], no more than twenty of them have been studied so far by molecular modeling techniques. As well, future advances in hardware and software, which allow further increase in reachable timescales for glycoproteins simulations, may be expected to provide further progress in relation to the comprehension of biological processes involving glycoproteins. An important aspect, for instance, comprise evaluating whether point modifications in carbohydrate or amino acid residues are capable to modify glycoproteins conformation, dynamics and, ultimately, function.

Whereas protein-protein interactions are considered vital to almost all cellular processes [112], the direct recognition between proteins and glycan chains are also considered important to a variety of biological processes [113,114]. While molecular docking is usually the main strategy employed for understanding such contacts, additional parameterization efforts are still required for both residues (amino acids and monosaccharides) flexibility and on the available scoring functions, in order to circumvent the high number of false positive results [115,116] and to support a more accurate description for the pertinent complexes. Moreover, free energy calculations may be expected to contribute in such goals, as previously achieved with protein-protein [117] and protein-carbohydrate [118,119] complexes. Unfortunately, the currently available data pointing to the importance of explicit solvent models for the obtaining of accurate conformational descriptions of carbohydrates [62] suggest a potential limitation of continuum solvent methods as MM-PBSA, even so efforts have been made recently in this area [120,121]. As well, force field parameterization and adequate conformational sampling may represent important challenges to a far-reaching use of Free Energy Perturbation [122,123] and Linear Interaction Energy [124,125] strategies in such systems.

Additionally, coarse graining approach emerge allowing micro and millisecond timescales to be sampled, with low computational cost, as well as the study of multi-macromolecular systems, comprising protein and carbohydrates separately [42,43]. As new parameters are developed and included in the currently available force fields, or new and more complete set of parameters are developed, including topologies for different types of carbohydrate-amino acid linkages and biologically important monosaccharide residues, important advances may expected to occur in comprehending biological processes involving glycoproteins.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] Dwek, R.A. Glycobiology: toward understanding the function of sugars. *Chem. Rev.*, **1996**, *96*, 683-720.

[2] Spiro, R.G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, **2002**, *12*, 43R-56R.

[3] Varki, A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, **1993**, *3*, 97-130.

[4] Sears, P.; Wong, C.-H. Enzyme action in glycoprotein synthesis. *Cell. Mol. Life Sci.*, **1998**, *54*, 223-252.

[5] Skropeta, D. The effect of individual N-glycans on enzyme activity. *Bioorg. Med. Chem.*, **2009**, *17*, 2645-2653.

[6] Tams, J.W.; Vind, J.; Welinder, K.G. Adapting protein solubility by glycosylation. N-Glycosylation mutants of Coprinus cinereus peroxidase in salt and organic solutions. *Biochim. Biophys. Acta*, **1999**, *1432*, 214-221.

[7] Price, N.J.; Pinheiro, C.; Soares, C.M.; Ashford, D.A.; Ricardo, C.P.; Jackson, P.A. A biochemical and molecular characterization of LEP1, an extensin peroxidase from lupin. *J. Biol. Chem.*, **2003**, *278*, 41389-41399.

[8] Wyss, D.F.; Wagner, G. The structural role of sugars in glycoproteins. *Curr. Opin. Biotechnol.*, **1996**, *7*, 409-416.

[9] 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.

[10] 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.

[11] Woods, R.J. Computational carbohydrate chemistry: what theoretical methods can tell us. *Glycoconjug. J.*, **1998**, *15*, 209-216.

[12] 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.

[13] Crispin, M.; Stuart, D.I.; Jones, E.Y. Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.*, **2007**, *14*, 354.

[14] Berman, H.M.; Henrick, K.; Nakamura, H.; Markley, J. Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.*, **2007**, *14*, 354-355.

[15] 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.

[16] Cumming, D.A.; Carver, J.P. Virtual and solution conformations of oligosaccharides. *Biochemistry*, **1987**, *26*, 6664-6676.

[17] Recacha, R.; Costanzo, M.J.; Maryanoff, B.E.; Carson, M.; DeLucas, L.; Chattopadhyay, D. Structure of human alpha-thrombin complexed with

RWJ-51438 at 1.7 A: unusual perturbation of the 60A-60I insertion loop. *Acta Crystallogr. D Biol. Crystallogr.*, **2000**, *56*, 1395-1400.

[18] Complex Carbohydrate Research Center, University of Georgia, Athens, GA. Woods Group (2005-2010). GLYCAM http:// www.glycam.com(Accessed April 28, 2010)

[19] Frank, M.; Bohne-Lang, A.; Wetter, T.; von der Lieth, C.-W. Rapid generation of a representative ensemble of N-glycan conformations. *In Silico Biol.*, **2002**, *2*, 427-439.

[20] Bohne-Lang, A.; von der Lieth, C.-W. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res.*, **2005**, *33*, W214-W219.

[21] Vitkup, D.; Melamud, E.; Moult, J.; Sander, C. Completeness in structural genomics. *Nat. Struct. Biol.*, **2001**, *8*, 559-566.

[22] Baker, D.; Sali, A. Protein structure prediction and structural genomics. *Science*, **2001**, *294*, 93-96.

[23] Cavasotto, C.N.; Phatak, S.S. Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today*, **2009**, *14*, 676-683.

[24] Imberty, A.; Peréz, S. Stereochemistry of the *N*-glycosylation sites in glycoproteins. *Protein Eng.*, **1995**, *8*, 699-709.

[25] 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.

[26] Imberty, A.; Pérez, S. Structure, conformation, and dynamics of bioactive oligosaccharides: theoretical approaches and experimental validations. *Chem. Rev.*, **2000**, *100*, 4567-4588.

[27] Pilobello, K.T.; Mahal, L.K. Deciphering the glycocode: the complexity and analytical challenge of glycomics. *Curr. Opin. Chem. Biol.*, **2007**, *11*, 300-305.

[28] Helenius, A.; Aebi, M. Intracellular functions of N-Linked glycans. *Science*, **2001**, *291*, 2364-2369.

[29] Kukuruzinska, M.A.; Lennon, K. Protein N-Glycosylation: molecular genetics and functional significance. *Crit. Rev. Oral Biol. Med.*, **1998**, *9*, 415-448.

[30] Stenutz, R.; Weintraub, A.; Widmalm, G. The structures of *Escherichia coli* O-polysaccharide antigens. *FEMS Microbiol. Rev.*, **2006**, *30*, 382-403.

[31] Beusen, D.D.; Shands, E.F.B. Systematic search strategies in conformational analysis. *Drug Discov. Today*, **1996**, *1*, 429-437.

[32] Kifli, N; Htar, T.T.; De Clercq, E.; Balzarini, J.; Simons, C. Novel bicyclic sugar modified nucleosides: synthesis, conformational analysis and antiviral evaluation. *Bioorg. Med. Chem.*, **2004**, *12*, 3247-3257.

[33] 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.

[34] Mukhopadhyay, C. Molecular dynamics simulation of glycoprotein-glycans of Immunoglobulin G and Immunoglobulin M. *Biopolymers*, **1998**, *45*, 177-190.

[35] 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.

[36] 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.

[37] Fernandes, C.L.; Sachett, L.G.; Pol-Fachin, L.; Verli, H. GROMOS96 43a1 performance in predicting oligosaccharides conformational ensemble within glycoproteins. *Carbohydr. Res.*, **2010**, *345*, 663-667.

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

[39] Stortz, C.A.; Cerezo, A.S. Disaccharide conformational maps: 3D contours or 2D plots? *Carbohydr. Res.*, **2002**, *337*, 1861-1871.

[40] Pérez, S.; Kouwijzer, M.; Mazeau, K.; Engelsen, S.B. Modeling polysaccharides: Present status and challenges. *J. Mol. Graph.*, **1996**, *14*, 307-321.

[41] Liwo, A.; Khalili, M.; Czaplewski, C.; Kalinowski, S.; Oldziej, S.; Wachucik, K.; Scheraga, H.A. Modification and Optimization of the United-Residue (UNRES) potential energy function for canonical simulations. i. temperature dependence of the effective energy function and tests of the optimization method with single training proteins. *J. Phys. Chem. B*, **2007**, *111*, 260-285.

[42] Monticelli, L.; Kandasamy, S.K.; Periole, X.; Larson, R.G.; Tieleman, D.P.; Marrink, S.J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.*, **2008**, *4*, 819-834.

[43] Lopez, C.A., Rzepiela, A.J., de Vries, A.H., Dijkhuizen, L., Hünenberger, P.H., and Marrink, S.J. Martini coarse-grained force field: extension to carbohydrates. *J. Chem. Theory Comput.*, **2009**, *5*, 3195-3210.

[44] Stortz, C.A. Disaccharide conformational maps: how adiabatic is an adiabatic map? *Carbohydr. Res.*, **1999**, *322*, 77-86.

[45] da Silva, C.O.; Nascimento, M.A. Ab initio conformational maps for disaccharides in gas phase and aqueous solution. *Carbohydr. Res.*, **2004**, *339*, 113-122.

[46] French, A.D.; Kelterer, A.-M.; Johnson, G.P.; Dowd, M.K.; Cramer, C.J. HF/6-31G* energy surfaces for disaccharide analogs. *J. Comput. Chem.*, **2001**, *22*, 65-78.

[47] Schnupf, U.; Willett, J.L.; Bosma, W.B.; Momany, F.A. DFT studies of the disaccharide, α-maltose: relaxed isopotential maps. *Carbohydr. Res.*, **2007**, *342*, 2270-2285.

[48] Yousfi, N.; Sekkal-Rahal, M.; Sayede, A.; Springborg, M. Relaxed energetic maps of κ-Carrabiose: A DFT study. *J. Comput. Chem.*, **2010**, *31*, 1312-1320.

[49] Tvaroška, I.; Pérez, S. Conformational-energy calculations for oligosaccharides: a comparison of methods and a strategy of calculation. *Carbohydr. Res.*, **1986**, *149*, 389-410.

[50] 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.

[51] 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.

[52] Scott, W.R.P.; Hunenberger, P.H.; Tironi, I.G.; Mark, A.E.; Billeter, S.R.; Fennen, J.; Torda, A.E.; Huber, T.; Kruger, P.; van Gunsteren, W.F. The GROMOS biomolecular simulation program. *J. Phys. Chem. A*, **1999**, *103*, 3596-3607.

[53] Verli, H.; Guimarães, J.A. Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. *Carbohydr. Res.*, **2004**, *339*, 281-290.

[54] Allinger, N.L.; Yuh, Y.H.; Lii, J.-H. Molecular mechanics. The MM3 Force Field for Hydrocarbons. 1. *J. Am. Chem. Soc.*, **1989**, *111*, 8551-8582.

[55] 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.

[56] 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.3rd.; 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.

[57] Mackerell, A.D.Jr. Empirical force fields for biological macromolecules: overview and issues. *J. Comput. Chem.*, **2004**, *25*, 1584-1604.

[58] Pérez, S.; Imberty, A.; Engelsen, S.B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J.-F.; van Eyck, B.P.; Johnson, G.; French, A.D.; Kouwijzer, M.L.C.E.; Grootenuis, P.D.J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. A comparison and chemometric analysis of several molecular mechanics force fields and parameter sets applied to carbohydrates. *Carbohydr. Res.*, **1998**, *314*, 141-155.

[59] Hemmingsen, L.; Madsen, D.E.; Esbensen, A.L.; Olsen, L.; Engelsen, S.B. Evaluation of carbohydrate molecular mechanics force fields by quantum mechanical calculations. *Carbohydr. Res.*, **2004**, *339*, 937-948.

[60] Hong, X.; Hopfinger, A.J. Construction, molecular modeling, and simulation of *Mycobacterium tuberculosis* cell walls. *Biomacromolecules*, **2004**, *5*, 1052-1065.

[61] Kirschner, K.N.; Woods, R.J. Solvent interactions determine carbohydrate conformation. *Proc. Natl. Acad. Sci. U. S. A.*, **2001**, *98*, 10541-10545.

[62] Almond, A.; Sheehan, J.K. Predicting the molecular shape of polysaccharides from dynamic interactions with water. *Glycobiology*, **2003**, *13*, 255-264.

[63] Eklund, R.; Widmalm, G. Molecular dynamics simulations of an oligosaccharide using a force field modified for carbohydrates. *Carbohydr. Res.*, **2003**, *338*, 393-398.

[64] González-Outeiriño, J.; Kadirvelraj, R.; Woods, R.J. Structural elucidation of type III group B *Streptococcus* capsular polysaccharide using molecular dynamics simulations: the role of sialic acid. *Carbohydr. Res.*, **2005**, *340*, 1007-1018.

[65] 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. *J. Mol. Struct.*, **2006**, *778*, 97-103.

[66] Woods, R.J.; Pathiaseril, A.; Wormald, M.R.; Edge, C.J.; Dwek, R.A. The high degree of internal flexibility observed for an oligomannose oligosaccharide does not alter the overall topology of the molecule. *Eur. J. Biochem.*, **1998**, *258*, 372-386.

[67] Stanca-Kaposta, E.C.; Gamblin, D.P.; Cocinero, E.J.; Frey, J.; Kroemer, R.T.; Fairbanks, A.J.; Davis, B.G.; Simons, J.P. Solvent interactions and conformational choice in a core n-glycan segment: gas phase conformation of the central, branching trimannose unit and its singly hydrated complex. *J. Am. Chem. Soc.*, **2008**, *130*, 10691-10696.

[68] 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-(β-D-Glycopyranosyl)haloacetamides. *J. Am. Chem. Soc.*, **2008**, *130*, 8317-8325.

[69] Corzana, F.; Busto, J.H.; Jiménez-Osés, G.; Asensio, J.L.; Jiménez-Barbero, J.; Peregrina, J.M.; Avenoza, A. New Insights into α-GalNAc-Ser Motif: influence of hydrogen bonding versus solvent interactions on the preferred conformation. *J. Am. Chem. Soc.*, **2006a**, *128*, 14640-14648.

[70] Corzana, F.; Busto, J.H.; Engelsen, S.B.; Jiménez-Barbero, J.; Asensio, J.L.; Peregrina, J.M.; Avenoza, A. Effect of β-O-Glucosylation on L-Ser and L-Thr Diamides: A Bias toward α-Helical Conformations. *Chemistry*, **2006b**, *12*, 7864-7871.

[71] Corzana, F.; Busto, J.H.; Jiménez-Osés, G.; García de Luis, M.; Asensio, J.L.; Jiménez-Barbero, J.; Peregrina, J.M.; Avenoza, A. Serine versus Threonine Glycosylation: the methyl group causes a drastic alteration on the carbohydrate orientation and on the surrounding water shell. *J. Am. Chem. Soc.*, **2007**, *129*, 9458-9467.

[72] Fernández-Tejada, A.; Corzana, F.; Busto, J.H.; Jiménez-Osés, G.; Jiménez-Barbero, J.; Avenoza, A.; Peregrina, J.M. Insights into the geometrical features underlying β-*o*-glcnac glycosylation: water pockets drastically modulate the interactions between the carbohydrate and the peptide backbone. *Chemistry*, **2009**, *15*, 7297-7301.

[73] Zhu, X.; Borchers, C.; Bienstock, R.J.; Tomer, K.B. Mass spectrometric characterization of the glycosylation pattern of HIV-gp120 expressed in CHO cells. *Biochemistry*, **2000**, *39*, 11194-11204.

[74] 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.

[75] 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.

[76] Perera, L.; Darden, T.A.; Pedersen, L.G. Predicted solution structure of zymogen human coagulation FVII. *J. Comput. Chem.*, **2002**, *23*, 35-47.

[77] 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.

[78] Lyman, E.; Zuckerman, D.M. Ensemble-Based convergence analysis of biomolecular trajectories. *Biophys. J.*, **2006**, *91*, 164-172.

[79] Lu, Z.; Hu, H.; Yang, W.; Marszalek, P.E. Simulating force-induced conformational transitions in polysaccharides with the SMD replica exchange method. *Biophys. J.*, **2006**, *91*, L57-L59.

[80] Kräutler, V.; Müller, M.; Hünenberger, P.H. Conformation, dynamics, solvation and relative stabilities of selected β-hexopyranoses in water: a molecular dyuamics study with the GROMOS 45A4 force field. *Carbohydr. Res.*, **2007**, *342*, 2097-2124.

[81] Pol-Fachin, L.; Verli, H. Depiction the forces participating in the 2-*O*-sulfo-α-L-iduronic acid conformational preference in heparin sequences within aqueous solutions. *Carbohydr. Res.*, **2008**, *343*, 1435-1445.

[82] Landström, J.; Widmalm, G. Glycan flexibility: insights into nanosecond dynamics from a microsecond molecular dynamics simulation explaining an unusual nuclear Overhauser effect. *Carbohydr. Res.*, **2010**, *345*, 330-333.

[83] Zagrovic, B.; Pande, V.S. How does averaging affect protein structure comparison on the ensemble level? *Biophys. J.*, **2004**, *87*, 2240-2246.

[84] van Gunsteren, W.F.; Berendsen, H.J.C. Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.*, **1990**, *29*, 992-1023.

[85] Koller, A.N.; Schwalbe, H.; Gohlke, H. starting structure dependence of NMR order parameters derived from MD simulations: implications for judging force-field quality. *Biophys. J.*, **2008**, *95*, L04-L06.

[86] Brigo, A.; Lee, K.W.; Iurcu Mustata, G.; Briggs, J.M. Comparison of multiple molecular dynamics trajectories calculated for the drug-resistant HIV-1 Integrase T66I/M154I catalytic domain. *Biophys. J.*, **2005**, *88*, 3072-3082.

[87] 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.

[88] 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.

[89] Erbel, P.J.A.; Karimi-Nejad, Y.; van Kuik, J.A.; Boelens, R.; Kamerling, J.P.; Vliegenthart, J.F.G. Effects of the N-Linked glycans on the 3D structure of the free r-subunit of human chorionic gonadotropin. *Biochemistry*, **2000**, *39*, 6012-6021.

[90] Reeves, R.E. The shape of pyranose rings. *J. Am. Chem. Soc.*, **1950**, *72*, 1499-1506.

[91] 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.

[92] 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.

[93] 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.

[94] 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.

[95] Bailey, D.; Renouf, D.V.; Large, D.G.; Warren, C.D.; Hounsell, E.F. Conformational studies of the glycopeptide Ac-Tyr-[Man$_9$GlcNAc-β-(1→4)GlcNAc-β-(1→Nδ)]-Asn-Leu-Thr-Ser-OBz and the constituent peptide and oligosaccharide. *Carbohydr. Res.*, **2000**, *324*, 242-254.

[96] Lommerse, J.P.; Kroon-Batenburg, L.M.; Kamerling, J.P.; Vliegenthart, J.F. Conformational analysis of the Xylose-containing N-Glycan of pineapple stem bromelain as part of the intact glycoprotein. *Biochemistry*, **1995**, *34*, 8196-8206.

[97] 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.

[98] 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.

[99] 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.

[100] Verli, H.; Calazans, A.; Brindeiro, R.; Tanuri, A.; Guimarães, J.A. Molecular dynamics analysis of HIV-1 matrix protein: clarifying differences between crystallographic and solution structures. *J. Mol. Graph. Model.*, **2007**, *26*, 62-68.

[101] Lima, L.M.; Becker, C.F.; Giesel, G.M.; Marques, A.F.; Cargnelutti, M.T.; de Oliveira Neto, M.; Monteiro, R.Q.; Verli, H.; Polikarpov, I. Structural and thermodynamic analysis of thrombin:suramin interaction in solution and crystal phases. *Biochim. Biophys. Acta*, **2009**, *1794*, 873-881.

[102] Price, D.J.; Brooks III, C.L. Modern protein force fields behave comparably in molecular dynamics simulations. *J. Comput. Chem.*, **2002**, *23*, 1045-1057.

[103] Petersen, B.O.; Vinogradov, E.; Kay, W.; Würtz, P.; Nyberg, N.T.; Duus, J.Ø.; Sørensen, O.W. H2BC: a new technique for NMR analysis of complex carbohydrates. *Carbohydr. Res.*, **2006**, *341*, 550-556.

[104] Church, T.; Carmichael, I.; Serianni, A.S. Two-bond ^{13}C-^{13}C spin-coupling constants in carbohydrates: effect of structure on coupling magnitude and sign. *Carbohydr. Res.*, **1996**, *280*, 177-186.

[105] Milton, M.J.; Harris, M.A.; Probert, M.A.; Field, R.A.; Homans, S.W. New conformational constraints in isotopically (^{13}C) enriched oligosaccharides. *Glycobiology*, **1998**, *8*, 147-153.

[106] Bose, B.; Zhao, S.; Stenutz, R.; Cloran, F.; Bondo, F.; Bondo, G.; Hertz, B.; Carmichael, I.; Serianni, A. Three-Bond C−O−C−C spin-coupling constants in carbohydrates: development of a karplus relationship. *J. Am. Chem. Soc.*, **1998**, *120*, 11158-11173.

[107] Haasnoot, C.A.G. Conformational analysis of six-membered rings in solution: ring puckering coordinates derived from vicinal NMR proton-proton coupling constants. *J. Am. Chem. Soc.*, **1993**, *115*, 1460-1468.

[108] Martin-Pastor, M.; Bush, C.A. New strategy for the conformational analysis of carbohydrates based on NOE and ^{13}C NMR coupling constants. Application to the flexible polysaccharide of *Streptococcus mitis* J22. *Biochemistry*, **1999**, *38*, 8045-8055.

[109] Cremer, D.; Pople, J.A. A General definition of ring puckering coordinates. *J. Am. Chem. Soc.*, **1975**, *97*, 1354-1358.

[110] Apweiler, R.; Hermjacob, H.; Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta*, **1999**, *1473*, 4-8.

[111] Ben-Dor, S.; Esterman, N.; Rubin, E.; Sharon, N. Biases and complex patterns in the residues flanking protein N-glycosylation sites. *Glycobiology*, **2004**, *14*, 95-101.

[112] Gray, J.J.; Moughon, S.E.; Kortemme, T.; Schueler-Furman, O.; Misura, K.M.; Morozov, A.V.; Baker, D. Protein-protein docking predictions for the CAPRI experiment. *Proteins*, **2003**, *52*, 118-122.

[113] Kerzmann, A.; Neumann, D.; Kohlbacher, O. SLICK - scoring and energy functions for protein-carbohydrate interactions. *J. Chem. Inf. Model.*, **2006**, *46*, 1635-1642.

[114] Agostino, M.; Jene, C.; Boyle, T.; Ramsland, P.A.; Yuriev, E. Molecular docking of carbohydrate ligands to antibodies: structural validation against crystal structures. *J. Chem. Inf. Model.*, **2009**, *49*, 2749-2760.

[115] Camacho, C.J.; Vajda, S. Protein-protein association kinetics and protein docking. *Curr. Opin. Struct. Biol.*, **2002**, *12*, 36-40.

[116] Englebienne, P.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 4. are popular scoring functions accurate for this class of proteins? *J. Chem. Inf. Model.*, **2009**, *49*, 1568-1580.

[117] Gohlke, H.; Kiel, C.; Case, D.A. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.*, **2003**, *330*, 891-913.

[118] Hill, A.D.; Reilly, P.J. A Gibbs Free energy correlation for automated docking of carbohydrates. *J. Comput. Chem.*, **2008**, *29*, 1131-1141.

[119] Gandhi, N.S.; Mancera, R.L. Free energy calculations of glycosaminoglycan–protein interactions. *Glycobiology*, **2009**, *19*, 1103-1115.

[120] Green, D.F. Optimized parameters for continuum solvation calculations with carbohydrates. *J. Phys. Chem. B*, **2008**, *112*, 5238-5249.

[121] Guimarães, C.R.W.; Mathiowetz, A.M. Addressing limitations with the MM-GB/SA scoring procedure using the watermap method and free energy perturbation calculations. *J. Chem. Inf. Model.*, **2010**, *50*, 547-559.

[122] Zwanzig, R.W. High temperature equation of state by a perturbation method. i. nonpolar gases. *J. Chem. Phys.*, **1954**, *22*, 1420-1426.

[123] Mezei, M. The finite difference thermodynamic integration, tested on calculating the hydration free energy difference between acetone and dimethylamine in water. *J. Chem. Phys.*, **1987**, *86*, 7084-7088.

[124] Åqvist, J.; Medina, C.; Samuelsson, J.E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.*, **1994**, *7*, 385-391.

[125] Åqvist, J.; Luzhkov, V.B.; Brandsdal, B.O. Ligand binding affinities from MD simulations. *Acc. Chem. Res.*, **2002**, *35*, 358-365.

Received: March 23, 2010 Revised: May 03, 2010 Accepted: May 12, 2010

4.3 Trabalho II

Este trabalho descreve o comportamento conformacional de diferentes conformações e glicoformas de AT, quando complexadas e não-complexadas à heparina, visando ao melhor entendimento do processo de ativação da serpina. Até então, 21 estruturas cristalográficas de AT humana estavam disponíveis no PDB, para suas formas nativa, intermediária e ativada. No entanto, a estrutura sacarídica completa de AT não foi observada nessas estruturas, e o efeito da glicosilação sobre a conformação e dinâmica de ativação da serpina por heparina havia sido apenas inferida a partir de estudos bioquímicos. Nesse sentido, foram realizados estudos de simulação computacional para as duas glicoformas circulantes de AT no plasma, α e β, em comparação com sua estrutura não glicosilada, a partir das conformações nativa e ativada, a fim de avaliar, a nível atômico, os efeitos da glicosilação sobre AT. Adicionalmente, a forma ativada foi estudada complexada à heparina, a fim de avaliar os efeitos da interação do polissacarídeo sobre a estrutura da serpina.

Os resultados obtidos indicam uma interferência da glicana ligada à Asn135 em α-AT na interação desta glicoforma com heparina, causando um dobramento na cadeia do GAG e uma diminuição na energia de interação proteína-carboidrato, o que pode estar associado à menor afinidade do polissacarídeo por tal glicoforma, tal como observado experimentalmente. A partir das simulações, também foi possível propor um modelo da via de transmissão conformacional desde a interação de heparina com as α-hélices A, D e P até o conjunto de β-fitas A, levando à sua ativação conformacional.

Effects of glycosylation on heparin binding and antithrombin activation by heparin

Laércio Pol-Fachin, Camila Franco Becker, Jorge Almeida Guimarães, Hugo Verli

Proteins: Structure, Function and Bioinformatics, **2011**, *79*; 2735-2745

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Effects of glycosylation on heparin binding and antithrombin activation by heparin

Laercio Pol-Fachin,[1] Camila Franco Becker,[1] Jorge Almeida Guimarães,[1] and Hugo Verli[1,2]*

[1] Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil

[2] Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Av Ipiranga 2752, Porto Alegre 90610-000, RS, Brazil

ABSTRACT

Antithrombin (AT), a serine protease inhibitor, circulates in blood in two major isoforms, α and β, which differ in their amount of glycosylation and affinity for heparin. After binding to this glycosaminoglycan, the native AT conformation, relatively inactive as a protease inhibitor, is converted to an activated form. In this process, β-AT presents the higher affinity for heparin, being suggested as the major AT glycoform inhibitor *in vivo*. However, either the molecular basis demonstrating the differences in heparin binding to both AT isoforms or the mechanism of its conformational activation are not fully understood. Thus, the present work evaluated the effects of glycosylation and heparin binding on AT structure, function, and dynamics. Based on the obtained data, besides the native and activated forms of AT, an intermediate state, previously proposed to exist between such conformations, was also spontaneously observed in solution. Additionally, Asn135-linked oligosaccharide caused a bending in AT-bounded heparin, moving such polysaccharide away from helix D, which supports its reduced affinity for α-AT. The obtained data supported the proposal of an atomic-level, solvent and amino acid residues accounting, putative model for the transmission of the conformational signal from heparin binding exosite to β-sheet A and the reactive center loop, also supporting the identification of differences in such transmission between the serpin glycoforms involving helix D, where the Asn135-linked oligosaccharide stands. Such intramolecular rearrangements, together with heparin dynamics over AT surface, may support an atomic-level explanation for the Asn135-linked glycan influence over heparin binding and AT activation.

Proteins 2011; 79:2735–2745.

Key words: glycan; glycosaminoglycan; intermediate; serpin; transmission pathway.

INTRODUCTION

Antithrombin III (AT) is a 432-amino acid residue plasma glycoprotein that plays a major role in the regulation of blood coagulation. AT belongs to the serine protease inhibitors (serpin) family, being involved in the inactivation of several coagulation factors, although its main physiological targets are thrombin, factor IXa, and factor Xa.[1] AT mechanism of action includes the formation of highly stable equimolar complexes with such enzymes through a trapping mechanism,[2–4] in which a covalent stable serpin-protease acyl-intermediate is generated and, after a series of conformational changes, the enzymes are possibly inactivated by structural deformation.[3] In contrast to most serpins, which rapidly inhibit their target enzymes, AT kinetic inhibitory rates are only moderate. However, such inactivating reactions are greatly accelerated by sulfated glycosaminoglycans (GAG), such as therapeutic heparin or vascular wall heparan sulfate proteoglycans,[5] upon AT binding to a minimum sequence-specific pentasaccharide site. Such GAG-AT interaction triggers the above-mentioned conformational changes in the serpin, resulting in its activation and, consequently, boosting its protease inhibitory action.[6]

AT can be found circulating in blood in two major isoforms, α and β, which differ in their amount of glycosylation.[7,8] The predominant glycoform, α-AT, which represents ~90% of the inhibitor levels in plasma, is fully glycosylated (at Asn96, Asn135, Asn155, and Asn192), whereas the minor glycoform, β-AT, lacks the oligosaccharide chain at Asn135.[7,9] The partial glycosylation at this site is thought to be derived from the presence of Ser rather than Thr residue in the third position of the Asn-X-Thr/Ser recognition sequence for glycan attachment.[10] Both AT isoforms are *N*-glycosylated with identical oligosaccharides, consisting of complex type, biantennary glycan structures with two terminal sialic acids,[11]

This work was perfomed at Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul.
Additional Supporting Information may be found in the online version of this article.
Grant sponsor: Conselho Nacional de Desenvolvimento Científico e Tecnológico, MCT; Grant number: CNPq#472174/2007-0; Grant sponsor: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), MEC, Brasília, DF, Brazil
*Correspondence to: Hugo Verli, Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil.
E-mail: hverli@cbiot.ufrgs.br or http://www.cbiot.ufrgs.br/bioinfo.
Received 14 March 2010; Revised 31 May 2011; Accepted 7 June 2011
Published online 16 June 2011 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.23102

although minor variations have been also identified.[12,13] In addition, such AT isoforms can be distinguished based on differences in their binding strength to heparin.[7,8,14] Of the two, β-AT exhibits a higher affinity for such GAG, being suggested as the major AT inhibitory form *in vivo*.[15,16]

Structural data, mainly derived from X-ray crystallography, have been crucial as part of the efforts to understand how GAGs activate this serpin. Namely, several structures of uncomplexed and oligosaccharide-complexed AT have supported studies on the conformational changes responsible for its allosteric activation.[17–19] However, the determination, at the atomic level, of the structural determinants for the differences between heparin binding to either of the two AT glycoforms is hindered by the high flexibility of their carbohydrate moieties, and the consequent difficulties for their crystallization.[20] As an alternative methodology, molecular dynamics (MD) simulations have been employed in the study of glycoproteins, effectively reproducing the conformation and flexibility of both their protein and carbohydrate moieties (as an example, see Refs. 21–23). This method has also been applied to evaluate the interaction between an activated, nonglycosylated AT and a synthetic pentasaccharide,[24] supporting new insights into AT-heparin mutual recognition.

In this context, the present work intends to characterize the effects of glycosylation on heparin binding and AT conformational activation in the presence of heparin through a series of MD simulations. For this purpose, AT was evaluated in its nonglycosylated and in its two circulating glycoforms, considering both native and activated conformational states, thus including in the computational model fundamental structural variables to the understanding of AT biological properties. Such data were accordingly compared with previous X-ray crystallography, biochemical and mutagenesis analysis, supporting the proposal of an atom-level, dynamical model, for the effects of glycosylation on AT flexibility and activity.

METHODS

Nomenclature and software

The recommendations and symbols of nomenclature as proposed by IUPAC[25] are used. The relative orientation of a pair of contiguous carbohydrate residues is described 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 ϕ and ψ are defined as shown in Eqs. (1) and (2):

$$\phi = O5—C1—OX—CX \quad (1)$$

$$\psi = C1—OX—CX—C(X-1) \quad (2)$$

For a (2→6) linkage, the Φ and Ψ are defined as shown in Eqs. (3) and (4):

$$\phi = O6—C2—O6—C6 \quad (3)$$

$$\psi = C2—O6—C6—C5 \quad (4)$$

Finally, for (1→6) and (2→6) linkages, the ω is defined as shown below:

$$\omega = O6—C6—C5—C4 \quad (5)$$

All saccharide topologies were generated with the PRODRG server,[26] the manipulation of structures was performed with MOLDEN,[27] VMD,[28] and PyMOL,[29] the homology modeling was performed with the Swiss-PDB Viewer,[30] and all the MD calculations and analyses were performed using GROMACS simulation suite, version 3.3.3,[31] and GROMOS96 43a1 force field.[32]

Carbohydrates topology construction

The monosaccharide fragments composing AT glycan chains and heparin polysaccharide fragment were constructed using MOLDEN software.[27] Each structure was then submitted to the PRODRG server,[26] from which the crude topologies were retrieved. Based on such information, these structures were described in GROMOS96 43a1 force field parameters, being further modified to include some refinements : (1) improper dihedrals, employed to preserve the conformational state of the hexopyranose rings in $^4C_1$ (D-GlcNAc, D-Man, D-Gal, and D-GlcNS,6S), ^2S_O (L-IdoA2S), or $^2C_5$ (NeuAc) forms; (2) proper dihedrals, as described in GROMOS96 43a1 force field for glucose, in order to support stable simulations,[23] and (3) Löwdin HF/6-31G** derived atomic charges, obtained from previous works of the group.[23,33,34]

Building of AT glycoforms

The native AT form was retrieved from PDB ID 2ANT,[17] comprising the complete protein structure with a single gap between residues 29 and 43. Otherwise, activated, heparin-bound AT structure was retrieved from PDB ID 1E03,[35] showing a gap between amino acid residues 28 and 37, five missing residues in its *N*-terminal (from residues 1–5) and the last lysine (residue 432) in its C-terminal. In both structures, such gaps were fulfilled by means of homology modeling using the Swiss-PDB Viewer program.[30] In order to generate initial models for both AT glycoforms, the *N*-glycosylation sites (N096, N155, and N192 for β-AT, added by N135 for α-AT) in both native and activated structures were filled with AT proposed oligosaccharide moiety (complex type, biantennary glycan structure, with two terminal sialic acids),[11] using glycosciences modeling tools.[36] Such models had

their glycosidic linkage geometries adjusted to the main conformational states for each linkage, based on their relative abundance in the isolated disaccharides in water, as previously described[37] and as calculated in the present work (Supporting Information, Fig. S1A, S1B). Regarding the heparin polysaccharide fragment, a heparin structure containing 21 carbohydrate residues (Supporting Information, Fig. S1E) was also built based on the most prevalent geometries obtained from solution MD simulations of heparin composing disaccharides (Supporting Information, Figure S1C, S1D) and as previously determined,[38] considering the specific pentasaccharide sequence for heparin-AT interaction and octasaccharide activating fragments[39] for determining its sequence. For this purpose, heparin was constructed in its most abundant commercial form, thus lacking 3-*O*-sulfation in the specific pentasaccharide motif,[40] which has been described as not required for AT allosteric activation mechanism.[41] Subsequently, such heparin molecule was complexed to AT through superimposition to the synthetic pentasaccharide observed in PDB ID 1E03 (Supporting Information, Fig. S1E–S1G).

MD simulations

Six uncomplexed AT structures were generated: three comprising native AT and three representing activated, unbounded AT. In each set of AT conformational states, the nonglycosylated, α- and β-glycoforms were considered. Additionally, three AT-heparin complexes were studied, considering the protein in its activated form and a heparin fragment with 21 carbohydrate residues. Accordingly, AT was simulated in its nonglycosylated, α- and β-glycosylated forms. Globally, nine AT systems were simulated, considering variations in its complexation to heparin, glycosylation, and conformational state. Such structures were solvated in triclinic boxes using periodic boundary conditions and SPC water model.[42] Counter ions (Na^+) were added to neutralize the system. The employed MD protocol was based on previous studies[43] as described.[23,33,34] The Lincs method[44] 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 with Particle Mesh Ewald method.[45] Temperature and pressure were kept constant by coupling (glyco)-proteins, heparin, ions, and solvent to external temperature and pressure baths with coupling constants of τ = 0.1 and 0.5 ps,[46] respectively. The dielectric constant was treated as ε = 1. The systems were heated slowly from 50 to 310 K, in steps of 5 ps, each one increasing the reference temperature by 50 K. After this heating, all simulations were further extended to 0.1 μs under a constant temperature of 310 K. In the case of activated AT-heparin complexes, before data collection, 10 ns MD simulations were employed as equilibration steps, in order to avoid protein structure bending and heparin escaping, in which an initial 5000 kJ mol^{-1} constant force was applied into the glycoprotein and heparin, being reduced by 1000 kJ mol^{-1} at every 2 ns, allowing water molecules and counter-ions to settle around the complexes.

RESULTS AND DISCUSSION

AT glycoforms overall structure and dynamics

In order to evaluate the effects of glycosylation over AT structure and dynamics, and to distinguish the effects of glycosylation from those of heparin binding, nine systems were simulated: (1) noncomplexed and nonglycosylated native AT; (2) noncomplexed, native α-AT; (2) noncomplexed, native β-AT; (4) noncomplexed and nonglycosylated activated AT; (5) noncomplexed, activated α-AT; (6) noncomplexed, activated β-AT; (7) heparin complexed and nonglycosylated activated AT; (8) heparin complexed, activated α-AT; and (9) heparin complexed, activated β-AT. The AT nonglycosylated structures were simulated as a control of potential glycosylation conformational effects.

The two glycosylated structures represent the two major AT glycoforms found circulating in human blood, namely, α and β. These *N*-glycans containing systems were filled with the proposed canonical biantennary complex-type oligosaccharide structure,[11–13,35] as presented in Figure 1(A). The conformational profiles of the glycan chains attached to AT were evaluated by means of their composing glycosidic linkages, being further compared with previous nuclear magnetic resonance (NMR) data.[47] As a general feature, the behavior obtained from MD simulations for the glycosidic linkages composing AT glycan chains are within the experimentally determined geometries, as presented in Supporting Information (Table S1), indicating an adequate description of their conformational ensemble. Furthermore, in the heparin containing systems, the GAG glycosidic linkage geometries are also within solution NMR[48] and solid state[49] determined conformations (Supporting Information, Table S2).

For all the simulated systems, most of the initial secondary structure content on the studied AT isoforms was retained during the MD simulations (Supporting Information, Fig. S2), maintaining the typical serpin tertiary structure. The major differences between the native and activated forms of the protein, observed both in crystallographic and MD derived structures, are different helix D extensions: up to Arg129 in the native state, and mostly from Gln118 to Asn135 in the activated conformation, together with the occurrence of helix P in the former conformational state [Fig. 1(B)]. In addition, during the three native AT glycoforms MD simulations, a ~30% helical prevalence may be observed from Ala134 to Ser137 [Fig. 1(B), and Supporting Information, Fig. S3]. Such propensity to adopt an helical conforma-

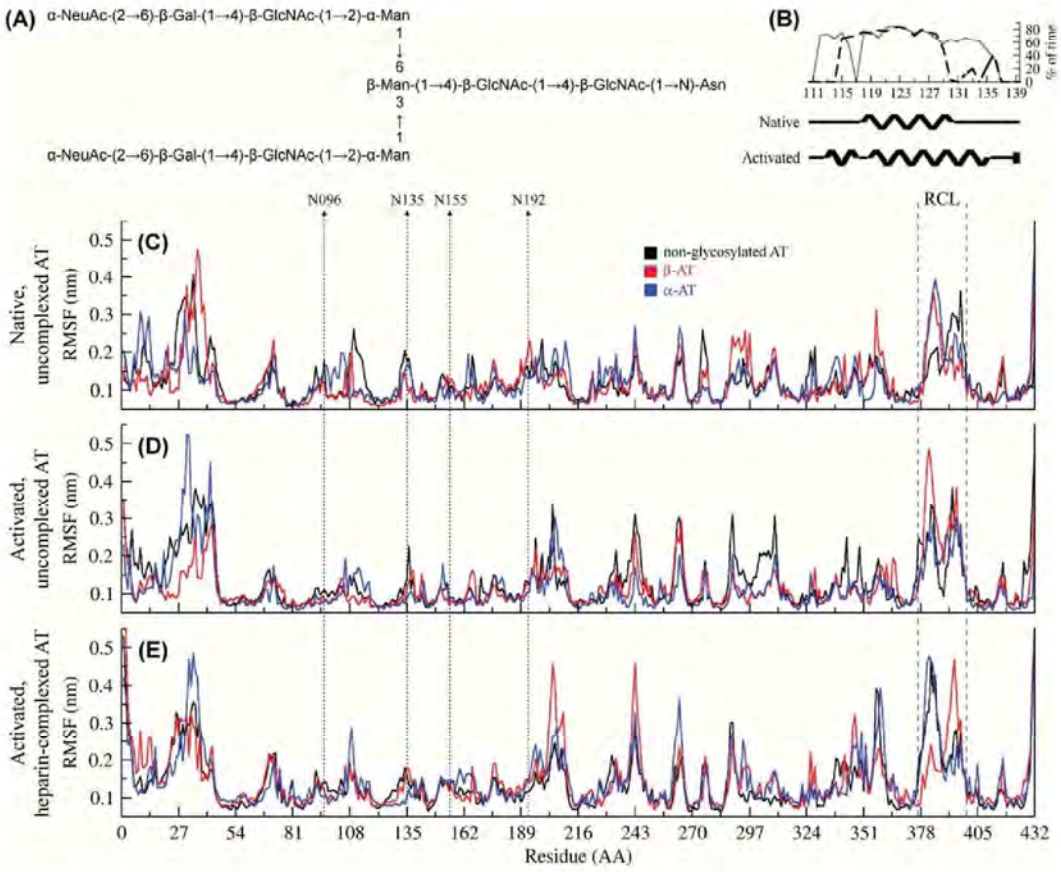


Figure 1

Native and activated AT isoforms overall flexibility. In (**A**), schematics of the biantennary, complex-type oligosaccharide used to construct AT glycosylated isoforms. In (**B**), the helix D content during 0.1 μs MD simulations of native (black dashed) and activated (gray straight) AT isoforms, in comparison to their X-ray structure (in the graph, a representative data, from nonglycosylated trajectories, are used). In addition, the root mean square fluctuation (RMSF) analysis for the (**C**) native, (**D**) activated but uncomplexed, and (**E**) activated, heparin-bound AT is shown for its nonglycosylated (black), β- (red) and α- (blue) glycoforms. Crossing the graphs, dashed arrows indicate AT *N*-glycosylation sites. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

tion may be related to previous data pointing that the deletion of these amino acids largely decouple conformational changes in the heparin-binding site for AT conformational activation.[50]

Regarding AT plasticity, the two most flexible regions in the polypeptide chain, independently on its adopted conformational state, include: (1) the *N*-terminal region, mainly in the loop between Arg24 and Pro41, which is usually not observed in serpins crystallographic structures, and (2) the reactive center loop (RCL), between residues Glu377 to Leu400 [Fig. 1(C–E)]. The flexibility behavior observed for the native and activated AT states, in its α-, β-, and nonglycosylated forms are in agreement with a B-factor profile obtained through averaging their per-residue values in a series of PDB entries (Supporting Information, Fig. S4). Still, stiff regions, described as presenting invariable structures during the heparin induced conformational change by previous works comparing several native and activated AT X-ray structures,[19] remained mostly stable when comparing their behavior during native and activated MD simulations. Specifically, such regions comprise residues 54–81 (mostly corresponding to helix A), 82–108 (comprising helixes B and C), 327-352 (containing helix I), and the C-terminal amino acids 400–431 [Fig. 1(C–E)]. In addition, the so described plastic fragments[19] appeared to behave randomly when varying AT conformation and glycoform, including the loops between residues 109–118 and 352–363 [Fig. 1(C–E)].

Glycosylation also appears to play a role in AT flexibility in solution. Hence, during the native AT MD simulations, the *N*-terminal segment of the RCL, also known as the hinge region, presented an increased flexibility in the glycosylated isoforms, although such region still appeared to be partially inserted into β-sheet A, indicating that glycosylation may contribute in the dynamics of such region. Minor plasticity differences may be also observed

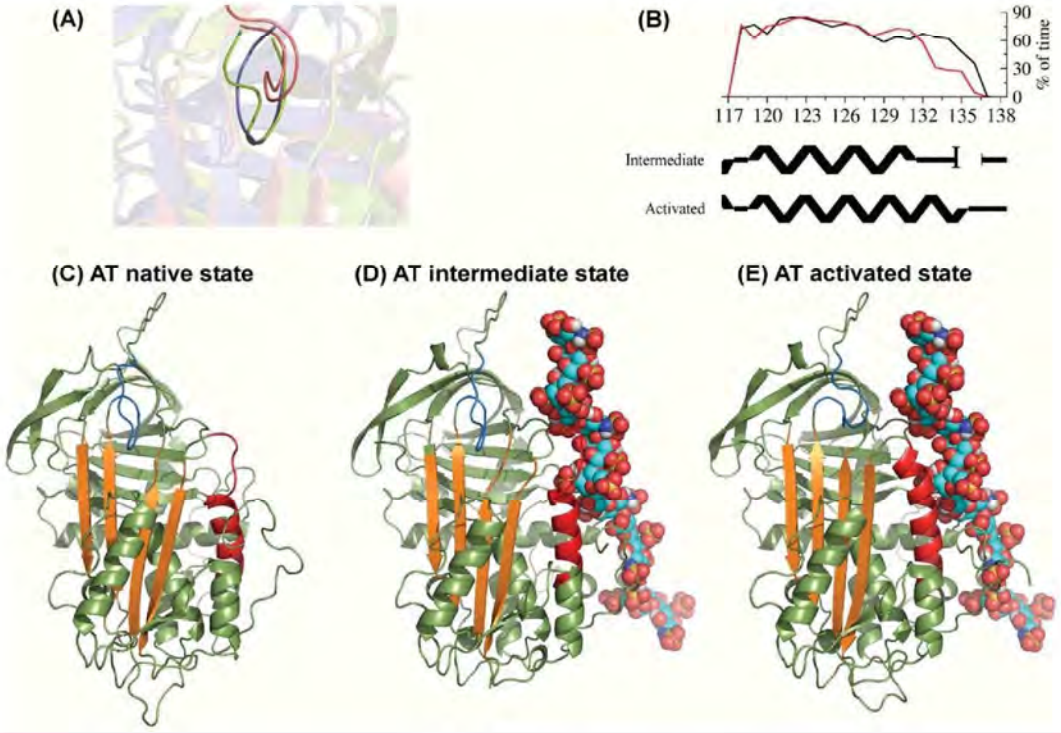


Figure 2

AT conformational equilibrium from native to activated states. (**A**) Hinge region partially insertion into β-sheet A during heparin-bound β-AT MD simulations (blue), in comparison with a previous AT intermediate state crystal structure (green) and to AT activated state crystal structure (red). In (**B**), helix D extension during AT intermediate state (red curve, from β-AT-heparin complex system) and activated (black curve, from nonglycosylated AT-heparin complex system) MD simulations, when compared with its extent in PDB ID 3EVJ (intermediate PROCHECK analysis) and PDB ID 1E03 (activated PROCHECK analysis). As well, the (**C**) native, (**D**) intermediate, and (**E**) activated AT conformational states, proposed to occur in solution, are presented. While heparin is presented as spheres, the AT isoforms are shown as ribbon diagrams in the classic orientation, with β-sheet A (orange) facing and RCL in top, with hinge region shown in blue and helix D in red. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

between the studied glycoforms. For instance, in native AT, the region comprising helix G (292–298) presented a high flexibility in β-AT when compared with α-AT, while the helix C composing residues (96–107) showed an increased plasticity in α-AT in comparison to β-AT [Fig. 1(C)]. In the activated state, β-AT presented an increased flexibility in the loops between helix F and strand A3 (residues 202–208) and between strands C3 and C4 (residues 242–244), while activated α-AT showed this behavior in the loop between strand B1 and strand B2 (residues 264–267) [Fig. 1(E)]. In relation to the effects of heparin binding over the polypeptide chain, it appears that AT global flexibility is slightly increased when bounded to the GAG, independently on its glycosylation pattern [Fig. 1(D), in comparison with Fig. 1(E)].

AT intermediate state

Previous X-ray crystallography studies have observed a heparin-complexed AT conformation in which most of the heparin-induced conformational changes have already occurred, except for the C-terminal extension of helix D and hinge-region expulsion from β-sheet A.[51–53] Based on these data, it has been suggested that such AT conformation would represent an intermediate state adopted after binding to heparin, but before reaching its activated conformation.[51] In this context, the simulation of heparin-complexed β-AT MD simulations, started from its activated form, presented a spontaneous conversion to the intermediate state of AT. Through this conformational path, the hinge region was re-inserted into β-sheet A [Fig. 2(A)], and helix D was prevalently unfolded after Arg132 [Fig. 2(B)], equivalently to previous AT intermediate state crystal structures.[51–53] On the other hand, none of the studied glycoforms returned to the native state during the AT activated, heparin-uncomplexed MD simulations, as microsecond or longer time scales may be required for such an amount of conformational changes to be observed.[54]

To our knowledge, this is the first evidence for occurrence of this conformational state in solution conditions, without the crystal environment forces, with an atomic-level

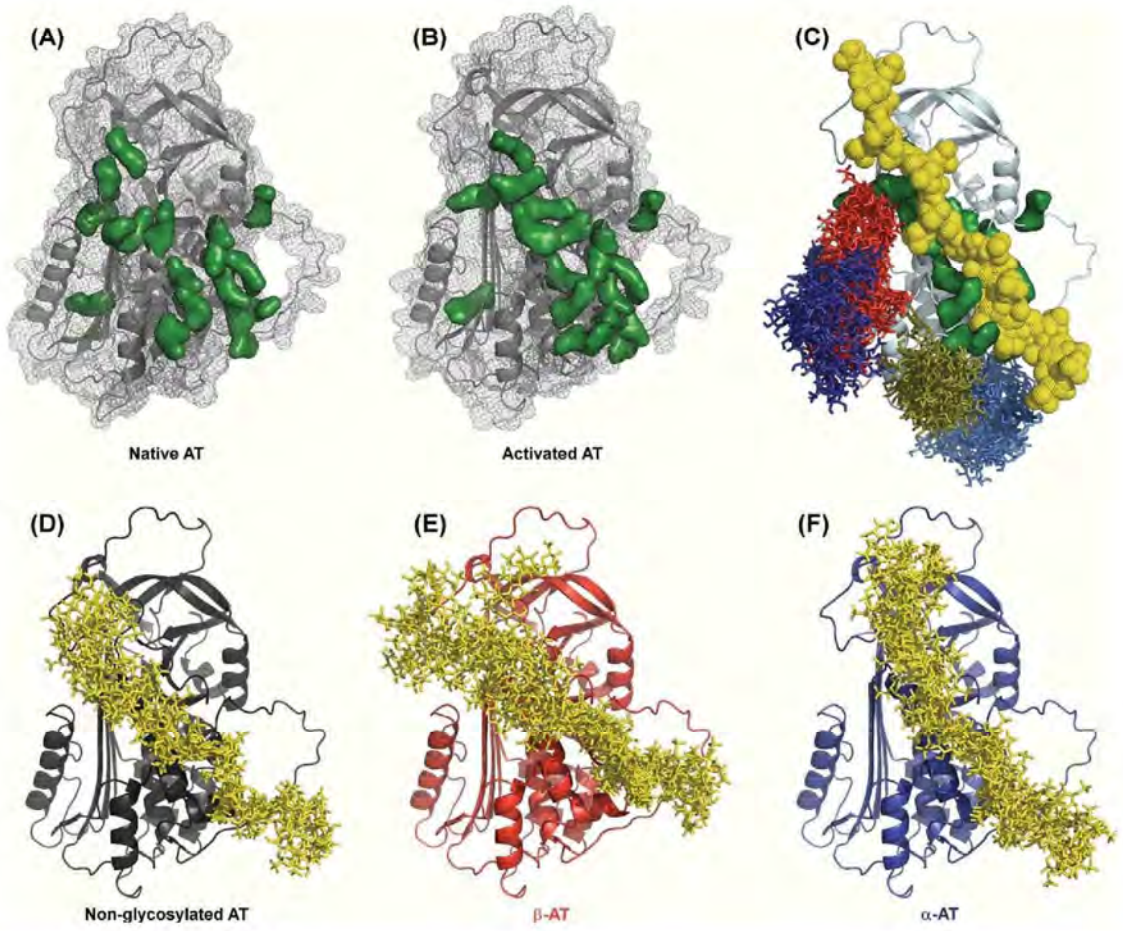


Figure 3

Heparin binding to AT isoforms. In (**A**) and (**B**), the protein core is shown in ribbon and mesh diagrams, while representative frames of AT glycan chains retrieved at every 10 ns of MD simulations are presented as sticks in (**C**) and the residues described as important for heparin binding in green surface in (A–C). As well, heparin conformational profile on the surface of (**D**) nonglycosylated, (**E**) β-AT, and (**F**) α-AT is showed, in which the protein is presented in ribbon diagram and heparin frames obtained at every 10 ns of MD simulations as yellow sticks (the first five residues of heparin chain were omitted for clarity). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

resolution. Such conformation, as previously suggested,[4,51] may fit the two-step, three-state process demonstrated by previous stopped-flow kinetic studies,[55] and is also in accordance to recent findings indicating that the activating conformational changes may occur in two stages.[56] Through this mechanism, after binding to heparin, a series of conformational changes would occur on AT native conformation [Fig. 2(C), first state] leading, at first, to intermediate AT [Fig. 2(D), second state] and, finally, to its activated form [Fig. 2(E), third state].

Heparin binding

Although it has been proposed that the initial and weak heparin binding to AT (presumably, to its native form) occurs equivalently on α- and β-AT,[14] the conformational changes on the serpin are recognized to increase heparin affinity for both forms of the protein.[4,55] In order to derive further structural evidences on heparin binding to the native and activated AT conformations, the amino acid residues proposed to participate in this interaction were selected, based on observations in crystal structures containing AT-heparin complexes and on studies comprising chemically modified or mutated Ats,[18,57–62] and highlighted in Figure 3(A–C). While in AT activated conformation [Fig. 3(B)] most of the AT heparin binding site residues appear to line a channel on the protein surface, as previously proposed,[57] in its native form [Fig. 3(A)] such amino acids seem to be more diffusely organized. In fact, most of the differences between such heparin binding site on native and activated AT (Supporting Information, Table S3) are related to amino acids located in loops (Lys11 and Arg13), involved in secondary structure formation during AT

Table I
Enthalpic Contribution for the Interaction between Heparin and AT in its Nonglycosylated, β- and α-Isoforms

	Average enthalpic contribution for interaction (kJ/mol)[a]		
Residues	Nonglycosylated AT	β-AT	α-AT
Lys11	−80 ± 24	−12 ± 26	−64 ± 25
Arg13	−96 ± 33	−39 ± 41	−124 ± 54
Arg24	0	−1 ± 9	−3 ± 11
Lys39	0	0	0
Thr44	−43 ± 18	−62 ± 20	−64 ± 26
Asn45	−82 ± 13	−54 ± 19	−67 ± 20
Arg46	−86 ± 36	−67 ± 34	−32 ± 32
Arg47	−106 ± 22	−113 ± 28	−60 ± 27
Glu113	−31 ± 10	−31 ± 11	−9 ± 15
Lys114	−107 ± 26	−127 ± 44	−88 ± 29
Lys125	−42 ± 32	−20 ± 27	−31 ± 32
Arg129	−64 ± 24	−61 ± 19	−65 ± 16
Arg132	−72 ± 51	−75 ± 29	−59 ± 35
Lys133	−30 ± 42	−32 ± 28	−19 ± 29
Lys136	−105 ± 43	−115 ± 44	−91 ± 37
Lys139	−39 ± 39	−35 ± 34	−6 ± 16
Arg145	0	0	0
Lys275	−44 ± 32	−106 ± 45	−86 ± 31
TOTAL	−1028 ± 138	−948 ± 117	−868 ± 120

[a]The entire employed timescale (0.1 μs) was considered for obtaining the average and standard deviation values.

conformational changes, that is, placed in helixes P (Glu113 and Lys114) and D (Arg129, Arg132, Lys133, and Lys136) or concerned in secondary structure rearrangement during AT activation, as in helix A (Thr44, Asn45, Arg46, and Arg47). Therefore, based on X-ray crystallographic and MD simulations data (Supporting Information, Table S3), it appears that the AT conformational changes, in order to increase AT inhibitory activity, pass through a distinct alignment of the residues responsible for heparin binding.

It may be observed, based on the glycans flexibility obtained from MD simulations, that the heparin binding site residues are apparently outlined by AT-linked oligosaccharide chains [Fig. 3(C)]. In addition, the Asn135- and Asn155-linked oligosaccharides have been suggested to have a major role on heparin affinity[8,14,63,64] without, however, further atomic insights. Accordingly, in their conformational fluctuation, they may stand near (and even overlap) heparin binding site residues [Fig. 3(C) and Supporting Information, Table S4]. Regarding Asn155 glycan chain, it lies near to some heparin binding site residues at the *N*-terminal segment of AT, but mostly Lys11 and Arg13 (Supporting Information, Table S4), possibly interfering in heparin binding to both circulating AT glycoforms, in relation to the nonglycosylated state.[65] Accordingly, previous data suggested that modifications in the oligosaccharide structures linked to Asn155 of AT, as fucosylation in the *N*-glycan core pentasaccharide,[63,64] may decrease heparin accessibility to its binding channel.

Concerning the Asn135-linked oligosaccharide, its presence only in the α-AT affects the heparin affinity for such glycoform,[7,8,14] such that β-AT has been shown to present an approximately 13-fold increased affinity for such GAG when compared with α-AT.[8] Accordingly, the total enthalpic contribution for the interaction between heparin and its binding site residues, as observed from nonglycosylated, β- and α-AT MD simulations (Table I), appear to be qualitatively correlated to their experimentally determined affinity for heparin.[7,8,14,65] Thus, the nonglycosylated isoform presents the higher interaction energy (-1028 ± 138 kJ mol^{-1}), consistent to previous results indicating that a decreased carbohydrate content in the serpin results in an increase of AT affinity for heparin.[65] As well, β-AT seems to be more tightly bounded to heparin than α-AT (-948 ± 117 kJ mol^{-1} against -868 ± 120 kJ mol^{-1}), which is in accordance to their experimentally determined AT-heparin dissociation constants[8,14] and fluorescence data.[7]

In fact, the decreased conformational flexibility caused by Asn135 oligosaccharide has been suggested as an important factor contributing to the reduction of α-AT affinity for heparin,[35] thus destabilizing its activated state, in relation to the native form.[14] In order to obtain further structural insights, at the atomic level, into the differential recognition of heparin by both AT glycoforms, and in relation to the nonglycosylated state, the polysaccharide dynamics on the three AT glycoforms surface was evaluated, as shown in Figure 3(D–F). Such data indicates that the GAG behaves comparably in the lower portion of the serpin, but is significantly diverse in the α-glycoform upper segment. Accordingly, a bending in the polysaccharide may be observed close to helix D as due to an obstruction caused by the Asn135-linked oligosaccharide [Fig. 3(F)], apparently moving a region of the heparin chain known to increase the GAG affinity for AT, outside the synthetic pentasaccharidic sequence (Supporting Information, Fig. S5) away from helix D. Such observation may consist in an effect associated to the glycoform-dependent AT-heparin recognition, additional to those previously reported.[14,35]

Mechanism of AT activation by heparin

While helix D extension was proposed as a major contributing factor to the stabilization of AT activated form,[66] the hinge region expulsion has been suggested as the most important AT activating conformational change, at least for fXa and fIXa proteases inhibition.[67] However, to date, the signal transmission mechanism from the heparin binding exosite, passing through helix D, to the RCL is not completely understood at the atomic level. Nevertheless, it has been proposed that helix D extension, stimulated by the GAG, induces shifts in strands A2 and A3, thus expelling the RCL.[68] On the other hand, studies on X-ray structures suggested that rearrangements on structural elements of the protein core, including helices D, E, and F[19,69] may contribute in this process. In order

Table II
Comparison of the Enthalpic Contribution of Interaction (kJ/mol) between Secondary Structure Elements in the Native and Activated AT States

Secondary structure elements		AT isoforms					
		Nonglycosylated AT		β-AT		α-AT	
		Native	Activated	Native	Activated	Native	Activated
Strand A2	Helix D	−139 ± 22	−69 ± 29	−95 ± 14	−80 ± 22	−81 ± 33	−75 ± 18
Strand A2	Helix B	−53 ± 15	−27 ± 11	−55 ± 20	−21 ± 9	−67 ± 19	−21 ± 9
Strand A3	Helix F	−50 ± 14	−32 ± 6	−42 ± 14	−23 ± 4	−41 ± 11	−25 ± 7
Helix B	Strand A3	−49 ± 16	−34 ± 9	−52 ± 12	−38 ± 13	−32 ± 8	−30 ± 10
Helix B	Helix E	−51 ± 21	−87 ± 20	−59 ± 10	−93 ± 14	−69 ± 11	−75 ± 15
Helix A	Helix D	−52 ± 14	−71 ± 7	−60 ± 13	−78 ± 16	−66 ± 10	−73 ± 10
Helix A	Helix P	−75 ± 28	−141 ± 25	−95 ± 23	−145 ± 19	−57 ± 26	−148 ± 22
Helix D	Helix P	−51 ± 6	−25 ± 4	−55 ± 7	−27 ± 4	−74 ± 10	−20 ± 6
Helix B	Helix P	−15 ± 10	−32 ± 14	−3 ± 4	−23 ± 13	−6 ± 8	−38 ± 12

to obtain further structural and dynamical insights into the conformational changes associated to AT activation, a comparison between protein-protein contacts in the native and heparin-bounded, activated AT states was performed (Table II).

The obtained data indicates that such conformational changes involve a dynamical set of forthcomings and withdrawals between AT secondary structure elements (see Fig. 4). Most of them were observed for the three AT studied glycosylation states, thus possibly constituting glycoform-independent conformational changes [Fig. 4(C–E)], from which the participation of specific amino acid residues could be proposed, as shown in Supporting Information, Table S5. (1) At first, upon heparin binding, an interaction enhancement between helixes A and P mainly through their composing residues Glu50 and

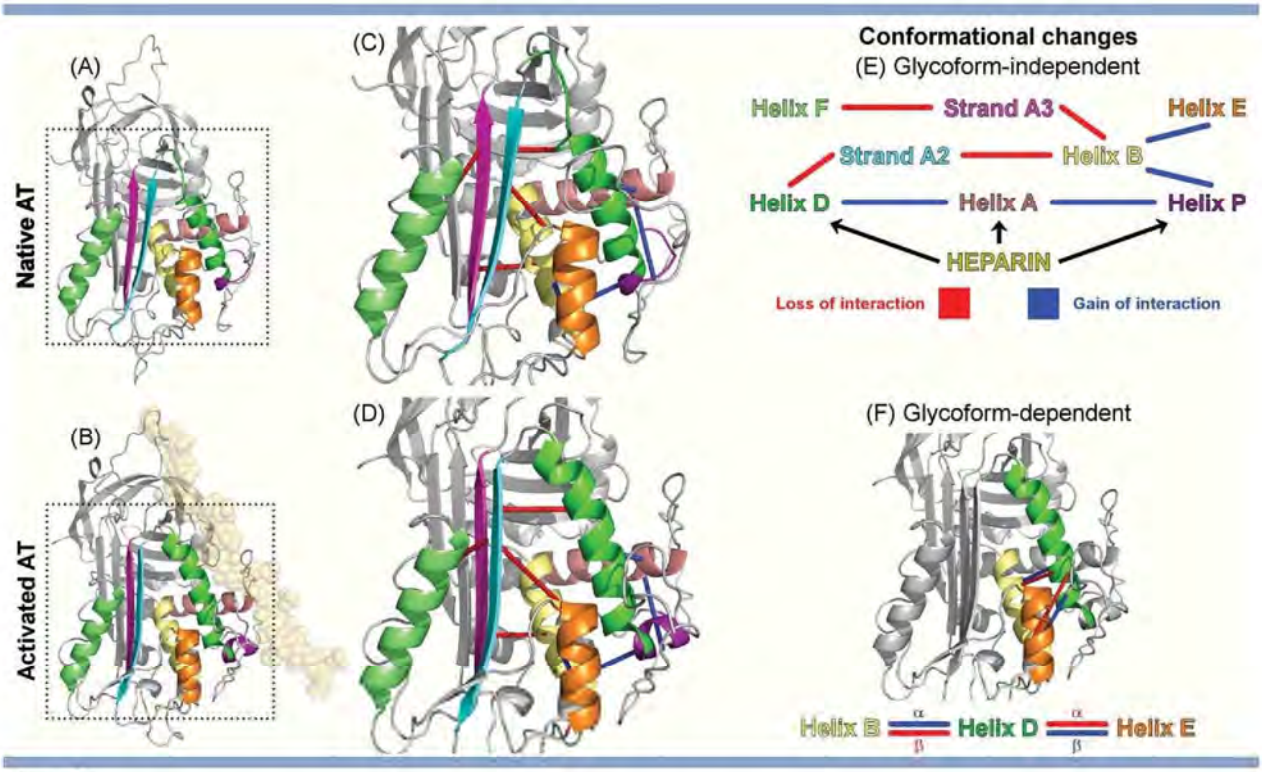


Figure 4
Proposal to the signal transmission pathway from heparin binding to β-sheet A. In the conformational pathway from (**A**) AT native form to the (**B**) AT activated state, several intramolecular rearrangements (**C–E**) would occur independently on the AT glycoform, comprising forthcomings and increases in interactions (blue) involving α-helices near heparin binding, with their consequent withdrawal and loss of interaction (red) from β-sheet A. As well, different interactions involving helix D were observed between the α- and β-glycoforms (**F**), possibly consisting of glycoform-dependent conformational changes within the serpin.

Thr115, respectively; (2) such approximation propagates from helix P to helix B through residues Lys91 and Asp117, (3) as well as from helix B (Met89) to helix E (Tyr158). Back to the heparin binding exosite, (4) helixes A and D also appear to be coupled, apparently through hydrophobic interactions involving residues Val48, Leu51, Phe123, and Leu126; subsequently, (5) both helixes B and D are withdrawn from strand A2: the first by departing Thr85 from Ser142, and helix D by disrupting a cluster around Tyr131, as previously proposed[70]; (6) finally, strand A3 is moved away from helix B, by displacing Pro87 from Leu215, and helix F. As a consequence of this process, the approach of helixes A, B, D, E, and P appears to allow a relaxation around strands A2 and A3, thus supporting a better organization of β-sheet A as a whole and the expulsion of the hinge region.

Additionally, different interactions involving helix D were observed in α-AT when compared with the nonglycosylated and β-AT as due to the presence of the additional *N*-glycan chain in the α-glycoform [Fig. 4(F)]. While the interaction between helixes B and D is more intense at the native state in the less or nonglycosylated forms, such behavior is inverted in α-AT (Supporting Information, Table S6). As well, the interaction between helices D and E is stronger at the activated state of nonglycosylated and β-AT, being also changed in the fully glycosylated form (Supporting Information, Table S6). Such glycoform-dependent interactions, although not directly involved with β-sheet A, may indicate a role of the Asn135-linked oligosaccharide chain in not only affecting heparin affinity by bending the polysaccharide chain and reducing heparin-AT interaction energy, but also altering intramolecular interactions within the serpin. This differential behavior may be also responsible for altering both the rate of heparin-induced conformational activation and the rate of its reversal,[14] possibly disturbing the signal transmission from the GAG binding site to β-sheet A, thus delaying the most important AT activating conformational change.

REFERENCES

1. Rau JC, Beaulieu LM, Huntington JA, Church FC. Serpins in thrombosis, hemostasis and fibrinolysis. J Thromb Haemost 2007;5: 102–115.
2. Gettins PGW. Serpin structure, mechanism, and function. Chem Rev 2002;102:4751–4803.
3. Huntington JA, Read RJ, Carrell RW. Structure of a serpin-protease complex shows inhibition by deformation. Nature 2000;407:923–926.
4. Olson ST, Richard B, Izaguirre G, Schedin-Weiss S, Gettins PGW. Molecular mechanisms of antithrombineheparin regulation of blood clotting proteinases. A paradigm for understanding proteinase regulation by serpin family protein proteinase inhibitors. Biochimie 2010;92:1587–1596.
5. Olson ST, Swanson R, Raub-Segall E, Bedsted T, Sadri M, Petitou M, Herault J-P, Herbert J-M, Björk I. Accelerating ability of synthetic oligosaccharides on antithrombin inhibition of proteinases of the clotting and fibrinolytic systems. Comparison with heparin and low-molecular-weight heparin Thromb Haemost 2004;92:929–939.
6. Olson ST, Björk I, Sheffer R, Craig PA, Shore JD, Choay J. Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. J Biol Chem 1992;267:12528–12538.
7. Peterson CB, Blackburn MN. Isolation and characterization of an antithrombin III variant with reduced carbohydrate content and enhanced heparin binding. J Biol Chem 1985;260:610–615.
8. Turko IV, Fan B, Gettins PGW. Carbohydrate isoforms of antithrombin variant N135Q with different heparin affinities. FEBS Lett 1993;335:9–12.
9. Brennan SO, George PM, Jordan RE. Physiological variant of antithrombin-III lacks carbohydrate sidechain at Asn 135. FEBS Lett 1987;219:431–436.
10. Picard V, Ersdal-Badju E, Bock SC. Partial glycosylation of antithrombin III asparagine-135 is caused by the serine in the third position of its *N*-glycosylation consensus sequence and is responsible for production of the β-antithrombin III isoform with enhanced heparin affinity. Biochemistry 1995;34:8433–8440.
11. Franzén LE, Svensson S, Larm O. Structural studies on the carbohydrate portion of human antithrombin III. J Biol Chem 1980;255: 5090–5093.
12. Plematl A, Demelbauer UM, Josic D, Rizzi A. Determination of the site-specific and isoform-specific glycosylation in human plasma-derived antithrombin by IEF and capillary HPLC-ESI-MS/MS. Proteomics 2005;5:4025–4033.
13. Demelbauer UM, Plematl A, Josic D, Allmaier G, Rizzi A. On the variation of glycosylation in human plasma derived antithrombin. J Chromatogr A 2005;1080:15–21.
14. Turk B, Brieditis I, Bock SC, Olson ST, Björk I. The oligosaccharide side chain on Asn-135 of α-antithrombin, absent in β-antithrombin, decreases the heparin affinity of the inhibitor by affecting the heparin-induced conformational change. Biochemistry 1997;36: 6682–6691.
15. Witmer MR, Hatton MW. Antithrombin III-beta associates more readily than antithrombin III-alpha with uninjured and de-endothelialized aortic wall in vitro and in vivo. Arterioscler Thromb 1991; 11:530–539.
16. Frebelius S, Isaksson S, Swedenborg J. Thrombin inhibition by antithrombin III on the subendothelium is explained by the isoform ATβ. Arterioscler Thromb Vasc Biol 1996;16:1292–1297.
17. Skinner R, Abrahams JP, Whisstock JC, Lesk AM, Carrell RW, Wardell MR. The 2.6 A structure of antithrombin indicates a conformational change at the heparin binding site. J Mol Biol 1997; 266:601–609.
18. Jin L, Abrahams JP, Skinner R, Petitou M, Pike RN, Carrell RW. The anticoagulant activation of antithrombin by heparin. Proc Natl Acad Sci USA 1997;94:14683–14688.
19. Whisstock JC, Pike RN, Jin L, Skinner R, Pei XY, Carrell RW, Lesk AM. Conformational changes in serpins. II. The mechanism of activation of antithrombin by heparin. J Mol Biol 2000;301:1287–1305.
20. Woods RJ. Computational carbohydrate chemistry: what theoretical methods can tell us. Glycoconj J 1998;15:209–216.
21. Zuegg J, Gready JE. Molecular dynamics simulation of human prion protein including both N-linked oligosaccharides and the GPI anchor. Glycobiology 2000;10:959–974.
22. Mandal TK, Mukhopadhyay C. Effect of glycosylation on structure and dynamics of MHC class I glycoprotein: a molecular dynamics study. Biopolymers 2001;59:11–23.
23. Pol-Fachin L, Fernandes CL, Verli H. GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations. Carbohydr Res 2009; 344:491–500.
24. Verli H, Guimarães JA. Insights into the induced fit mechanism in antithrombin-heparin interaction using molecular dynamics simulations. J Mol Graph Model 2005;24:203–212.

25. IUPAC-IUBMB Commission on Biochemical Nomenclature. Nomenclature of carbohydrates. Pure Appl Chem 1996;68:1919–2008.
26. Schuettelkopf AW, van Aalten DMF. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr D Biol Crystallogr 2004;60:1355–1363.
27. Schaftenaar G, Noordik JH. Molden: a pre- and post-processing program for molecular and electronic structures. J Comput Aided Mol Des 2000;14:123–134.
28. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33–38.
29. DeLano WL. The PyMOL molecular graphics system. San Carlos, CA: DeLano Scientific LCC; 2002.
30. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 2006;22:195–201.
31. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: fast, flexible, and free. J Comput Chem 2005;26:1701–1718.
32. Scott WRP, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Krueger P, van Gunsteren WF. The GROMOS biomolecular simulation program. J Phys Chem A 1999;103:3596–3607.
33. Verli H, Guimarães JA. Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. Carbohydr Res 2004;339:281–290.
34. Becker CF, Guimarães JA, Verli H. Molecular dynamics and atomic charge calculations in the study of heparin conformation in aqueous solution. Carbohydr Res 2005;340:1499–1507.
35. McCoy AJ, Pei XY, Skinner R, Abrahams JP, Carrell RW. Structure of β-antithrombin and the effect of glycosylation on antithrombin's heparin affinity and activity. J Mol Biol 2003;326:823–833.
36. Bohne-Lang A, von der Lieth C-W. Gly Prot: in silico glycosylation of proteins Nucleic Acids Res 2005;33:W214–W219.
37. Fernandes CL, Sachett LG, Pol-Fachin L, Verli H. GROMOS96 43a1 performance in predicting oligosaccharide conformational ensembles within glycoproteins. Carbohydr Res 2010;345:663–671.
38. Pol-Fachin L, Verli H. Depiction of the forces participating in the 2-*O*-sulfo-α-L-iduronic acid conformational preference in heparin sequences in aqueous solutions. Carbohydr Res 2008;343:1435–1445.
39. Guerrini M, Guglieri S, Casu B, Torri G, Mourier P, Boudier C, Viskov C. Antithrombin-binding octasaccharides and role of extensions of the active pentasaccharide sequence in the specificity and strength of interaction. Evidence for very high affinity induced by an unusual glucuronic acid residue. J Biol Chem 2008;283:26662–26675.
40. Richard B, Swanson R, Olson ST. The signature 3-*O*-sulfo group of the anticoagulant heparin sequence is critical for heparin binding to antithrombin but is not required for allosteric activation. J Biol Chem 2008;284:27054–27064.
41. Kusche M, Torri G, Casu B, Lindahl U. Biosynthesis of heparin. Availability of glucosaminyl 3-*O*-sulfation sites J Biol Chem 1990;265:7292–7300.
42. Berendsen HJC, Grigera JR, Straatsma TP. The missing term in effective pair potentials. J Phys Chem 1987;91:6269–6271.
43. de Groot BL, Grubmüller H. Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. Science 2001;294:2353–2357.
44. Hess B, Bekker H, Berendsen HJC, Fraaije ,JGEM. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18:1463–1472.
45. 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.
46. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
47. Pol-Fachin L, Verli H. Molecular modeling on the characterization of glycoproteins conformation. In: Verli H, Guimarães JA, editors. Strategies for the determination of carbohydrates structure and conformation. Kerala, India: Transworld Research Network; 2010. pp 103–131.
48. Mulloy B, Forster MJ, Jones C, Davies DB. N.m.r. and molecular-modelling studies of the solution conformation of heparin. Biochem J 1993;293:849–858.
49. Khan S, Gor J, Mulloy B, Perkins SJ. Semi-rigid solution structures of heparin by constrained X-ray scattering modelling: new insight into heparin-protein complexes. J Mol Biol 2010;395:504–521.
50. Meagher JL, Olson ST, Gettins PGW. Critical role of the linker region between helix D and strand 2A in heparin activation of antithrombin. J Biol Chem 2000;275:2698–2704.
51. Johnson DJD, Huntington JA. Crystal structure of antithrombin in a heparin-bound intermediate state. Biochemistry 2003;42:8712–8719.
52. Johnson DJD, Langdown L, Li W, Luis SA, Baglin TA, Huntington JA. Crystal structure of monomeric native antithrombin reveals a novel reactive center loop conformation. J Biol Chem 2006;281:35478–35486.
53. Langdown L, Belzar KJ, Savory WJ, Baglin TA, Huntington JA. The critical role of hinge-region expulsion in the induced-fit heparin binding mechanism of antithrombin. J Mol Biol 2009;386:1278–1289.
54. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nat Struct Biol 2002;9:646–652.
55. Olson ST, Srinivasan KR, Björk I, Shore JD. Binding of high affinity heparin to antithrombin III. Stopped flow kinetic studies of the binding interaction. J Biol Chem 1981;256:11073–11079.
56. Schedin-Weiss S, Richard B, Olson ST. Kinetic evidence that allosteric activation of antithrombin by heparin is mediated by two sequential conformational changes. Arch Biochem Biophys 2010;504:169–176.
57. Ersdal-Badju E, Lu A, Zuo Y, Picard V, Bock SC. Identification of the antithrombin III heparin binding site. J Biol Chem 1997;272:19393–19400.
58. Desai U, Swanson R, Bock SC, Björk I, Olson ST. Role of arginine 129 in heparin binding and activation of antithrombin. J Biol Chem 2000;275:18976–18984.
59. Arocas V, Turk B, Bock SC, Olson ST, Björk I. The region of antithrombin interacting with full-length heparin chains outside the high-affinity pentasaccharide sequence extends to Lys136 but not to Lys139. Biochemistry 2000;39:8512–8518.
60. Arocas V, Bock SC, Raja S, Olson ST, Björk I. Lysine 114 of antithrombin is of crucial importance for the affinity and kinetics of heparin pentasaccharide binding. J Biol Chem 2001;276:43809–43817.
61. Schedin-Weiss S, Desai UR, Bock SC, Gettins PGW, Olson ST, Björk I. Importance of lysine 125 for heparin binding and activation of antithrombin. Biochemistry 2002;41:4779–4788.
62. Jairajpuri MA, Lu A, Desai UR, Olson ST, Björk I, Bock SC. Antithrombin III phenylalanines 122 and 121 contribute to its high affinity for heparin and its conformational activation. J Biol Chem 2003;278:15941–15950.
63. Garone L, Edmunds T, Hanson E, Bernasconi R, Huntington JA, Meagher JL, Fan B, Gettins PGW. Antithrombin-heparin affinity reduced by fucosylation of carbohydrate at asparagine 155. Biochemistry 1996;35:8881–8889.
64. Olson ST, Frances-Chmura AM, Swanson R, Björk I. Zettlmeissl G. Effect of individual carbohydrate chains of recombinant antithrombin on heparin affinity and on the generation of glycoforms differing in heparin affinity Arch Biochem Biophys 1997;341:212–221.
65. Björk I, Ylinenjärvi K, Olson ST, Hermentin P, Conradt HS, Zettlmeissl G. Decreased affinity of recombinant antithrombin for heparin due to increased glycosylation. Biochem J 1992;286:793–800.

66. Belzar KJ, Zhou A, Carrell RW, Gettins PGW, Huntington JA. Helix D elongation and allosteric activation of antithrombin. J Biol Chem 2002;277:8551–8558.
67. Langdown J, Johnson DJ, Baglin TP, Huntington JA. Allosteric activation of antithrombin critically depends upon hinge region extension. J Biol Chem 2004;279:47288–47297.
68. Van Boeckel CAA, Grootenhuis PDJ, Visser A. A mechanism for heparin-induced potentiation of antithrombin III. Nat Struct Biol 1994;1:423–425.
69. Huntington JA, McCoy A, Belzar KJ, Pei XY, Gettins PGW, Carrell RW. The conformational activation of antithrombin. A 285-Å structure of a fluorescein derivative reveals an electrostatic link between the hinge and heparin binding regions J Biol Chem 2000;275: 15377–15383.
70. dela Cruz RG, Jairajpuri MA, Bock SC. Disruption of a tight cluster surrounding tyrosine 131 in the native conformation of antithrombin III activates it for factor Xa inhibition. J Biol Chem 2006;281: 31668–31676.

4.4 Trabalho III

Em continuação ao trabalho anterior, este trabalho descreve o comportamento conformacional de proteases da cascata de coagulação sanguínea, também complexadas e não-complexadas à heparina, a fim de avaliar o processo de inibição mediado pelo GAG. A ação de inibidores sobre fIIa e fXa é bastante explorado na literatura, e descrito para ligantes em seus sítios ativos e no exosítio-1, localizado na superfície das enzimas. No entanto, a ação de inibidores sobre o exosítio 2, onde heparina se liga, é pouco descrita. Adicionalmente, fIIa é uma glicoproteína, no entanto a influência da glicosilação sobre sua estrutura e processo de inibição não é bem compreendida. Neste contexto, foram realizadas simulações por DM de fIIa, glicosilado e não-glicosilado, visando estudar os efeitos da glicosilação sobre a dinâmica e função da protease, bem como de fIIa e fXa complexados ao GAG, a fim de avaliar os efeitos da interação de heparina sobre a dinâmica de inibição das enzimas.

Os resultados obtidos indicam que a glicosilação, em fIIa, é responsável pelo aumento do tamanho da cavidade do sítio ativo da enzima, o que pode contribuir para fIIa glicosilado possuir uma atividade catalítica maior que fIIa não-glicosilado, tal como descrito experimentalmente. A interação de heparina com as proteases levou a uma desorganização da geometria da tríade catalítica das enzimas, o que está de acordo com estudos prévios da literatura, e sugere um efeito de inibição alostérica sobre elas. A partir disso, o envolvimento de alguns resíduos de aminoácido com esse alosterismo foi proposto. Finalmente, a dinâmica de heparina na superfície das enzimas sugere que a orientação do GAG determina o mecanismo de ação a qual elas serão inibidas por AT.

Structural glycobiology of heparin dynamics on the exosite 2 of coagulation cascade proteases: Implications for glycosaminoglycans antithrombotic activity

Laercio Pol-Fachin, Hugo Verli

Glycobiology, **2014**, *24*; 97-105

Glycobiology vol. 24 no. 1 pp. 97–105, 2014
doi:10.1093/glycob/cwt095
Advance Access publication on November 7, 2013

Structural glycobiology of heparin dynamics on the exosite 2 of coagulation cascade proteases: Implications for glycosaminoglycans antithrombotic activity

Laercio Pol-Fachin[2] and Hugo Verli[1,2]

[2]Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av Bento Gonçalves 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil

Received on June 4, 2013; revised on October 16, 2013; accepted on October 16, 2013

fIIa and fXa are two of the main targets of antithrombin, a serine proteases inhibitor that plays a major role in the regulation of blood clotting. The formation of ternary complexes between such molecules and glycosaminoglycans, as heparin, is the main path for inhibiting those enzymes, which may occur through two distinct mechanisms of action. While these serine proteases present distinct susceptibilities to these paths, in which fIIa demands an interaction with heparin, neither the molecular basis of this differential inhibition nor the role of fIIa glycosylation on this process is fully understood. Thus, the present work evaluated through molecular dynamics simulations the effects of glycosylation on fIIa and the consequences of heparin binding to both proteases function and dynamics. Based on the obtained data, fIIa N-linked glycan promoted an increase in the active site pocket size by stabilizing regions that encircle it, while heparin binding was observed to reverse such an effect. Additionally, heparin orientation observed on the surface of fIIa, but not fXa, allows a linear long-chain heparin binding to antithrombin in ternary complexes. Finally, the enzymes catalytic triad organization was disrupted due to a strong glycosaminoglycan binding to the proteases exosite 2. Such data support an atomic-level explanation for the higher inhibition constant of the antithrombin–heparin complex over fIIa than fXa, as well as for the different susceptibilities of those enzymes for antithrombin mechanisms of action.

Keywords: fXa / glycosylation / hemostasis / molecular dynamics / thrombin

Introduction

Blood coagulation involves the sequential activation of serine proteases, which culminates in the generation of thrombin (fIIa) and a subsequent conversion of fibrinogen into insoluble fibrin (Furie and Furie 1988). Several studies point out that, among such reactions, those related to factor X activation into factor Xa (fXa) by the intrinsic tenase complex is the rate-limiting step for fIIa generation and, consequently, fibrin production (Lawson et al. 1994; Rand et al. 1996). The major inhibitor of plasma proteinases, antithrombin (AT), a member of serine proteinase inhibitors (serpin) family, shows as its two main targets both fIIa and fXa (Rau et al. 2007). AT mechanism of action includes the formation of highly stable equimolar complexes with such enzymes, which are possibly inactivated by structural deformation (Huntington et al. 2000). Additionally, sulfated glycosaminoglycans (GAG), as therapeutic heparin or endothelial heparan sulfate proteoglycans, are known to accelerate such inactivating reactions upon AT binding to a minimum sequence-specific pentasaccharide site (Olson et al. 2004). GAGs are a group of sulfated polysaccharides composed of repeating disaccharide units, varying in their glycosidic linkage geometries and monosaccharide constitution (Gandhi and Mancera 2008). Heparin, specifically, consists of repeating units of $1 \rightarrow 4$ linked glucosamine and iduronic or glucuronic acid residues, with varying sulfation pattern, depending on the polysaccharide source (Nader et al. 2001).

It is well established that heparin can exert its effect on AT-proteases inhibition rate through [1] a conformational activation mechanism, in which the sequence-specific pentasaccharide fragment binds uniquely to AT, altering its conformation to improve its reactivity with coagulation proteases, or [2] a bridge (template) mechanism, in which long-chain heparins concomitantly bind both serpin and protease (Gettins 2002; Olson et al. 2004). Regarding such mechanisms of action, while the specific heparin pentasaccharide alone is sufficient to enhance the rate of inactivation over fXa (Choay et al. 1983), as observed from its corresponding X-ray structure (Johnson et al. 2006), at least 13 additional saccharide residues are required for a significant heparin-mediated inhibition of fIIa (Laurent et al. 1978; Oosta et al. 1981; Lane et al. 1984). In this scenario, while an AT-mediated thrombin inhibition demands protease–heparin interaction, as observed through X-ray crystallography (Li et al. 2004), only under certain conditions is fXa capable of binding such GAG in a bridging complex (Rezaie 1998; Rezaie and Olson 2000). For both enzymes, such protein–heparin interaction

[1]To whom correspondence should be addressed: Tel: +55-51-3308-7770; Fax: +55-51-3308-7309; e-mail: hverli@cbiot.ufrgs.br; URL: http://www.ufrgs.br/bioinfo

is described to occur through their exosite 2-composing amino acid residues (Sheehan and Sadler 1994; Rezaie 2000). In addition, GAGs binding to both proteases has been shown to induce conformational changes in their active sites (Fredenburgh et al. 1997; O'Brien et al. 2003). In spite of such an amount of information regarding fIIa and fXa inhibition, a structural determination, at atomic level, of the possible differences on such enzymes binding to heparin and the consequences of such interaction over their dynamics is still absent. Considering the highly flexible nature of the carbohydrate chains composing such systems, including heparin oligosaccharides, molecular dynamics (MD) simulations emerge as a promising tool for describing such complexes considering both spatial and temporal properties, thus accounting for most of its containing structural and dynamical variables, as solvation and thrombin *N*-glycosylation (Nilsson et al. 1983).

In this context, the present work intends to evaluate heparin binding to the exosite 2 of fIIa and fXa, as well as the mutual influence of this interaction between polysaccharide and enzymes, through a series of MD simulations. Additionally, under the evaluated conditions, thrombin was studied comparatively in its non-glycosylated and glycosylated (physiological) forms. Such data were accordingly compared with those of previous biochemical, mutagenesis-derived analysis and X-ray crystallography structures, as well as reinforced by dynamic network and principal component analyses. Based on these data, insights into the formation of ternary complexes involving the GAG, AT and two of the main coagulation cascade serine proteases could be obtained, supporting the proposal of an atomic model for the dynamics of heparin and other endogenous GAGs when bound to such enzymes.

Results

Systems preparation

In order to assess the effects of heparin interaction with the evaluated serine proteases, each enzyme was studied both unbound and bound to the GAG. fIIa was considered in its physiological, glycosylated form and in a non-glycosylated state, which was studied as a control to evaluate the effects of glycosylation on its structure and flexibility. In addition, fXa was assessed in systems both including and not including Ca^{2+}, as the presence of such divalent cation allows fXa to bind heparin in a bridging complex with AT (Rezaie 1998; Rezaie and Olson 2000). Thus, eight systems were studied, comprising non-glycosylated fIIa, glycosylated fIIa, fXa in a Ca^{2+}-absent system and fXa in a system with Ca^{2+}, each unbound and bound to heparin. The studied enzymes show very different non-catalytic domains, both in terms of size (Supplementary data, Figure S1A) and function, especially regarding fXa light-chain γ-carboxyglutamic acid (GLA) domain and its calcium-dependent interaction with anionic phospholipid surfaces (Furie and Furie 1988). As no interaction of heparin with these regions was observed, no data were obtained from the simulations for either fIIa or fXa light chains. Of interest to the present work, while the catalytic domain of thrombin and fXa presents only about 40% of sequence identity (Supplementary data, Figure S1B), their tertiary structure is quite similar, showing RMSD < 3 Å (Supplementary data, Figure S1C), and the exosite for heparin interaction is located in the same 3D region (Supplementary data, Figure S1D). In view of that, the heparin orientation on the surface of thrombin in previous X-ray structures (Li et al. 2004; Carter et al. 2005) was also employed as a starting structure for fXa. The structures and parameters for such heparin oligosaccharide, as well as for thrombin N-linked glycan, were obtained from previous publications on heparin-glycoprotein complexes (Pol-Fachin et al. 2011).

Thrombin glycosylation

The effects of N-glycans on thrombin were initially evaluated by its heavy-chain root mean square fluctuation (RMSF), from which it may be observed that the protein global flexibility is mostly unchanged due to glycosylation (Figure 1D). Still, the Asn60G-linked oligosaccharide appeared to be capable of reducing the flexibility of four regions that encircle the enzyme active site (Figure 1B–D), including the 60-60G region, where the *N*-glycan is attached. As the entrance to thrombin active site has been described to be partially occluded by the 60-60G loop in X-ray structures (Bode et al. 1992), we evaluated whether such flexibility difference would also influence the active site pocket size. Thus, based on CASTp measurement of cavity and pocket sizes (Dundas et al. 2006), it was observed that glycosylation, while stabilizing the motion of these regions, is capable of increasing the active site pocket size (Table I, $P \leq 0.05$). Such modifications were accompanied by alterations in the interaction of those regions with their surroundings, including a decreased interaction with the solvent (except for the 165–180 loop) and an increased interaction with the remaining protein moiety (Supplementary data, Table S1). In addition, the 60–60G and 84–98 regions were observed to interact with the N-glycan in the glycosylated fIIa system (Supplementary data, Table S1). Hence, their decreased motions in glycosylated systems may be attributed to their direct interaction with the Asn60G-linked oligosaccharide. In addition, from a dynamic network analysis (Supplementary data, Tables S2 and 3), it was observed that most residues comprising regions 84–98 (Figure 1D, region 2) and 165–180 (Figure 1D, region 4) are within the same community during all thrombin simulations, which indicates that their movement is correlated. Such data were confirmed by principal component analysis (PCA) on the MD trajectories, in which such regions showed decreased motions in the principal component of the flexibility of glycosylated thrombin (Supplementary data, Figure S2E). In this context, while region 4 has no interaction with fIIa N-glycan, and showed a different interaction pattern with the solvent, its decreased flexibility may be associated with region 2 reduced motions.

Serine proteases–heparin binding. Heparin binding to the studied serine proteases was initially evaluated by the energy interaction profile between such a polysaccharide and basic residues in the proteinases surface (Table II). For this comparison, the amino acid residues previously identified, through site-directed mutagenesis, as important components of exosite 2 of both fIIa and fXa (Sheehan and Sadler 1994; Rezaie 2000), were considered. Based on such analysis, it was observed that the amino acid residues showing more intense interaction energies with heparin are indeed those previously recognized as important for exosite 2-GAG recognition, mainly Arg93,

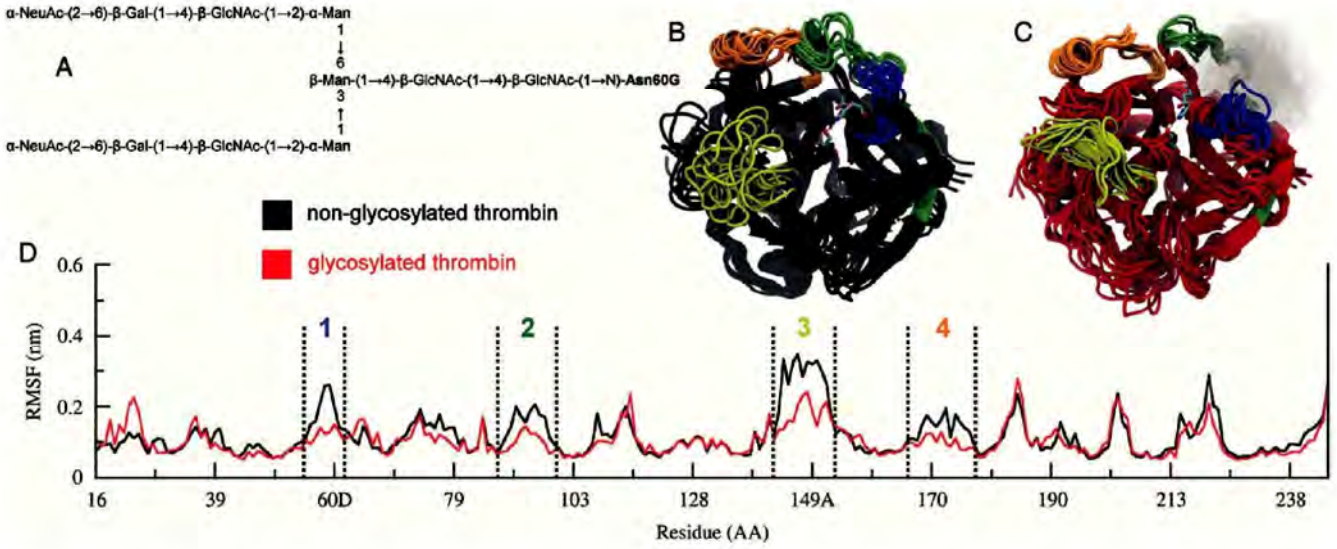


Figure 1. Effects of glycosylation over thrombin dynamics. In (**A**), schematics of the biantennary, complex-type oligosaccharide used to construct glycosylated thrombin. In (**B**) and (**C**), the conformational profile inherent to regions that presented reduced flexibility in the presence of N-glycosylation is shown, in which frames at every 10 ns of MD simulations are superimposed (the catalytic triad is presented as cyan sticks and thrombin N-glycan as gray mesh). In (**D**), the root mean square fluctuation (RMSF) analysis for non-glycosylated (black) and glycosylated (red) thrombin is shown. Dashed lines indicate regions between resides 60-60G (blue), 84–98 (green), 140–150 (yellow) and 165–180 (orange) regions.

Table I. Thrombin active site pocket size

fIIa system[a]	Non-glycosylated	Glycosylated (Å^3)
Heparin absent	463 ± 328	790 ± 293
Complexed to heparin	423 ± 282	624 ± 363

[a]The active site pocket volume was measured and averaged, with ±standard deviation, from frames collected at every 10 ns of MD simulations.

Table II. Interaction energy between heparin and selected amino acid residues on the surface of the enzymes

Amino acid residues (fIIa/fXa)	Interaction energy with Heparin (kJ/mol)[a]			
	fIIa		fXa	
	Non-glycosylated	Glycosylated	Without Ca^{2+}	With Ca^{2+}
Lys87/Val87	−50 ± 31	−2 ± 11	−14 ± 5	−20 ± 4
Ile90/Lys90	−19 ± 14	0	−61 ± 32	−51 ± 29
His91/His91	−31 ± 14	−26 ± 12	−10 ± 3	−20 ± 17
Arg93/Arg93	−85 ± 26	−78 ± 26	−115 ± 26	−74 ± 39
Arg97/Lys96	−3 ± 15	−20 ± 28	−42 ± 31	−1 ± 7
Arg101/Phe101	−74 ± 35	−37 ± 24	−16 ± 7	−4 ± 4
Arg126/Arg125	−63 ± 38	−82 ± 41	−27 ± 26	−34 ± 36
Arg165/Arg165	0	0	−112 ± 37	−8 ± 19
Arg169/Lys169	0	0	−78 ± 38	0
Arg233/Ala233	−39 ± 26	−42 ± 54	0	0
Lys235/Leu235	−11 ± 24	−36 ± 27	0	−4 ± 7
Lys236/Lys236	−96 ± 51	−155 ± 43	−37 ± 42	−87 ± 53
Lys240/Arg240	−46 ± 43	−100 ± 42	−68 ± 19	−93 ± 36
Exosite-2 residues	−339 ± 91	−492 ± 103	−478 ± 94	−296 ± 105
All evaluated residues	−518 ± 101	−579 ± 97	−580 ± 96	−395 ± 117

Underlined residues were previously identified as important for the enzymes interaction with heparin.
[a]The entire employed timescale (0.1 μs) was considered for obtaining the average and standard deviation values.

Arg126, Lys236 and Arg240 (Sheehan and Sadler 1994; Rezaie 2000). Markedly, Arg165 and Lys169, identified as exosite 2-composing residues in fXa (Rezaie 2000), strongly interacted with the polysaccharide in the fXa–heparin system, while their counterparts in thrombin, Arg165 and Arg169, not identified as components of fIIa exosite 2, presented no interaction with the GAG (Table II). Such a behavior is also observed for fIIa Arg233 and Lys235 positively charged amino acid residues, described to compose thrombin exosite 2 (Sheehan and Sadler 1994), which interact with heparin during simulations, and their counterparts in fXa, Ala233 and Leu235, which showed no significant interaction with the GAG (Table II). Moreover, frames were collected at every 10 ns of the entire 0.1 μs trajectories, from which the heparin orientation in relation to the serine proteases was assessed (Figure 2A–D). As a general feature, during MD simulations, the oligosaccharides indeed remained around the proteases exosite-2 (Figure 2A–D), resembling the fIIa-heparin orientation (Figure 2E) in the X-ray structure of their ternary complex with AT (Li et al. 2004). However, in non-glycosylated thrombin (Figure 2A) and fXa with Ca^{2+} (Figure 2D) systems, heparin orientation showed an increased fluctuation on the surface of the enzymes when compared with glycosylated thrombin (Figure 2B) and fXa in the system without Ca^{2+} (Figure 2C). Moreover, the interaction energy between heparin and exosite-2 composing residues was higher in these systems in which the oligosaccharide presented reduced fluctuations (Table II). Furthermore, in contrast to what

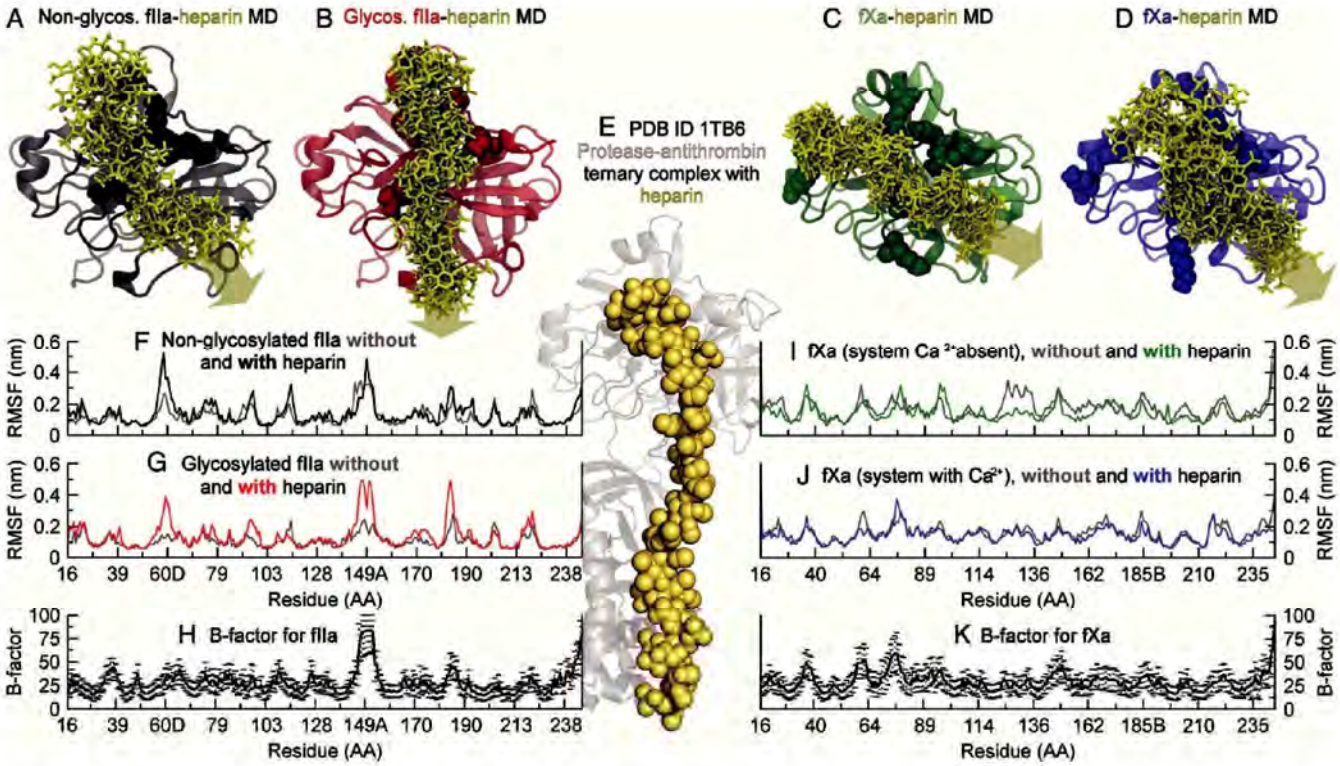


Figure 2. Comparison between heparin–proteases interaction. In (**A**–**D**), heparin dynamics on the surface of the enzymes is represented by frames obtained at every 10 ns of MD simulations. In (**E**), fIIa–heparin orientation in the ternary complex structure present in PDB ID 1TB6 is presented. In (**F**–**G**) and (**I**–**J**), the root mean square fluctuation (RMSF) analysis comparing the flexibility of the enzymes when bound and unbound to heparin is shown. B-factor values averaged from (**H**) ~200 available PDB entries for fIIa and (**K**) ~80 PDB entries for fXa are also presented.

happened due to glycosylation, heparin binding to fIIa appeared to promote a flexibility increase in the 60–60G, 84–98, 140–150 and 165–180 regions that encircle the thrombin active site (Figure 2F–G). Regarding fXa, no major flexibility differences may be observed due to heparin binding, but a decreased motion around residue 60 (Figure 2I and J), in a region that directly interacts with heparin during simulations. The only exception is the 120–134 region of fXa system without Ca^{2+} (Figure 2I), consisting of a helix that unfolded in the trajectory without heparin. Still, the flexibility behavior of both proteases is in agreement with a B-factor profile obtained by averaging their per residue values in more than 200 PDB entries for fIIa (Figure 2H, in comparison with Figure 2F and G) and in approximately 80 PDB entries for fXa (Figure 2K, in comparison with Figure 2I and J). In addition, the RMSD associated with such domains reflect the stability achieved during the second half of the simulations (Supplementary data, Figure S2), considered for the RMSF analyses.

Active site environment. Based on fluorescence evidences that the interaction of GAGs to fIIa and fXa exosites-2 induce conformational changes to their active sites (Fredenburgh et al. 1997; O'Brien et al. 2003), and in order to explore the effects of heparin binding on the serine proteases, the environment around the catalytic triad (His57, Asp102 and Ser195) of such enzymes was evaluated (Figure 3). From dynamic network analyses performed on the trajectories (Supplementary data, Table S2 and S3), it was observed that His57 and Asp102 were always within the same community, which reflects that their backbone movements were correlated in spite of heparin binding to either fIIa or fXa. On the other hand, His57 and Asp102 side-chain conformations were observed to be modified in previous publications due to inhibitors binding to fIIa exosite 2 (Fernández et al. 2013). Accordingly, in the present work, we observed that the hydrogen bond interaction between His57 Nδ1 and Asp102 Oδ2 was partially disrupted due to a strong heparin binding to the enzymes exosite-2-composing residues (Figure 3C and D), that is, in glycosylated fIIa and fXa in the system without Ca^{2+} (Table II). As a consequence of this binding, the distance between His57 Nδ1 and Asp102 Oδ2, which predominantly showed distances compatible with a hydrogen bond in the absence of heparin (~3 Å), presented a new conformational state when complexed to the GAG (~5 Å), thus departing such residues (Figure 3A) and partially disrupting the catalytic triad organization. Such increase in the distance between His57 Nδ1 and Asp102 Oδ2 was accompanied by modifications in His57 rotameric state (Hansen and Kay 2011). In the system containing heparin-unbound fXa in the absence of Ca^{2+}, His57 mostly populated the *gauche+* ($\chi^1 = 60$ degrees) state, and in glycosylated fIIa both *gauche+* and *trans* ($\chi^1 = 180$ degrees) states (Supplementary data, Figure S3). In the presence of heparin, in the two systems in which the catalytic triad organization was disturbed, His57 showed as equilibrium between the three possible rotameric states (Supplementary data, Figure S3), including a high prevalence of the *gauche–*

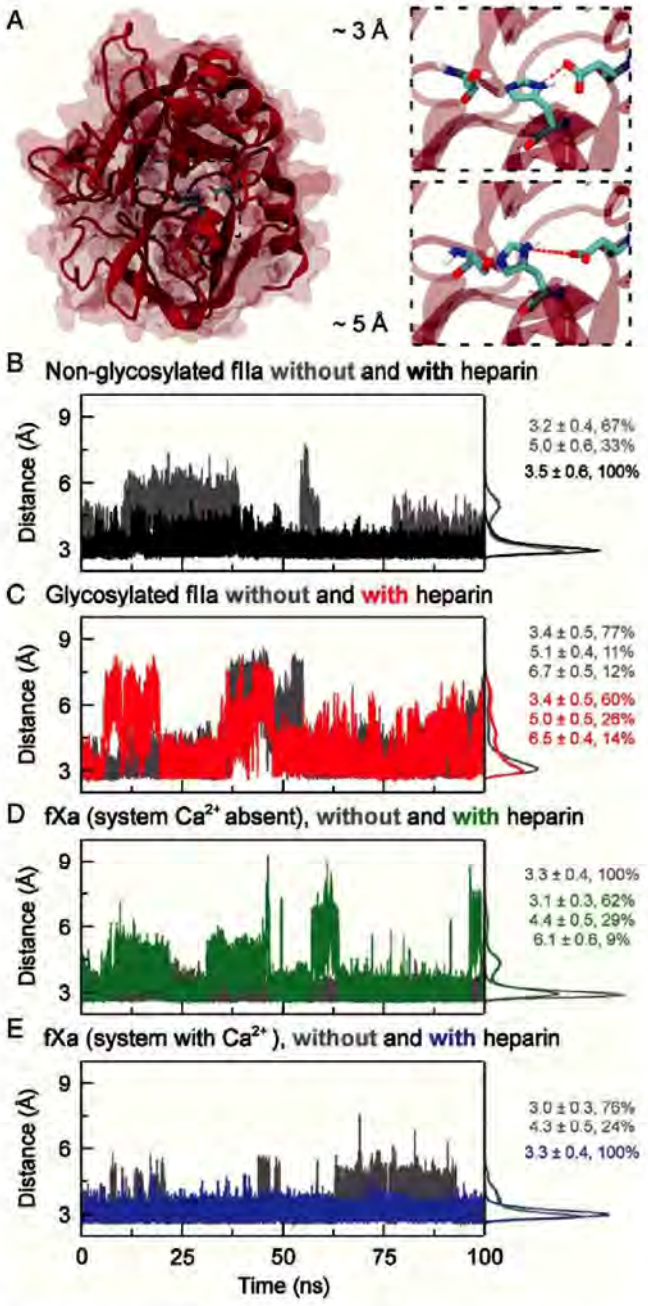


Figure 3. Catalytic triad arrangement in the studied enzymes. In (**A**), the different conformational states of the enzymes active site are presented (exemplified with thrombin, which structure is shown in ribbon diagrams and catalytic triad amino acid residues as sticks). In (**B**–**E**), the distances between the Nδ1 atom of His57 and Oδ1 atom of Asp102, which perform a hydrogen bond within the enzymes catalytic triad, are presented in their unbound (gray) and heparin-bound (coloured, online; foreground graphs in print) states.

($\chi^1 = -60$ degrees) conformation. In the systems containing non-glycosylated fIIa and fXa with Ca^{2+}, on the other hand, heparin interaction with the enzymes exosite 2-composing residues was weaker (Table II). Accordingly, the GAG also appeared to act modulating the distance between His57 Nδ1 and Asp102 Oδ2, so that the ~5 Å conformation was not observed (Figure 3B and E) and His57 rotameric state remained unaltered, as in the case of fXa, or was stabilized in the *gauche+* state, in non-glycosylated thrombin (Supplementary data, Figure S3).

Insights into heparin mechanisms of action

Extrapolating the obtained results in the context of AT-proteases-GAG ternary complexes, the orientation of heparin on the surface of the physiologically occurring, glycosylated fIIa, but not fXa—neither in the presence nor in the absence of Ca^{2+}—would allow a linear long-chain heparin binding to AT considering the bridge (template) mechanism (Figure 2A–E, shaded arrows). Such assumption is in accordance with previous suggestions based on the AT-fXa complex, in which it was proposed that a direct linear heparin bridge is not possible (Johnson et al. 2006). Therefore, it may be suggested that fXa is less susceptible to the bridge mechanism since its preferred orientation in relation to AT, as observed in previous X-ray structures (Johnson et al. 2006), does not allow the GAG to assume the linear conformation expected for such polysaccharide (Mulloy et al. 1993). Thus, heparin, in AT-fXa-GAG bridge complexes should bend perpendicularly on the surface of fXa in comparison with its observed orientation in previous X-ray studies for the AT-fIIa complex (Li et al. 2004). It may be observed, however, that the GAG orientation on the surface of fXa in the system with Ca^{2+} was more related to the available fIIa-heparin-AT ternary complex structure than in the absence of such a cation (Figure 2C–E, shaded arrows). In this context, it may be proposed that Ca^{2+} approaches heparin orientation on the surface of fXa to a linear long-chain GAG binding with AT. Such heparin reorientation appeared to be mediated by an intense Ca^{2+} interaction with the heparin region responsible for the main contacts between fXa exosite 2-composing residues and the GAG (Supplementary data, Table S4 and Figure S4). Regarding thrombin, heparin dynamics on the surface of glycosylated fIIa most perfectly fitted the required GAG arrangement in an AT–proteases–heparin ternary complex, while in non-glycosylated fIIa, it was quite comparable to its behavior when bound to fXa in the system with Ca^{2+}. Thus, it may be proposed that thrombin glycosylation, present in the physiologically occurring form of the enzyme, facilitates a proper AT–fIIa–heparin complex formation.

Discussion

While little is known about the effects of glycosylation on thrombin (Nilsson et al. 1983; Rosenfeld and Danishefsky 1984), the presence of *N*-glycans has been related to influence a broad range of properties from other glycoproteins, such as structure, folding and function (Pol-Fachin and Verli 2011). Still, conformational stabilization is the property reported to be most frequently modified by glycosylation (Pol-Fachin and Verli 2011). Accordingly, insights into the molecular basis for the flexibility decrease of the regions 60-60G, 84–98 and 165–180, caused by glycosylation, could be suggested in the present study. However, those related to the flexibility difference of 140–150 residues could be not proposed, as such loop had no interaction with thrombin N-glycan, and was separated in different communities to those of the remaining evaluated regions, as observed from dynamic network analyses, and confirmed by PCA. Still, based on the results obtained in the present work, it may be proposed that glycosylation shows a conformational

stabilization effect on fIIa, by decreasing the mobility of the above-mentioned regions, around the enzyme active site.

Heparin is recognized to be mostly involved with a serpin-mediated thrombin inhibition (Olson et al. 2004), and presumably does not alter such enzyme catalytic activity (Henry et al. 2007) while interacting with its exosite 2. Still, based on fluorescence data, conformational changes on fIIa and fXa active sites could be proposed due to GAGs binding to the enzymes surface (Fredenburgh et al. 1997; O'Brien et al. 2003). In this context, a partial disruption of such proteases catalytic triad could be observed due to a strong heparin binding to such enzymes exosites 2 (Figure 3), which may be proposed as the conformational changes observed in fluorescence experiments. Moreover, heparin binding was observed to additionally increase the flexibility of 60-60G, 84–98, 140–150 and 165–180 regions (Figure 2F and G) and to decrease its active site pocket size (Table I). Such aspects, which are possibly not related to thrombin direct inhibition by heparin, may indeed contribute to its inhibition by AT-GAG complexes, in which fIIa demands an interaction with the polysaccharide. In this context, the conformational modulation observed in the present work in regions that encircle thrombin active site are reinforced by findings that 140–150 and 165–180 loops were also observed to show an increased mobility when thrombin is inhibited by PPACK (Fuglestad et al. 2012). Also, no analogous regions exist in fXa, and thus, only the disruption of His Nδ1 to Asp Oδ2 distance occurs. Therefore, the observed influence of heparin binding on those regions encircling thrombin active site may also contribute to the higher antithrombin–heparin complex inhibition constant over fIIa than fXa (Craig et al. 1989).

Finally, the orientation of heparin on the surface of the evaluated serine proteases could be correlated with their diverse susceptibilities to heparin mechanisms of action. In fact, the orientation of heparin on the surface of fXa, in the absence of Ca^{2+}, would not allow a long linear arrangement of the GAG in a ternary complex with AT, which may contribute to its higher susceptibility to the conformational change-based mechanism, in which heparin does not directly interact with fXa. Markedly, considering previous studies indicating that fXa is only capable of forming a bridging complex with heparin in the presence of Ca^{2+} (Rezaie 1998; Rezaie and Olson 2000), the presence of this cation maintained heparin orientation on the surface of fXa similar to that observed for non-glycosylated fIIa, thus closer to the linear arrangement for a long-chain heparin binding to AT. Additionally, in the presence of glycosylation, thrombin interacted more intensely with the GAG, and its orientation on the surface of the enzyme mostly superimposed the required heparin arrangement in ternary complexes. Such data are expected to support future studies of the development of new anticoagulant agents, mimicking heparin interactions, both on AT and serine proteases, or focused on the enzymes exosite 2.

Materials and methods

Nomenclature and software

The recommendations and symbols of nomenclature as proposed by IUPAC (IUPAC 1996) are used. All saccharide topologies were generated with the PRODRG server (Schuettelkopf and van Aalten 2004), the manipulation of structures was performed with MOLDEN (Schaftenaar and Noordik 2000), VMD (Humphrey et al. 1996) and PyMOL (The PyMOL Molecular Graphics System), the homology modeling was performed with MODELLER v9.8 (Sanchez et al. 2000), the dynamic network analysis was performed with NetworkView (Eargle and Luthey-Schulten 2012) extension of VMD and all MD calculations and remaining analyses were performed using the GROMACS simulation suite (van der Spoel et al. 2005) and the GROMOS96 43A1 force field (Scott et al. 1999).

Topology construction for non-usual amino acid and carbohydrate residues

The employed protocol for carbohydrate topology construction was based on previous studies, for both N-glycan structures (Pol-Fachin et al. 2009, 2011) and sulfated polysaccharides (Verli and Guimarães 2004, 2005; Pol-Fachin and Verli 2008; Fernández et al. 2013). Such parameters, based on the GROMOS 43A1 force field, have been extensively validated against X-ray crystallography and NMR data, which substantiate the use of these topologies for carbohydrate MD simulations. Specifically, the monosaccharide fragments composing thrombin N-glycan chain and heparin polysaccharide fragments were constructed using the MOLDEN software (Schaftenaar and Noordik 2000). Each structure was then submitted to the PRODRG server (Schuettelkopf and van Aalten 2004), from which the crude topologies were retrieved. Based on such information, these structures were described in GROMOS96 43A1 force field parameters and further refined through: (1) improper dihedrals, employed to preserve the conformational state of the hexopyranose rings in $^4C_1$ (D-GlcNAc, D-Man, D-Gal and D-GlcNS,6S), ^2S_O (L-IdoA2S) or $^2C_5$ (NeuAc) forms; (2) proper dihedrals, as described in the GROMOS96 43a1 force field for glucose, in order to support stable simulations (Pol-Fachin et al. 2009) and (3) Löwdin HF/6-31G**-derived atomic charges, obtained from previous works of the group (Verli and Guimarães 2004; Pol-Fachin et al. 2009). In addition, the post-translational modifications known to occur along fXa light-chain polypeptide structure were also considered, that is, the presence of GLA residues at the GLA-domain and β-hydroxyaspartate (BHD) at position 63 (Hansson and Stenflo 2005). Thus, the parameters for such non-usual amino acid residues were compiled based on groups present within the GROMOS96 43A1 force field: the carboxylate group was used to build GLA from Glu residues, and the additional hydroxyl group of BHD was parameterized based on Ser side chain.

Thrombin N-glycosylation. The non-glycosylated thrombin protein core, comprising its complete light- and heavy-chain sequences, was retrieved from PDB ID 1PPB (Bode et al. 1989). In order to generate an initial model for its glycosylated, physiologically occurring form, the N-glycosylation site, located at the heavy-chain Asn60G (according to chymotrypsin numbering) residue, was modified to include a complex type, biantennary glycan structure, with two terminal sialic acids (Figure 1A), proposed as thrombin oligosaccharide moiety (Nilsson et al. 1983), using glycosciences modeling tools (Lütteke et al. 2006). Such models had their glycosidic linkage geometries adjusted to the main conformational states for each linkage, based on their relative abundance in the isolated

disaccharides in water, as previously described (Fernandes et al., 2010; Pol-Fachin et al. 2011).

fXa light-chain homology modeling. While the 3D structure for fXa heavy chain was retrieved from PDB ID 1XKA (Kamata et al. 1998), comprising its complete polypeptide sequence, a model for complete fXa light chain was performed using homology modeling techniques, employing MODELLER9v8 (Sanchez et al. 2000), as previously described (Mulinari et al. 2011). The modeling was carried out by combining templates from PDB ID 1WHE (Sunnerhagen et al. 1996) for residues 1–53 and from PDB ID 1XKA (Kamata et al. 1998) for 54-139. The selection and validation of the best model, from fifty constructed structures, was based on stereochemical evaluation with PROCHECK (Laskowski et al. 1993), while additional validations with Verify3D (Lüthy et al. 1992) and QMEAN6 (Benkert et al. 2011) in the "Structure Assessment" tool of SwissModel (Arnold et al. 2006) were also employed (Supplementary note 1; Supplementary data, Figure S5). The complete fXa structure was then obtained by superimposing the obtained model for fXa light chain (residues 1–139) to that previously contained within PDB ID 1XKA (residues 54–139).

Building the initial models for heparin–proteases complexes. A heparin fragment containing 10 carbohydrate residues was built based on the most prevalent geometries obtained from solution MD simulations of heparin-composing disaccharides (Supple-mentary data, Figure S6A), as previously determined (Pol-Fachin and Verli 2008; Pol-Fachin et al. 2011). Such polysaccharide sequence was constructed aiming to mimic the known monosaccharide proportions in the regions involved with AT binding to heparin (Guerrini et al. 2008) as, in ternary complexes involving the studied enzymes and AT, the constructed heparin sequence would be vicinal to the GAG region interacting with the serpin. Subsequently, such a heparin molecule was complexed to both fIIa and fXa through superimposition to the oligosaccharide orientation as observed in PDB ID 1TB6 (Li et al. 2004). Modeling the fXa–GAG complex based on the structure within PDB ID 2GD4 (Johnson et al. 2006), which contains an AT-fXa-pentasacharide complex, was not possible (1) as no direct fXa–heparin interaction is observed in such a structure and (2) because, while superimposing a long-chain heparin to the pentasaccharide, the GAG portion near the enzyme remained away from fXa exosite 2. The obtained complexes fulfill heparin orientation on the surface of thrombin (Supplementary data, Figure S6), as observed on previous X-ray crystallography studies (Li et al. 2004; Carter et al. 2005). The validation of heparin dynamics during MD was performed by comparing the GAG glycosidic linkages conformation with NMR-derived models (Mulloy et al. 1993), X-ray structures (Khan et al. 2010) and a free heparin MD simulation (Supplementary note 2; Supple-mentary data, Table S5).

MD simulations

Each non-glycosylated and glycosylated thrombin structure, as well as the modeled fXa in the presence and absence of Ca^{2+} ions, was generated either in an uncomplexed or in a heparin-complexed state. Globally, eight systems were simulated up to 0.1 μs: (1) non-glycosylated thrombin, heparin-absent, (2) glycosylated thrombin, heparin-absent, (3) non-glycosylated thrombin, complexed to heparin, (4) glycosylated thrombin, complexed to heparin, (5) fXa, heparin-absent, system without Ca^{2+} ions, (6) fXa, heparin-absent, system with Ca^{2+} ions, (7) fXa, complexed to heparin, system without Ca^{2+} ions, and (8) fXa, complexed to heparin, system with Ca^{2+} ions. Additionally, a ninth system, comprising a free 0.1 μs heparin MD simulation, was also performed for reference. Such structures were solvated in triclinic boxes using periodic boundary conditions and SPC water model (Berendsen et al. 1987), employing a 10 Å distance from the outside of each structure and the box edge. In the systems in which the presence Ca^{2+} was considered, such divalent cation was added in its known physiological concentration. After that, in order to neutralize the charge of the eight simulated systems, which were highly negative, due to heparin sulfate and sulfonamide groups, counter ions (Na^+ or Cl^-) were added. The MD protocol, including general simulation parameters and thermalization procedures, was employed on previous studies, as described for other heparin–protein complexes (Verli and Guimarães 2005; Pol-Fachin et al. 2011). The Lincs method (Hess et al. 1997) 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 through the Particle Mesh Ewald method (Darden et al. 1993). Temperature and pressure were kept constant by coupling (glyco)proteins, heparin, ions and solvent to external temperature and pressure baths with coupling constants of $\tau = 0.1$ and 0.5 ps (Berendsen et al. 1984), respectively. The systems were heated slowly from 50 to 310 K, in steps of 5 ps, each one increasing the reference temperature by 50 K. In the case of proteases–heparin complexes, before data collection, 10 ns MD simulations were also employed as equilibration steps, as previously described (Pol-Fachin et al. 2011). After this heating/equilibration, all simulations were further simulated to 0.1 μs under a constant temperature of 310 K.

Dynamic network analysis

The general idea of a dynamic network analysis is to assess interaction networks within a system, usually assuming that each amino acid residue is a node, and the contacts between them are represented by links (Sethi et al. 2009). In this context, a weight is given to those connections, depending on the intensity of these interactions. Of interest to the present work, a community of amino acids is established when the nodes representing the amino acid residues within that group have a higher number or stronger links to each other. Such concept can be compared to a structural domain in a polypeptide (Eargle and Luthey-Schulten 2012), but here it is defined by the protein dynamics, and not by its tertiary structure.

Supplementary data

Supplementary data for this article are available online at http://glycob.oxfordjournals.org/.

Funding

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) from the Ministério de

Ciência e Tecnologia; by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) from the Ministério da Educação e Cultura, Brasília, DF, Brazil; and by Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), Porto Alegre, RS, Brazil.

Conflict of interest

None.

Abbreviations

3D, three-dimensional; AT, antithrombin; BHD, β-hydroxyaspartate; fIIa, thrombin; fXa, factor Xa; GAG, glycosaminoglycan; GLA, γ-carboxyglutamic acid; MD, molecular dynamics; PCA, principal component analysis; PDB, Protein Data Bank; RMSF, root mean square fluctuation

References

Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: A web-based environment for protein structure homology modeling. *Bioinformatics*. 22:195–201.

Benkert P, Biasini M, Schwede T. 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*. 27:343–350.

Berendsen HJC, Grigera JR, Straatsma TP. 1987. The missing term in effective pair potentials. *J Phys Chem*. 91:6269–6271.

Berendsen HJC, Postma JPM, DiNola A, Haak JR. 1984. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 81:3684–3690.

Bode W, Mayr I, Baumann U, Huber R, Stone SR, Hofsteenge J. 1989. The refined 1.9 A crystal structure of human alpha-thrombin: Interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J*. 8:3467–3475.

Bode W, Turk D, Karshikov A. 1992. The refined 1.9-A X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human alpha-thrombin: Structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships. *Protein Sci*. 1:426–471.

Carter WJ, Cama E, Huntington JA. 2005. Crystal structure of thrombin bound to heparin. *J Biol Chem*. 280:2745–2749.

Choay J, Petitou M, Lormeau J-C, Sinay P, Casu B, Gatti G. 1983. 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*. 116:492–499.

Craig PA, Olson ST, Shore JD. 1989. Transient kinetics of heparin-catalyzed protease inactivation by antithrombin III. Characterization of assembly, product formation, and heparin dissociation steps in the factor Xa reaction. *J Biol Chem*. 264:5452–5461.

Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J Chem Phys*. 98:10089–10092.

Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. 2006. CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res*. 34: W116–W118.

Eargle J, Luthey-Schulten Z. 2012. Networkview: 3D display and analysis of protein-RNA interaction networks. *Bioinformatics*. 28:3000–3001.

Fernandes CL, Sachett LG, Pol-Fachin L, Verli H. 2010. GROMOS96 43a1 performance in predicting oligosaccharide conformational ensembles within glycoproteins. *Carbohydr Res*. 345:663–671.

Fernández PV, Quintana I, Cerezo AS, Caramelo JJ, Pol-Fachin L, Verli H, Estevez JM, Ciancia M. 2013. Anticoagulant activity of a unique sulfated pyranosic (1→3)-β-L-arabinan through direct interaction with thrombin. *J Biol Chem*. 288:223–233.

Fredenburgh JC, Stafford AR, Weitz JI. 1997. Evidence for allosteric linkage between exosites 1 and 2 of thrombin. *J Biol Chem*. 272:25493–25499.

Fuglestad B, Gasper PM, Tonelli M, McCammon JA, Markwick PRL, Komives EA. 2012. The dynamic structure of thrombin in solution. *Biophys J*. 103:79–88.

Furie B, Furie BC. 1988. The molecular basis of blood coagulation. *Cell*. 33:505–518.

Gandhi NS, Mancera RL. 2008. The structure of glycosaminoglycans and their interactions with proteins. *Chem Biol Drug Des*. 72:455–482.

Gettins PGW. 2002. Serpin structure, mechanism, and function. *Chem Rev*. 102:4751–4803.

Guerrini M, Guglieri S, Casu B, Torri G, Mourier P, Boudier C, Viskov C. 2008. Antithrombin-binding octasaccharides and role of extensions of the active pentasaccharide sequence in the specificity and strength of interaction. Evidence for very high affinity induced by an unusual glucuronic acid residue. *J Biol Chem*. 283:26662–26675.

Hansen DF, Kay LE. 2011. Determining valine side-chain rotamer conformations in proteins from methyl ^{13}C chemical shifts: Application to the 360 kDa half-proteasome. *J Am Chem Soc*. 133:8272–8281.

Hansson K, Stenflo J. 2005. Post-translational modifications in proteins involved in blood coagulation. *J Thromb Haemost*. 3:2633–2648.

Henry BL, Monien BH, Bock PE, Desai UR. 2007. A novel allosteric pathway of thrombin inhibition: Exosite II mediated potent inhibition of thrombin by chemo-enzymatic, sulfated dehydropolymers of 4-hydroxycinnamic acids. *J Biol Chem*. 282:31891–31899.

Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. 1997. LINCS: A linear constraint solver for molecular simulations. *J Comput Chem*. 18:1463–1472.

Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual molecular dynamics. *J Mol Graphics*. 14:33–38.

Huntington JA, Read RJ, Carrell RW. 2000. Structure of a serpin-protease complex shows inhibition by deformation. *Nature*. 407:923–926.

International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology Joint Commission on Biochemical Nomenclature (IUPAC). 1996. Nomenclature of carbohydrates. *Pure Appl Chem*. 68:1919–2008.

Johnson DJD, Li W, Adams TE, Huntington JA. 2006. Antithrombin–S195A factor Xa-heparin structure reveals the allosteric mechanism of antithrombin activation. *EMBO J*. 25:2029–2037.

Kamata K, Kawamoto H, Honma T, Iwama T, Kim SH. 1998. Structural basis for chemical inhibition of human blood coagulation factor Xa. *Proc Natl Acad Sci USA*. 95:6630–6635.

Khan S, Gor J, Mulloy B, Perkins SJ. 2010. Semi-rigid solution structures of heparin by constrained X-ray scattering modelling: New insight into heparin-protein complexes. *J Mol Biol*. 395:504–521.

Lane DL, Denton J, Flynn AM, Thunberg L, Lindahl U. 1984. Anticoagulant activities of heparin oligosaccharides and their neutralization by platelet factor 4. *Biochem J*. 218:725–732.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallog*. 26:283–291.

Laurent TC, Tengblad A, Thunberg L, Höök M, Lindahl U. 1978. The molecular-weight-dependence of the anti-coagulant activity of heparin. *Biochem J*. 175:691–701.

Lawson JH, Kalafatis M, Stram S, Mann KG. 1994. A model for the tissue factor pathway to thrombin. I. An empirical study. *J Biol Chem*. 269:23357–23366.

Li W, Johnson DJD, Esmon CT, Huntington JA. 2004. Structure of the antithrombin–thrombin–heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nat Struct Mol Biol*. 11:857–862.

Lüthy R, Bowie JU, Eisenberg D. 1992. Assessment of protein models with three-dimensional profiles. *Nature*. 356:83–85.

Lütteke T, Bohne-Lang A, Loss A, Goetz T, Frank M, von der Lieth CW. 2006. GLYCOSCIENCES.De: An Internet portal to support glycomics and glycobiology research. *Glycobiology*. 16:71R–81R.

Mulinari F, Becker-Ritt AB, Demartini DR, Ligabue-Braun R, Stanisçuaski F, Verli H, Fragoso RR, Schroeder EK, Carlini CR, Grossi-de-Sá MF. 2011. Characterization of JBURE-IIb isoform of Canavalia ensiformis (L.) DC urease. *Biochim Biophys Acta*. 1814:1758–1768.

Mulloy B, Forster MJ, Jones C, Davies DB. 1993. N.m.r. and molecular- modelling studies of the solution conformation of heparin. *Biochem J*. 293:849–858.

Nader HB, Pinhal MA, Baú EC, Castro RA, Medeiros GF, Chavante SF, Leite EL, Trindade ES, Shinjo SK, Rocha HA, et al. 2001. Development of new heparin-like compounds and other antithrombotic drugs and their interaction with vascular endothelial cells. *Braz J Med Biol Res*. 34:699–709.

Nilsson B, Horne MKr, Gralnick HR. 1983. The carbohydrate of human thrombin: Structural analysis of glycoprotein oligosaccharides by mass spectrometry. *Arch Biochem Biophys*. 224:127–133.

O'Brien LA, Stafford AR, Fredenburgh JC, Weitz JI. 2003. Glycosaminoglycans bind factor Xa in a Ca2+-dependent fashion and modulate its catalytic activity. *Biochemistry*. 42:13091–13098.

Olson ST, Swanson R, Raub-Segall E, Bedsted T, Sadri M, Petitou M, Herault J-P, Herbert J-M, Björk I. 2004. Accelerating ability of synthetic oligosaccharides on antithrombin inhibition of proteinases of the clotting and fibrinolytic systems. Comparison with heparin and low-molecular-weight heparin. *Thromb Haemostasis*. 92:929–939.

Oosta GM, Gardner WT, Beeler DL, Rosenberg RD. 1981. Multiple functional domains of the heparin molecule. *Proc Natl AcadSci USA*. 78:829–833.

Pol-Fachin L, Becker CF, Guimarães JA, Verli H. 2011. Effects of glycosylation on heparin binding and antithrombin activation by heparin. *Proteins*. 79:2735–2745.

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

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, Verli H. 2011. Assessment of glycoproteins dynamics from computer simulations. *Mini Rev Org Chem*. 8:229–238.

Rand MD, Lock JB, vant Veer C, Gaffney DP, Mann KG. 1996. Blood clotting in minimally altered whole blood. *Blood*. 88:3432–3445.

Rau JC, Beaulieu LM, Huntington JA, Church FC. 2007. Serpins in thrombosis, hemostasis and fibrinolysis. *J Thromb Haemost*. 5:102–115.

Rezaie AR. 1998. Calcium enhances heparin catalysis of the antithrombin-factor Xa reaction by a template mechanism. Evidence that calcium alleviates GLA domain antagonism of heparin binding to factor Xa. *J Biol Chem*. 273:16824–16827.

Rezaie AR. 2000. Identification of basic residues in the heparin-binding exosite of factor Xa critical for heparin and factor Va binding. *J Biol Chem*. 275:3320–3327.

Rezaie AR, Olson ST. 2000. Calcium enhances heparin catalysis of the antithrombin-factor Xa reaction by promoting the assembly of an intermediate heparin-antithrombin-factor Xa bridging complex. Demonstration by rapid kinetics studies. *Biochemistry*. 39:12083–12090.

Rosenfeld L, Danishefsky I. 1984. Effects of enzymatic deglycosylation on the biological activities of human thrombin and antithrombin. *Arch Biochem Biophys*. 229:359–367.

Sanchez R, Pieper U, Mirkovic N, de Bakker PIW, Wittenstein E, Sali A. 2000. Modbase, a database of annotated comparative protein structure models. *Nucleic Acids Res*. 28:250–253.

Schaftenaar G, Noordik JH. 2000. Molden: A pre- and post-processing program for molecular and electronic structures. *J Comput Aided Mol Des*. 14:123–134.

Schuettelkopf AW, van Aalten DMF. 2004. PRODRG: A tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallogr Sect D*. 60:1355–1363.

Scott WRP, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Krüger P, van Gunsteren WF. 1999. The GROMOS biomolecular simulation program package. *J Phys Chem A*. 103:3596–3607.

Sethi A, Eargle J, Black AA, Luthey-Schulten Z. 2009. Dynamical networks in tRNA: Protein complexes. *Proc Natl Acad Sci USA*. 106:6620–6625.

Sheehan JP, Sadler JE. 1994. Molecular mapping of the heparin-binding exosite of thrombin. *Proc Natl Acad Sci USA*. 91:5518–5522.

Sunnerhagen M, Olah GA, Stenflo J, Forsen S, Drakenberg T, Trewhella J. 1996. The relative orientation of Gla and EGF domains in coagulation factor X is altered by Ca2+ binding to the first EGF domain. A combined NMR-small angle X-ray scattering study. *Biochemistry*. 35:11547–11559.

The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.

van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. 2005. GROMACS: Fast, flexible, and free. *J Comput Chem*. 26: 1701–1718.

Verli H, Guimarães JA. 2004. Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. *Carbohydr Res*. 339: 281–290.

Verli H, Guimarães JA. 2005. Insights into the induced fit mechanism in antithrombin-heparin interaction using molecular dynamics simulations. *J Mol Graph Model*. 24:203–212.

4.5 Trabalho IV

Este trabalho propõe um novo conjunto de parâmetros (GROMOS 53A6$_{GLYC}$) para o estudo de carboidratos utilizando campos de força da série GROMOS. O conjunto de parâmetros até então distribuído para carboidratos, GROMOS 45A4, é capaz de descrever adequadamente a geometria de ligações glicosídicas de dissacarídeos, o efeito *gauche* (que descreve a orientação do grupamento hidroximetil, localizado na posição C6 de aldohexopiranose) e o efeito *exo*-anomérico em monossacarídeos. No entanto, baseado em estudos prévios da literatura, a conformação esperada para determinados resíduos (avaliados através de coordenadas *puckering*) não era reproduzida. Nesse sentido, os termos de energia potencial para diedros componentes do anel interno desses resíduos foram testados e, caso necessário, aprimorados, visando à descrição dos estados conformacionais supracitados, não descritos pelo campo de força antigo.

A partir dos estudos realizados, observou-se que os potenciais utilizados na descrição de diedros envolvendo C-C-C-O precisavam ser modificados, a partir dos quais dois novos termos de energia foram propostos. A utilização desses novos potenciais manteve a reprodução da geometria de ligações glicosídicas de dissacarídeos, dos efeitos *gauche* e *exo*-anomérico, e permitiu a reprodução, de forma qualitativa, das conformações esperadas para as dezesseis aldohexopiranoses avaliadas. Tais parâmetros, inseridos em um novo campo de força, intitulado GROMOS 53A6$_{GLYC}$, possuem adicionalmente ampla adaptabilidade a programas de simulação de dinâmica molecular e possibilidade de utilização em conjunto com parâmetros pré-existentes para proteínas, lipídeos ou ácidos nucleicos.

GROMOS 53A6$_{GLYC}$, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates

Laércio Pol-Fachin, Victor Holanda Rusu, Hugo Verli, Roberto Dias Lins Neto

Journal of Chemical Theory and Computation, **2012**, *8*; 4681-4690

JCTC Journal of Chemical Theory and Computation Article
pubs.acs.org/JCTC

GROMOS 53A6$_{GLYC}$, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates

Laercio Pol-Fachin,[†] Victor H. Rusu,[‡] Hugo Verli,*[,†,§] and Roberto D. Lins*[,‡]

[†]Center of Biotechnology, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil
[‡]Department of Fundamental Chemistry, Federal University of Pernambuco, Recife, PE, Brazil
[§]College of Pharmacy, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil

ABSTRACT: An improved parameter set for explicit-solvent simulations of carbohydrates (referred to as GROMOS 53A6$_{GLYC}$) is presented, allowing proper description of the most stable conformation of all 16 possible aldohexopyranose-based monosaccharides. This set includes refinement of torsional potential parameters associated with the determination of hexopyranose rings conformation by fitting to their corresponding quantum-mechanical profiles. Other parameters, as the rules for third and excluded neighbors, are taken directly from the GROMOS 53A6 force field. Comparisons of the herein presented parameter set to our previous version (GROMOS 45A4), the GLYCAM06 force field, and available NMR data are presented in terms of ring puckering free energies, conformational distribution of the hydroxymethyl group, and glycosidic linkage geometries for 16 selected monosaccharides and eight disaccharides. The proposed parameter modifications have shown a significant improvement for the above-mentioned quantities over the two tested force fields, while retaining full compatibility with the GROMOS 53A6 and 54A7 parameter sets for other classes of biomolecules.

INTRODUCTION

In recent years, glycobiology has become a critical facet of postgenomic science,[1] as carbohydrates have been associated with several biological events, from participating in large spatial or temporal scale processes, such as immune defense and cellular growth,[2,3] to influencing minor but complex properties of biomolecules, such as those related to protein glycosylation.[4] Such a broad range of functions achieved by carbohydrates is possibly related to their great structural diversity. In contrast to most nucleic acids and proteins, which are linear and have a unique type of linkage, glycans can be branched, linked through one of two anomeric configurations[5] through different atoms of their monomeric units, monosaccharides. Proper understanding of carbohydrate roles over biological systems requires an appropriate characterization, at the atomic level, of their conformation and dynamics conveyed by means of experimental and/or theoretical techniques.

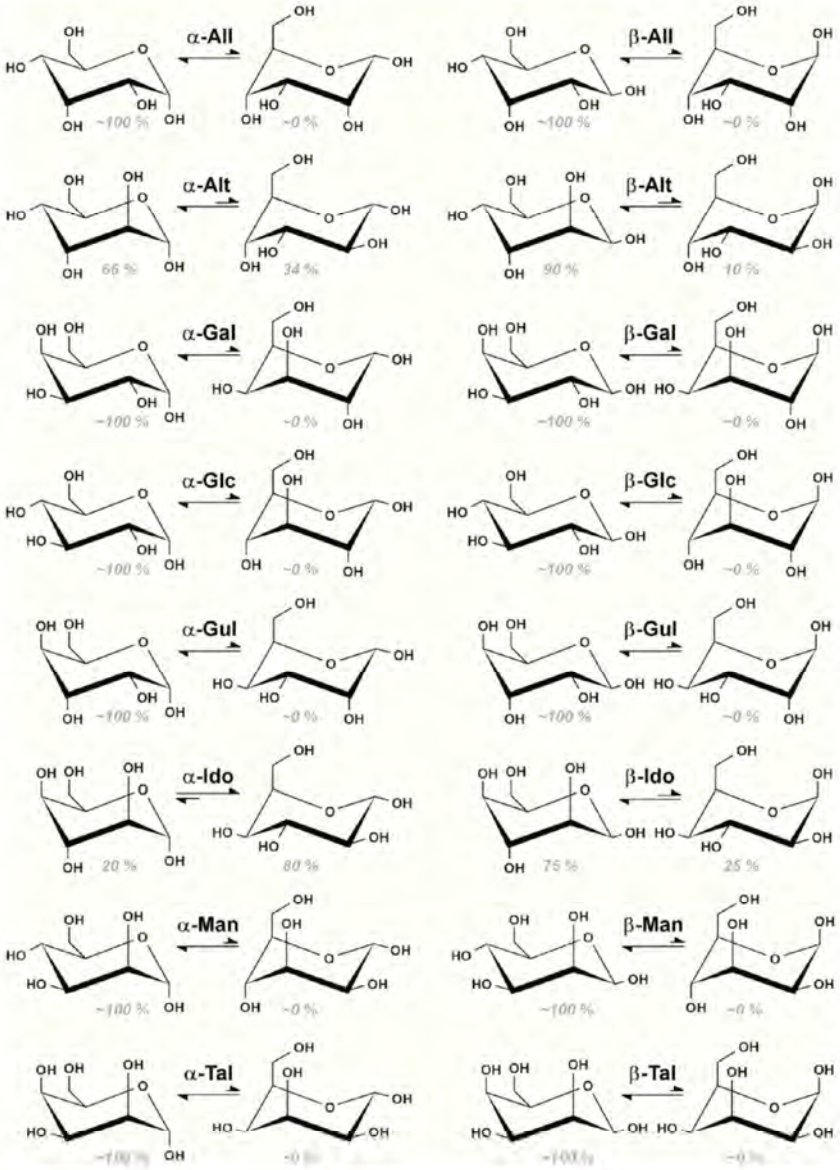
Among the experimental methods, X-ray crystallography usually provides the most complete description of a structure, once a crystal is obtained for the system under study.[5] However, due to the flexible nature of oligosaccharides, together with the lack of strong lipophilic or dipolar interresidue interactions and their high degree of coordination to water molecules,[6] carbohydrates are usually resistant to crystallization. Alternatively, NMR spectroscopy, based on the nuclear Overhauser effect (NOE), is frequently employed for obtaining or validating carbohydrate conformational profiles.[7−9] A caveat is that atomic coordinates are not provided, but a set of solution-averaged spatial constraints that restrict the number of possible conformations is.[5] In addition, NMR usually supplies a low number of NOE signals for carbohydrates,[6] which causes difficulties for completely determining glycan conformations. Other indirect experimental techniques, such as electron microscopy, light/neutron diffraction, or circular dichroism, have also been used to provide useful information about carbohydrates' three-dimensional organization.[10]

In this scenario, atomistic molecular dynamics (MD) simulations emerge as a powerful tool, providing data about carbohydrates typically inaccessible to experimental methods, with both atomic and temporal resolution.[5,6,11,12] The accuracy of such simulations is determined by the underlying computational model, including the accessible system size and time scale, the employed methodology, and parameters set.[12] In the context of the last property, among the parameter sets available for carbohydrates simulations in explicit solvent, some of the most used are included within the AMBER-GLYCAM,[13,14] CHARMM,[15−18] OPLS,[19,20] and GROMOS[21−23] force fields. Of interest to the present work, the GROMOS 45A4/53A6 force field for carbohydrates[21,24] has been shown to accurately describe a wide variety of polysaccharide properties[25−30] such as conformation,[28,29] dynamics,[25] and function.[27,30] However, it has been suggested that such parameters may present difficulties in reproducing internal ring puckering of a few residues and, to a lesser extent, the most stable form of specific aldohexopyranose monosaccharides.[31] On the basis of this information, modifications have been proposed by tuning torsional potentials by trial and error. However, a proper description of ring conformer populations was not achieved for some aldohexopyranoses.[31,32] Additionally, a reparameterized force field has been also suggested,[12] adequately describing aldohexopyranose preferred puckering conformations. The latter requires extensive modifications upon the existing parameters, i.e., new charges and atom types, several new torsional potentials, and nonstandard van der Waals scalings.[12]

Received: June 10, 2012
Published: September 6, 2012

dx.doi.org/10.1021/ct300479h | J. Chem. Theory Comput. 2012, 8, 4681−4690

Chart 1. Chemical Structures of the 16 Aldohexopyranoses Considered in the Present Study, in Their $^4C_1$ (left) and $^1C_4$ (right) Forms along with Their NMR Predicted Puckering Percentage Preferences



Considering that the improvement of glycans' parameter set should maintain compatibility with the general GROMOS53A6 force field for other biomolecules, that is, proteins, nucleic acids, and lipids, the present work intends to expand the GROMOS 45A4/53A6 force field for carbohydrates. As the previous studies have shown, the limitation of such parameters resided in the adequate description of the preferred aldohexopyranose puckering conformations. To overcome this issue, the torsional energy profiles of model compounds representing the ring atoms were derived to describe their quantum-mechanical (QM) profiles. On the basis of these data, two novel potentials are proposed, which compose the new version of the herein modified GROMOS force field, referred to as GROMOS 53A6$_{GLYC}$. Accordingly, such new parameters allow a proper description of the most stable conformation of hexopyranose-based monosaccharides, suitably enabling the study of saccharides either in solution or complexed to other classes of biomolecules, in a landscape of structures resembling the existing biological complexity for aldohexopyranoses.

METHODS

Nomenclature and Software. The nomenclature recommendations and symbols were used as proposed by IUPAC.[33] The QM calculations were performed with GAUSSIAN 98.[34] The equilibrium MD simulations were carried out with the standard version of GROMACS simulation suite version 4.5.1,[35] while metadynamics calculations[36] were performed using a modified version of GROMACS 4.5.1 interfaced with the PLUMED plugin package, version 1.2.2.[37] The free energy surfaces were obtained through the PLUMED package *sum_hills* tool. The relative orientation of a contiguous pair of carbohydrate residues is described by two or three torsional angles at the glycosidic linkage. For a (1→X) linkage, where 'X'

is '3', '4', or '6' for the (1→3), (1→4), or (1→6) linkages, respectively, the ϕ and ψ are defined as shown in eqs 1 and 2:

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

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

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

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

Comparatively, in unmodified monosaccharides, the ϕ, employed for evaluating the hydroxymethyl group orientation, is defined as shown below:

$$\varpi = O6 - C6 - C5 - O5 \quad (4)$$

Finally, for a (1→1) linkage, the ϕ and ϕ' are defined as shown below:

$$\phi = O5 - C1 - O1 - C1' \quad (5)$$

$$\phi' = C1 - O1 - C1' - O5' \quad (6)$$

QM Calculations. The QM torsional profiles for the rotations around C−C and C−O bonds, which comprise the dihedral angles determining pyranose ring conformations (Cx−Cx−Cx−Cx, Cx−Cx−O5−Cx, and Cx−Cx−Cx−Ox) were determined, respectively, for the heavy atoms of model compounds $CH_3-CH_2-CH_2-CH_3$, $CH_3-CH_2-O-CH_3$, and $CH_3-CH_2-CH_2-OH$ by increments of 30° at the HF/6-31G* level of theory. This level of theory was chosen to ensure compatibility with the previously derived torsional parameters for the force field. However, it is worth noting that calculations at the MP2/6-31G* and MP2/6-31G** levels of theory were also carried out and showed a negligible difference in the relative energies for the different conformers as calculated with HF/6-31G* (data not shown for conciseness). Each single-point calculation involved starting from the fully optimized conformation of each conformer, which were generated by rotating the selected dihedral angle to its target value, and then fully reoptimizing all other degrees of freedom. Subsequently, the same procedure was repeated at the molecular-mechanical level using fully optimized structures (conjugated-gradients) with a harmonic dihedral-angle restraining potential of 10 kJ mol^{-1} deg^{-2} force constant on each specific dihedral angle. The differences between QM and molecular-mechanical profiles were fitted by a cosine series, which served to determine whether changes in force field torsional parameters were required.

Molecular Systems. For metadynamics and unbiased MD simulations, 16 aldopyranoses were evaluated, comprised of D-allose (All), D-altrose (Alt), D-galactose (Gal), D-glucose (Glc), D-gullose (Gul), D-idose (Ido), D-mannose (Man), and D-talose (Tal) in both α and β anomeric states (Chart 1). Additionally, eight disaccharides were also studied, α-Glc−(1→1)-Glc (trehalose), β-Gal−(1→4)-Glc (lactose), α-Glc−(1→4)-Glc (maltose), β-Glc−(1→4)-Glc (cellobiose), α-Glc−(1→3)-Glc (nigerose), β-Glc−(1→3)-Glc (laminarabiose), α-Gal−(1→6)-Glc (melibiose), and β-Glc−(1→6)-Glc (gentiobiose). The systems consisted of each carbohydrate centered in a triclinic box, with each axis presenting 3.6 nm, filled with ca. 1512 SPC water molecules.[38] Monosaccharides were assessed under three different conditions: (1) described by the current GROMOS 45A4/53A6 parameters,[21] (2) the new GROMOS 53A6$_{GLYC}$ parameter set, and (3) employing the GLYCAM06 force field parameters.[14] The disaccharide units were only evaluated under condition 2, by the new GROMOS 53A6$_{GLYC}$ parameter set to assess their glycosidic linkage conformation in comparison to previous NMR data.[39−45]

MD Simulation. During all of the calculations, the Lincs method[46] was applied to constrain covalent bond lengths, allowing an integration step of 2 fs, while long-range electrostatic interactions were treated by the reaction field method[47] with ε = 66. A 1.2 nm cutoff was used for the short-range electrostatics and van der Waals interactions. The temperature was maintained at 298 K by coupling solute and solvent separately with Nosé−Hoover thermostats[48,49] with a relaxation time of 0.1 ps. Pressure was kept at 1 bar using a Parrinello−Rahman barostat[50,51] via an isotropic coordinate scaling with a coupling constant of 1.0 ps and a compressibility of 4.5×10^{-5} bar^{-1}. A 1 ns MD simulation was performed as an equilibration period and was not taken into account to calculate the average ensemble properties. Metadynamics calculations for the studied aldohexopyranoses consisted of 8 ns MD simulations, employing a height of 0.1 for the Gaussian height and a σ of 0.5 to each of the θ and ϕ angular coordinates of Cremer and Pople.[52] For the evaluated disaccharides, a 10 ns MD simulations was used, employing a height of 0.1 for the Gaussian height and a σ of 0.5 to each of the ϕ and ψ glycosidic linkage dihedral angles. Also, 1 μs MD simulations were performed employing unbiased MD simulations for all aldohexopyranoses and disaccharides studied.

■ RESULTS AND DISCUSSION

Force Field Parametrization. The atomic charges for carbohydrate residues (Table 1), the potentials for bond stretching, bond-angle bending, and improper dihedral deformation (Table 2), as well as van der Waals interactions terms were retrieved directly from the GROMOS 45A4/53A6 functional form for carbohydrates,[21,24] previously validated elsewhere.[21,53−55]

The functional form of the potential-energy term, associated with the stretching of bond m, is given by

$$V_{b,m} = (1/4)k_{b,m}[b_m^2 - b_{o,m}^2]^2 \quad (7)$$

and it is applied to all unique pairs of covalently linked atoms that match those specified in Table 2, where b_m is the bond-length distance, b_o its reference value, and $k_{b,m}$ the corresponding (quartic) force constant. The functional form of the potential-energy term, associated with the bending of bond angle m, is given by

$$V_{\theta,m} = (1/2)k_{\theta,m}[\cos\theta_m - \cos\theta_{o,m}]^2 \quad (8)$$

and it is applied to all unique triplets of covalently linked atoms that match those specified in Table 2, where θ_m is the bond-angle value, θ_o its reference value, and $k_{\theta,m}$ the corresponding (cosine-harmonic) force constant. The functional form of the potential-energy term, associated with the deformation of improper-dihedral angle m, is given by

$$V_{\xi,m} = (1/2)k_{\xi,m}[\xi_m - \xi_{m,o}]^2 \quad (9)$$

and it is only applied, for each hexopyranose monomer, to the subset of improper-dihedral angles specified in Table 2, where ξ_m is the improper-dihedral angle value, ξ_o its reference value, and $k_{\xi,m}$ the corresponding (harmonic) force constant. While some of the torsional potentials were preserved from the GROMOS 53A6 parameter set, those determining the conformation of the hexopyranose rings (Cx−Cx−Cx−Cx, Cx−Cx−O5−Cx, and Cx−Cx−Cx−Ox) were re-evaluated

Table 1. Atom Types, Atomic Partial Charges and Charge-Group Definitions of the New GROMOS 53A6$_{GLYC}$ Force Field for Hexopyranose-Based Carbohydrates[a,b]

atom	charge group	atom type	partial atomic charge
		4-OH initiation patch[c]	
O4	1	OA	−0.642
HO4	1	H	0.410
		monomeric unit	
C4	1	CH_1	0.232
C3	2	CH_1	0.232
O3	2	OA	−0.642
HO3	2	H	0.410
C2	3	CH_1	0.232
O2	3	OA	−0.642
HO2	3	H	0.410
C6	4	CH_2	0.232
O6	4	OA	−0.642
HO6	4	H	0.410
C5	5	CH_1	0.376
O5	5	OA	−0.480
		1-OH termination patch[d]	
C1	5	CH_1	0.232
O1	5	OA	−0.538
HO1	5	H	0.410
		(1→X) glycosidic linkage (X corresponding to 2, 3, 4, or 6)	
C1	5	CH_1	0.232
O1	5	OA	−0.360
Cx	5	CH_1	0.232
		1-OCH_3 termination patch[d]	
C1	5	CH_1	0.232
O1	5	OA	−0.360
C_{Me}	5	CH_3	0.232

[a]The charge set is reported in the context of a (1→4)-linked hexopyranose unit, but the table is easily transferable to other linkages. An example for building a (1→1) termination patch may be found elsewhere in the literature.[21] [b]The rules for excluded atoms and third-neighbors follow the GROMOS conventions,[53] except for an additional exclusion between the ring oxygen O5 and the lactol hydrogen HO1 in residues terminated by a 1-OH patch. [c]A monosaccharide molecule or the first residue in an unbranched saccharide must be initiated by a 4-OH initiation patch, with the inclusion of the shown atoms. [d]A monosaccharide molecule or the last residue in an unbranched saccharide must be terminated by a 1-OH or a 1-OCH3 termination patch.

based on fitting to QM data due to the so far reported difficulties of 53A6 parameters to reproduce ring puckering conformational properties. The functional form of the potential energy term, associated with the torsion around dihedral angle m, is given by

$$V_{\phi,m} = k_{\phi,m}[1 + \cos\delta_m \cos(n_m\phi_m)] \quad (10)$$

where ϕ_m is the dihedral angle value, n_m the multiplicity of the term, δ_m the associated phase shift, and $k_{\phi,m}$ the corresponding force constant, which are applied, for each hexopyranose monomer, to a subset of dihedral angles as specified in Table 3.

It is worth noting that a given dihedral angle may be involved in more than one torsional potential energy term with different multiplicities and/or phase shifts. Accordingly, the classical energy profiles obtained from their rotation were compared to energy profiles obtained from QM calculations, as presented in Figure 1. From such analyses, torsional parameters for Cx−Cx−Cx−Cx and Cx−Cx−O5−Cx dihedral angles, as presented in

Table 2. Bond Stretching, Bond-Angle Bending, and Improper-Dihedral Deformation Parameters Employed in GROMOS 53A6$_{GLYC}$ Force Field for Hexopyranose-Based Carbohydrates

bond type	k_b [10^6 kJ mol^{-1} nm^{-4}][a]	b_0 [nm]
C−C	5.43	0.152
C−O	6.10	0.144
C−H	15.70	0.100
bond-angle type	k_θ [kJ mol^{-1}][b]	θ_0 [deg]
C−C−C	285	109.5
C−C−O	320	109.5
O−C−O	320	109.5
C−O−C	380	109.5
C−O−H	450	109.5
improper dihedral (assuming a chair $^4C_1$ puckering)	k_ξ [kJ mol^{-1} deg^{-2}][c]	ξ_0 [deg]
C1−O1−O5−C2 (β-anomer)	0.102	35.2644
C1−O5−O1−C2 (α-anomer)	0.102	35.2644
C2−O2−C3−C1 (equatorial 2-OH)	0.102	35.2644
C2−C3−O2−C1 (axial 2-OH)	0.102	35.2644
C3−O3−C2−C4 (equatorial 3-OH)	0.102	35.2644
C3−C2−O3−C4 (axial 3-OH)	0.102	35.2644
C4−C3−O4−C5 (equatorial 4-OH)	0.102	35.2644
C4−O4−C3−C5 (axial 4-OH)	0.102	35.2644
C5−O5−C6−C4 (equatorial 6-CH_2OH)	0.102	35.2644
C5−C6−O5−C4 (axial 6-CH_2OH)	0.102	35.2644

[a]The parameters refer to eq 7. [b]The parameters refer to eq 8. [c]The parameters refer to eq 9.

the GROMOS 45A4/53A6 force field, are nearly identical to that from QM gas phase calculations (Figure 1A,B). On the other hand, important divergences between QM-calculated and 45A4/53A6 energy profiles for the Cx−Cx−Cx−Ox dihedral are observed (Figure 1C). Although accounting for two minimum-energy geometries, at 60°and 300°, the conformational barriers obtained for the current classical parameters are not sufficiently elevated to properly describe the rotation of such a dihedral. Additionally, the 180° conformer is described as the energy global maximum, instead of a local minimum (Figure 1C). On the basis of these data, two new torsional dihedral potentials associated with the rotation of the C−C−C−O dihedral angle were obtained by fitting the corresponding classical energy profiles to energy profiles obtained from QM calculations (Figure 1C). The resulting potentials, (1) a constant force of 5.88 kJ mol^{-1}, a phase shift of 0, and multiplicity 1 and (2) a constant force of 7.67 kJ mol^{-1}, phase shift of 0, and multiplicity of 3, were shown to adequately reproducing the QM-obtained energy profile related to such torsion. Consequently, those potentials were introduced to the torsional parameter set list (Table 3) for the GROMOS 53A6$_{GLYC}$ force field.

Hexopyranoses Puckering Evaluation through Pseudorotational Free Energy. Adaptative biasing or flat histogram methods have been successfully used to compare force field accuracy, including the study of sugar puckering profiles.[31,56,57] Similarly, here a series of metadynamics calculations[36] was performed in order to evaluate whether the new parameter set was capable of reproducing the conformational profile of aldohexopyranose-based carbohydrates. For comparison purposes, the current GROMOS 45A4/53A6 force field was evaluated. It shows the $^4C_1$ chair form as the most stable configuration for the 16 aldohexopyranoses (Table 4,

Table 3. Torsional Dihedral Interaction Parameters Used in the GROMOS 53A6$_{GLYC}$ Force Field for Hexopyranose-Based Carbohydrates

dihedral angle	k_ϕ [kH mol $^{-1}$][a]	δ	n
C1−C2−C3−C4	5.92	0	3
C2−C3−C4−C5	5.92	0	3
C3−C4−C5−C6	5.92	0	3
O4−C4−C3−O3	2.09	0	2
O3−C3−C2−O2	2.09	0	2
O2−C2−C1−O1	2.09	0	2
C2−C1−O5−C5	3.77	0	3
C1−O5−C5−C4	3.77	0	3
O1−C1−C2−C3	*5.88*	*0*	*1*
O1−C1−C2−C3	*7.67*	*0*	*3*
O2−C2−C3−C4	*5.88*	*0*	*1*
O2−C2−C3−C4	*7.67*	*0*	*3*
O3−C3−C2−C1	*5.88*	*0*	*1*
O3−C3−C2−C1	*7.67*	*0*	*3*
O3−C3−C4−C5	*5.88*	*0*	*1*
O3−C3−C4−C5	*7.67*	*0*	*3*
O4−C4−C3−C2	*5.88*	*0*	*1*
O4−C4−C3−C2	*7.67*	*0*	*3*
O4−C4−C5−C6	*5.88*	*0*	*1*
O4−C4−C5−C6	*7.67*	*0*	*3*
O5−C1−C2−C3	*5.88*	*0*	*1*
O5−C1−C2−C3	*7.67*	*0*	*3*
O5−C5−C4−C3	*5.88*	*0*	*1*
O5−C5−C4−C3	*7.67*	*0*	*3*
C1−C2−O2−HO2	3.90	0	3
C2−C3−O3−HO3	3.90	0	3
C3−C4−O4−HO4	3.90	0	3
C5−C6−O6−HO6	3.90	0	3
O5−C5−C6−O6[b]	9.50	0	3
O5−C5−C6−O6[b]	9.35	180	1
O5−C5−C6−O6[c]	7.69	0	3
O5−C5−C6−O6[c]	6.66	180	1
C4−C5−C6−O6[c]	2.67	180	1
O5−C1−O1−HO1 (α anomer)	3.65	0	3
O5−C1−O1−HO1 (α anomer)	9.45	180	1
O5−C1−O1−Ox (α anomer)	3.65	0	3
O5−C1−O1−Ox (α anomer)	9.45	180	1
O5−C1−O1−HO1 (β anomer)	4.69	0	3
O5−C1−O1−HO1 (β anomer)	3.41	180	1
O5−C1−O1−Ox (β anomer)	4.69	0	3
O5−C1−O1−Ox (β anomer)	3.41	180	1

[a]The parameters refer to eq 10. [b]Term to be used when the O4 and C6 atoms are on opposite sides of the ring plane (that is, equatorial−equatorial or axial−axial, as in glucose). [c]Term to be used when the O4 and C6 atoms are on the same side of the ring plane (that is, equatorial−axial or axial−equatorial, as in galactose).

first column). Proper description of ring pucker was obtained for all but one monosaccharide (α-Ido) using the GROMOS 45A4/53A6 force field. It is worth noting that Autieri and co-workers have reported the inability of this parameter set to properly describe the ring pucker of two monosaccharides, α-Ido and β-Tal.[31,32] Such discrepancy is likely due to the different simulation setups. Besides adoption of different values of σ and height, the present metadynamics simulations were carried out over an 8 ns period, in comparison with the previously published 4 ns time interval.[31] In contrast, the GROMOS 53A6$_{GLYC}$ has properly described the conformational preferences of all evaluated monosaccharides, that is, the $^1C_4$ conformation as the most stable ring puckering form for α-Ido and $^4C_1$ as the preferred state for the remaining aldohexopyranoses (Table 4, second column). In addition, a broader comparison with the GROMOS 53A6$_{GLYC}$ force field was also carried out by obtaining topology and coordinate files for GLYCAM06 from the GLYCAM Web site[14] and by converting them into GROMACS readable files with the *acpype* code.[58] Corresponding metadynamics simulations were carried out and the puckering free energies calculated. The latter differed from those performed with GROMOS 53A6$_{GLYC}$ only by the use of the TIP3P water model,[59] instead of SPC.[38] Our results show that GLYCAM06 parameters did not successfully describe the ring puckering preferential configurations for the α-anomers of three monosaccharides, All, Alt, and Gul (Table 4, third column). A common point is that they present an axial C-3 linked hydroxyl group. Although Ido is not available for study in the GLYCAM Web site,[14] a previous publication involving an Ido-derived monosaccharide, α-L-iduronic acid, also suggested that the GROMOS96 parameter set can better predict ring puckering conformational transitions in comparison to experimental data than GLYCAM06.[60] Altogether, based on the performed metadynamics calculations, the GROMOS 53A6$_{GLYC}$ force field was observed as capable of qualitatively reproducing the most stable ring puckering conformation of the 16 aldohexopyranose-based monosaccharides

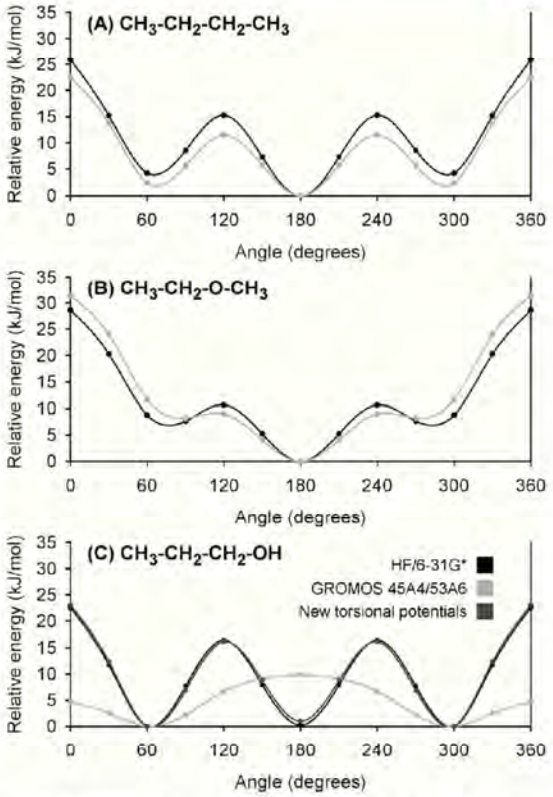


Figure 1. Comparison of energy profiles calculated at classical (45A4/53A6) and QM (HF/6-31G*) levels in the gas phase. In A−C, the energy profiles for the rotation of the heavy atoms composing the model compounds $CH_3-CH_2-CH_2-CH_3$, $CH_3-CH_2-O-CH_3$, and $CH_3-CH_2-CH_2-OH$ dihedral angles are respectively presented, as obtained from QM (black), the current torsional potentials (light gray), and the novel, derived torsional potentials (dark gray).

Table 4. Difference between of Relative Free Energies (kJ mol^{-1}) of $^4C_1$ and $^1C_4$ Puckering Conformations Associated with the Simulated Hexopyranoses, in Comparison with Chair Conformer Populations Obtained from NMR Experimental Data[a]

aldohexopyranose		GROMOS 53A6	GROMOS 53A6$_{GLYC}$	GLYCAM06	NMR $^4C_1/^1C_4$ (%/%)
All	α	−26	−14	**+31**	100:0
	β	−31	−46	−4	100:0
Alt	α	−28	−5	**+25**	66:34
	β	−21	−28	−7	90:10
Gal	α	−10	−40	−18	100:0
	β	−18	−54	−38	100:0
Glc	α	−18	−49	−11	100:0
	β	−31	−77	−41	100:0
Gul	α	−28	−6	**+25**	100:0
	β	−30	−20	−15	100:0
Ido	α	**−26**	+20	-[b]	20:80
	β	−25	−13	-[b]	75:25
Man	α	−11	−23	−6	100:0
	β	−11	−59	−31	100:0
Tal	α	−20	−2	−7	100:0
	β	−22	−40	−35	100:0

[a]The difference between the $^4C_1$ and $^1C_4$ chair conformations' relative free energy values are bolded when the most stable form of the aldohexopyranose monosaccharide is not properly described. [b]For GLYCAM06, the topology and coordinate files for α- and β-Ido were not available for download in the Web site "Carbohydrate 3D Structure Predictor" section.

ides evaluated, with better results when compared to the current GROMOS 45A4/53A6 and GLYCAM06 force fields.

Hexopyranoses Puckering Evaluation through Unrestrained MD. In order to reinforce the results obtained with the metadynamics approach, each studied monosaccharide was further evaluated during a 1 μs unrestrained MD simulation, described as the three parameter sets shown in Table 4, starting from each respective most stable conformation in solution ($^1C_4$ for α-Ido and $^4C_1$ for the remaining aldohexopyranoses). From these data, as shown in Figure 2, it may be observed that, for the new GROMOS 53A6$_{GLYC}$ force field, each monosaccharide remained in their expected ring puckering conformation, and no transitions between minor conformational states could be observed. However, from the cited NMR data, there should be a 2:1 ratio of $^4C_1$ to $^1C_4$ conformers. From the NMR distributions, one could estimate the $^4C_1 \rightarrow {}^1C_4$ conversion free energies (at 298.15 K) for α-Alt (66:34), β-Alt (90:10), α-Ido (20:80), and β-Ido (75:25) to be 1.7 kJ mol^{-1}, 22.5 kJ mol^{-1}, −3.5 kJ mol^{-1}, and 2.7 kJ mol^{-1}, respectively. The obtained values for α-Alt (5 kJ mol^{-1}) and β-Alt (28 kJ mol^{-1}) are near quantitative to the estimated free energies, while the most probable conformations for α-Ido and β-Ido seem overstabilized. These results suggest the possibility of an intrinsic overall stabilization of the most probable puckering conformation of β-aldohexopyranoses. Nevertheless, variations in NMR derived data are not uncommon, as it is shown in Table 5 for the preferred conformation of the hydroxymethyl group in β-aldohexopyranoses.

For the GROMOS 45A4/53A6 force field, in agreement with the metadynamics data, α-Ido was observed to converge from $^1C_4$ to a $^4C_1$ conformation at the beginning of the simulation, and to remain in such ring puckering state for the rest of the trajectory. Additionally, the monosaccharides' conformational behavior with the GLYCAM06 parameter set presented a more variable profile (Figure 2). In accordance with the metadynamics results, α-All, α-Alt, and α-Gul converged to the $^1C_4$ ring puckering form, and all the β-anomers remained during all trajectories in their $^4C_1$ state. The conformational stability of the GROMOS parameters are in contrast with GLYCAM generated ensembles, which have also shown multiple conformational states in previous publications.[60]

Hydroxymethyl Group Rotation. In comparison to the parameters set previously published for GROMOS 45A4/53A6 force field,[21] a discrepancy in one torsional potential related to the ϕ dihedral angle (O5−C5−C6−O6) is found in the force field distribution within the GROMACS package. For the monosaccharides presenting the O4 and C6 atoms on the same side of the ring plane (that is, Gal, Gul, Ido, and Tal), one of such dihedrals presents a reduced constant force of 4.97 kJ mol^{-1}, instead of the proposed 6.66 kJ mol^{-1}. In order to assess whether such modification would alter the conformational profile of the ϕ dihedral angle, we have analyzed its variation during the performed 1 μs MD simulations. It is verified that the use of the decreased constant force promotes the higher population on the *tg* rotamer in Gal. The increased constant force is required for a proper description of the ϕ angle of this monosaccharide, *i.e.*, a dominance of the *gt* rotamer in both anomers of Gal (Table 5), as well as similar populations of *tg* and *gg* rotamers, as shown by NMR data.[61,62] On the basis of the MD obtained data, among the remaining monosaccharides presenting the O4 and C6 atoms on the same side of the ring plane, α- and β-Tal are those showing the same proportions of rotamers of Gal (Table 5), while the $^1C_4$ conformation of α-Ido promoted an increased population on the *gt* rotamer, similarly to Gal, and a small population related to the *tg* rotamer. On the other hand, for both anomers of Gul and for β-Ido, the *gg* rotamer is dominant (Table 5). The data suggest that such a difference is a feature of axial O3-presenting monosaccharides (Gul and Ido), which must be experimentally confirmed. Moreover, as a robust sampling could be obtained for the dynamics related to such dihedral, we have also analyzed its conformational profile for the remaining monosaccharides, presenting the O4 and C6 atoms on opposite sides of the ring plane (All, Alt, Glc, and Man), although no torsional parameter related to ϕ was changed in the here presented GROMOS 53A6$_{GLYC}$ force field with respect to the GROMOS 45A4/53A6 parameter set. The relative populations for the *gg* (−60°), *gt* (60°), and *tg* (180°) conformers, as reported in Table 5, reproduce the small population related to the *tg* rotamer, as well as *gg* and *gt* rotamers with similar populations in Glc and Man, as observed from NMR-derived data.[61−66] Moreover, both anomers of All and Alt presented the same conformational profile.

Glycosidic Linkages Conformation. In order to assess whether the herein presented GROMOS 53A6$_{GLYC}$ force field is capable of properly describing the conformation of glycosidic linkage, we have performed 1 μs MD simulations in eight selected disaccharides presenting known glycosidic linkage conformations, as previously described by NMR methods.[39−45] As displayed in Figure 3, the simulations reproduced very well the glycosidic linkage conformations in all cases. Particularly, the employed force field was capable of properly reproducing the *exo*-anomeric effect, in which the ϕ dihedral angle of α-anomers (Figure 3A, C, E, and G) is more populated at positive geometries (~60°), while the β-anomers (Figure 3B, D, F, and H) may adopt two staggered conformations, at approximately

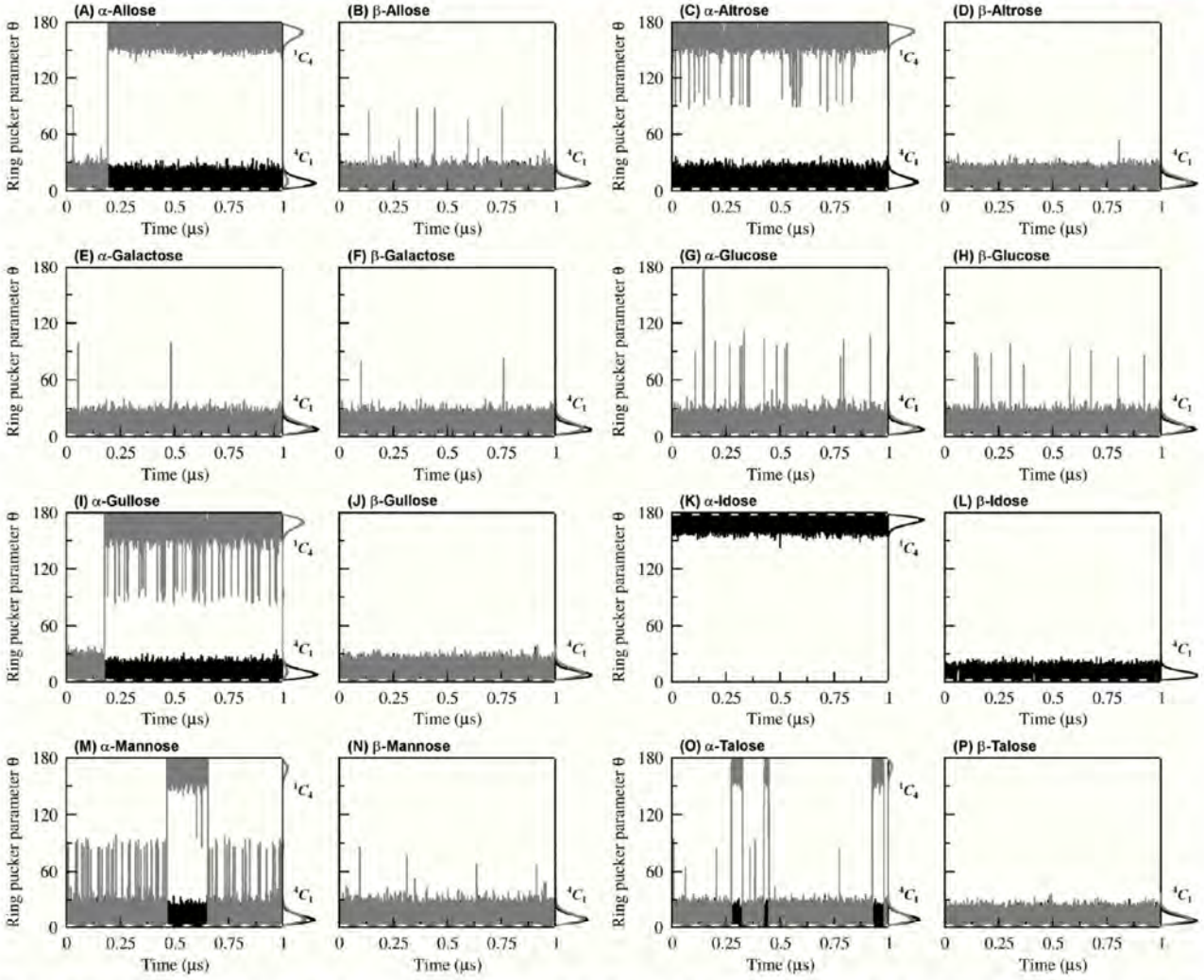


Figure 2. Fluctuations (curves) and distribution (side histograms) of the θ ring pucker parameter during the 1 μs MD simulations performed with the sixteen studied aldohexopyranoses, described as the new GROMOS 53A6$_{GLYC}$ (black) and GLYCAM06 (gray) force fields. θ values close to 0 indicate a $^4C_1$ conformation, while values close to 180 point to a $^1C_4$ state. In the graphs, white dashed lines indicate the most populated ring pucker conformational state for each aldohexopyranose, as observed from experimental data.

−60° and 60°, but mostly negative values.[67] Regarding the ω dihedral angle (O1−C6−C5−C4) of (1→6) linkages, only the three possible staggered conformations (−180°, −60°, and 60°) were populated, in accordance with Shefter and Trueblood's convention.[68] The GROMOS 53A6$_{GLYC}$ parameter set was also capable of adequately reproducing the *gauche* effect,[69] in which the −60° conformation was mostly not observed in solution.

CONCLUSIONS

In the current study, a new parameter set (GROMOS 53A6$_{GLYC}$) was presented, in which we have proposed two new torsional potentials to be included within the previous GROMOS 45A4/53A6 force field for the simulation of hexopyranose-based carbohydrates. The major advantage over previously proposed modifications of the force field is that this approach is still compatible with the general GROMOS parameter set for other classes of biomolecules. The addition of new torsional potentials, instead of largely altering atom types and van der Waals nonbonded interactions, is a similar approach to that recently performed for improving GROMOS parameters for proteins.[70,71] Moreover, the employment of a 6.66 kJ mol^{-1} force constant for a torsional potential parameter describing the ϕ dihedral angle (O5−C5−C6−O6) of monosaccharides presenting the O4 and C6 atoms on the same side of the ring plane, as Gal, was confirmed necessary.

The new GROMOS 53A6$_{GLYC}$ force field is capable of properly describing the most stable ring pucker conformation of a set of 16 aldohexopyranoses, showing more accurate results than the GROMOS 45A4/53A6 and an ACPYPE-converted version of GLYCAM06. The new parameter set was also able to adequately reproduce the glycosidic linkage conformation of eight different disaccharides tested, as well as the relative abundance of the hydroxymethyl group in Gal, Glc, and Man anomers, both based on comparisons of previous solution NMR data with 1 μs MD simulations achieved in the present study. Furthermore, as previously reported for the GROMOS 45A4/53A6 force field,[21] the GROMOS 53A6$_{GLYC}$ parameter set is suitable for MD simulation studies of any unbranched hexopyranose-based mono-, di-, oligo-, or polysaccharide. Nevertheless, as a continual upgrading is required to meet

Table 5. Relative Populations Associated with the Three Staggered Conformations of the Hydroxymethyl Group, As Obtained from 1 μs MD Simulations for the 16 Studied Aldohexopyranoses

aldohexopyranose		ϕ^a		
		−60° (*gg*)	60° (*gt*)	180° (*tg*)
All	α	62	38	0
	β	60	40	0
Alt	α	56	44	0
	β	53	47	0
Gal	α	34 (21,[b] 3[d]) [24[e]]	43 (54,[b] 74[d]) [30[e]]	23 (25,[b] 23[d]) [46[e]]
	β	32 (22,[b] 3[d]) [21[e]]	42 (53,[b] 72[d]) [30[e]]	26 (25,[b] 25[d]) [49[e]]
Glc	α	60 (56,[b] 47,[c] 40[d])	40 (44,[b] 54,[c] 53[d])	0 (0,[b] −1,[c] 7[d])
	β	55 (53,[b] 45,[c] 31[d])	45 (45,[b] 62,[c] 61[d])	0 (2,[b] −7,[c] 8[d])
Gul	α	48 [36[e]]	33 [25[e]]	19 [39[e]]
	β	45 [33[e]]	34 [26[e]]	21 [41[e]]
Ido	α	16 [15[e]]	80 [75[e]]	4 [10[e]]
	β	43 [31[e]]	37 [30[e]]	20 [39[e]]
Man	α	53	47	0
	β	49 (48[e])	51 (49[e])	0 (3[e])
Tal	α	38 [27[e]]	41 [31[e]]	21 [42[e]]
	β	35 [24[e]]	42 [30[e]]	23 [46[e]]

[a]The two letters in the notations *gt*, *tg*, and *gg* refer successively to the relative orientations of O6 relative to O5 (first letter, gauche or trans) and C4 (second letter, gauche or trans). The values between parentheses for Glc, Gal, and Man anomers correspond to relative populations derived from NMR experiments by [b]Nishida et al.[61,63,64] [c]Brochier−Salon and Morin.[65] [d]Thibaudeau et al.[62] [e]Hori et al.[66] Rotamer population using a force constant of 4.97 kJ mol^{-1}.

the demands of the scientific community, we are looking forward to further refinement of the parameter set, as well as its extension for a broader set of carbohydrates, such as acetylated sugar derivatives and furanoses.

AUTHOR INFORMATION

Corresponding Author

*E-mail: roberto.lins@ufpe.br.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was funded by CNPq, NanoBiotec-BR/CAPES, INCT-INAMI, FACEPE, nBioNet, and the Swedish Foundation for International Cooperation in Research and Higher Education (STINT). Part of the computational resources were provided by the Environmental Molecular Sciences Laboratory, a U.S. national scientific user facility sponsored by the U.S. Department of Energy located at the Pacific Northwest National Laboratory.

REFERENCES

(1) Turnbull, J. E.; Field, R. A. Emerging glycomics technologies. *Nat. Chem. Biol.* **2007**, *3*, 74−77.

(2) Varki, A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **1993**, *3*, 97−130.

(3) Dwek, R. A. Glycobiology: toward understanding the function of sugars. *Chem. Rev.* **1996**, *96*, 683−720.

(4) Lowe, J. B.; Marth, J. D. A genetic approach to mammalian glycan function. *Annu. Rev. Biochem.* **2003**, *72*, 643−691.

(5) 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.

(6) Woods, R. J. Computational carbohydrate chemistry: what theoretical methods can tell us. *Glycoconjugate J.* **1998**, *15*, 209−216.

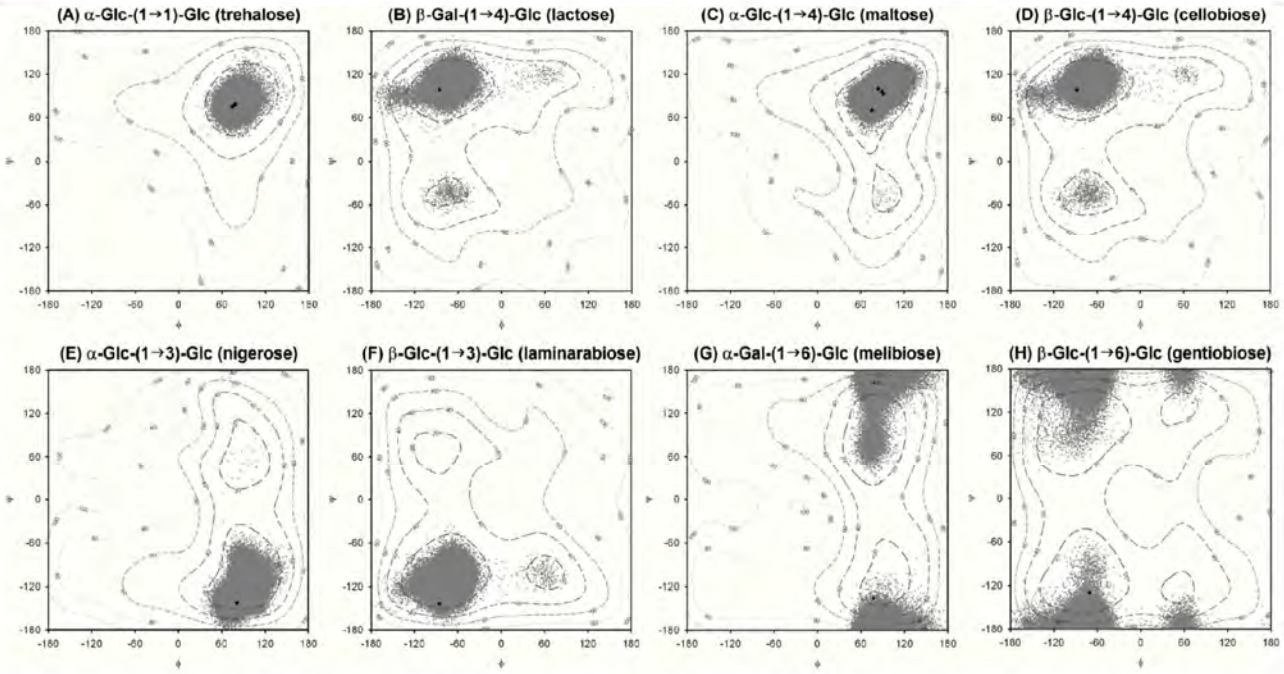


Figure 3. Contour plots for the glycosidic linkages of eight disaccharides, as obtained from metadynamics calculations. Free energy maps are shown at every 20 kJ mol^{-1}, from 20 to 100 kJ mol^{-1}. The corresponding solution profiles for the disaccharides from 1 μs MD simulations are presented as gray dots at every 10 ps interval. Black dots indicate experimental glycosidic linkage geometries obtained from NMR.[39−45]

(7) 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.
(8) Bubb, W. NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concepts Magn. Reson., Part A* **2003**, *19A*, 1−13.
(9) Blanchard, V.; Chevalier, F.; Imberty, A.; Leeflang, B. R.; Basappa; Sugahara, K.; Kamerling, J. P. Conformational studies on five octasaccharides isolated from chondroitin sulfate using NMR spectroscopy and molecular modeling. *Biochemistry* **2007**, *46*, 1167−1175.
(10) Pérez, S.; Kouwijzer, M. In *Carbohydrates: Structures, Dynamics and Syntheses*; Finch, P., Ed.; Kluwer Academic Press: Dordrecht, The Netherlands, 1999; pp 258−293.
(11) Imberty, A.; Pérez, S. Structure, Conformation, and Dynamics of Bioactive Oligosaccharides: Theoretical Approaches and Experimental Validations. *Chem. Rev.* **2000**, *100*, 4567−4588.
(12) Hansen, H. S.; Hünenberger, P. H. A Reoptimized GROMOS Force Field for hexopyranose-Based Carbohydrates Accounting for the Relative Free Energies of Ring Conformers, Anomers, Epimers, Hydroxymethyl Rotamers, and Glycosidic Linkage Conformers. *J. Comput. Chem.* **2011**, *32*, 998−1032.
(13) 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.
(14) 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.
(15) Kuttel, M.; Brady, J. W.; Naidoo, K. J. Carbohydrate solution simulations: Producing a force field with experimentally consistent primary alcohol rotational frequencies and populations. *J. Comput. Chem.* **2002**, *23*, 1236−1243.
(16) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. J. Additive empirical force field for hexopyranose monosaccharides. *J. Comput. Chem.* **2008**, *29*, 2543−2564.
(17) Guvench, O.; Hatcher, E. R.; Venable, R. M.; Pastor, R. W.; Mackerell, A. D. J. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353−2370.
(18) Raman, E. P.; Guvench, O.; Mackerell, A. D. J. CHARMM Additive All-Atom Force Field for Glycosidic Linkages in Carbohydrates Involving Furanoses. *J. Phys. Chem. B* **2010**, *114*, 12981−12994.
(19) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. OPLS all-atom force field for carbohydrates. *J. Comput. Chem.* **1997**, *18*, 1955−1970.
(20) Kony, D.; Damm, W.; Stoll, S.; van Gunsteren, W. F. An improved OPLS-AA force field for carbohydrates. *J. Comput. Chem.* **2002**, *23*, 1416−1429.
(21) Lins, R. D.; Hünenberger, P. H. New GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Comput. Chem.* **2005**, *26*, 1400−1412.
(22) Verli, H.; Guimarães, J. A. Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution. *Carbohydr. Res.* **2004**, *339*, 281−290.
(23) Pol-Fachin, L.; Fernandes, C. L.; Verli, H. GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations. *Carbohydr. Res.* **2009**, *344*, 491−500.
(24) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(25) Pereira, C. S.; Kony, D.; Baron, R.; Müller, M.; van Gunsteren, W. F.; Hünenberger, P. H. Conformational and Dynamical Properties of Disaccharides in Water: a Molecular Dynamics Study. *Biophys. J.* **2006**, *90*, 4337−4344.
(26) Franca, E. F.; Lins, R. D.; Freitas, L. C. G.; Straatsma, T. P. Characterization of Chitin and Chitosan Molecular Structure in Aqueous Solution. *J. Chem. Theory Comput.* **2008**, *4*, 2141−2149.
(27) Horta, B. A. C.; Peric-Hassler, L.; Hünenberger, P. H. Interaction of the disaccharides trehalose and gentiobiose with lipid bilayers: A comparative molecular dynamics study. *J. Mol. Graphics Modell.* **2010**, *29*, 331−346.
(28) Horta, B. A. C.; Peric-Hassler, L.; Hünenberger, P. H. Interaction of the disaccharides trehalose and gentiobiose with lipid bilayers: A comparative molecular dynamics study. *J. Mol. Graphics Modell.* **2010**, *29*, 331−346.
(29) Franca, E. F.; Freitas, L. C. G.; Lins, R. D. Chitosan Molecular Structure as a Function of N-Acetylation. *Biopolymers* **2011**, *95*, 448−460.
(30) Fedorov, M. V.; Goodman, J. M.; Neruck, D.; Schumm, S. Self-assembly of trehalose molecules on a lysozyme surface: the broken glass hypothesis. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2294−2299.
(31) Autieri, E.; Sega, M.; Pederiva, F.; Guella, G. Puckering free energy of pyranoses: a NMR and metadynamics-umbrella sampling investigation. *J. Chem. Phys.* **2010**, *133*, 095104.
(32) Autieri, E.; Sega, M.; Pederiva, F.; Guella, G. Erratum: "Puckering free energy of pyranoses: a NMR and metadynamics-umbrella sampling investigation" [J. Chem. Phys. 133, 095104 (2010)]. *J. Chem. Phys.* **2011**, 134.
(33) McNaught, A. D. International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology - Joint Commission on Biochemical Nomenclature - Nomenclature of carbohydrates - Recommendations 1996. *Pure Appl. Chem.* **1996**, *68* (10), 1919−2008.
(34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Rega, N.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Laham, A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Gordon, H.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.11.4.; Gaussian, Inc.: Pittsburgh, PA, 2002.
(35) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2009**, *4*, 435−447.
(36) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *WIREs Comput. Mol. Sci.* **2011**, *1*, 826−843.
(37) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: a portable plugin for free energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.
(38) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(39) Cheetham, N. W. H.; Dasgupta, P.; Ball, G. E. NMR and modelling studies of disaccharide conformation. *Carbohydr. Res.* **2003**, *338*, 955−962.
(40) Parfondy, A.; Cyr, N.; Perlin, A. S. ^{13}C-^{1}H inter-residue, coupling in disaccharides, and the orientations of glycosidic bonds. *Carbohydr. Res.* **1977**, *59*, 299−309.
(41) Shashkov, A. S.; Lipkind, G. M.; Kochetkov, N. K. Nuclear overhauser effects for methyl β-maltoside and the conformational states of maltose in aqueous solution. *Carbohydr. Res.* **1986**, *147*, 175−182.
(42) Pérez, S.; Taravel, F.; Vergelati, C. Experimental evidences of solvent induced conformational changes in maltose. *Nouv. J. Chim.* **1985**, *4*, 561−564.

(43) Poveda, A.; Vicent, C.; Penades, S.; Jimenez-Barbero, J. NMR experiments for the detection of NOEs and scalar coupling constants between equivalent protons in trehalose-containing molecules. *Carbohydr. Res.* **1997**, *301*, 5−10.
(44) Batta, G.; Kover, K. E.; Gervay, J.; Hornyak, M.; Roberts, G. M. Temperature dependence of molecular conformation, dynamics, and chemical shift anisotropy of α,α-trehalose in D_2O by NMR relaxation. *J. Am. Chem. Soc.* **1997**, *119*, 1336−1345.
(45) Hayes, M. L.; Serianni, A. S.; Barker, R. Methyl β-lactoside: 600-MHz 1H- and 75-MHz 13C-n.m.r. studies of 2H- and 13C-enriched compounds. *Carbohydr. Res.* **1982**, *100*, 87−100.
(46) 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.
(47) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(48) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255−268.
(49) Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(50) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(51) Nosé, S.; Klein, M. L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **1983**, *50*, 1055−1076.
(52) Cremer, D.; Pople, J. A. A General Definition of Ring Puckering Coordinates. *J. Am. Chem. Soc.* **1975**, *97*, 1354−1358.
(53) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krueger, P.; van Gunsteren, W. F. The GROMOS biomolecular simulation program. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(54) Schuler, L. D.; van Gunsteren, W. F. On the Choice of Dihedral Angle Potential Energy Functions for *n*-Alkanes. *Mol. Simul.* **2000**, *25*, 301−319.
(55) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase. *J. Comput. Chem.* **2001**, *22*, 1205−1218.
(56) Spiwok, V.; Králová, B.; Tvaroska, I. Modelling of β-D-glucopyranose ring distortion in different force fields: a metadynamics study. *Carbohydr. Res.* **2010**, *345*, 530−537.
(57) Barnett, C. B.; Naidoo, K. J. Ring Puckering: A Metric for Evaluating the Accuracy of AM1, PM3, PM3CARB-1, and SCC-DFTB Carbohydrate QM/MM Simulations. *J. Phys. Chem. B* **2010**, *114*, 17142−17154.
(58) Sousa da Silva, A. W.; Vranken, W. F.; Laue, E. D. ACPYPE − AnteChamber PYthon Parser interfacE. Manuscript to be submitted.
(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(60) Gandhi, N. S.; Mancera, R. L. Can current force fields reproduce ring puckering in 2-O-sulfo-alpha-L-iduronic acid? A molecular dynamics simulation study. *Carbohydr. Res.* **2010**, *345*, 689−695.
(61) Nishida, Y.; Ohrui, H.; Meguro, H. ^{1}H-NMR Studies of (6R)- and (6S)-Deuterated D-Hexoses: Assignment of the Preferred Rotamers about C5-C6 Bond of D-Glucose and D-Galactose Derivatives in Solutions. *Tetrahedron Lett.* **1984**, *25*, 1575−1578.
(62) Thibaudeau, C.; Stenutz, R.; Hertz, B.; Klepach, T.; Zhao, S.; Wu, Q.; Carmichael, I.; Serianni, A. S. Correlated C-C and C-O bond conformations in saccharide hydroxymethyl groups: parametrization and application of redundant ^{1}H-^{1}H, ^{13}C-^{1}H, and ^{13}C-^{13}C NMR J-couplings. *J. Am. Chem. Soc.* **2004**, *126*, 15668−15685.
(63) Ohrui, H.; Nishida, Y.; Higuchi, H.; Hori, H.; Meguro, H. The preferred rotamer about the C5-C6 bond of D-galactopyranoses and the stereochemistry of dehydrogenation by D-galactose oxidase. *Can. J. Chem.* **1987**, *65*, 1145−1153.
(64) Nishida, Y.; Hori, H.; Ohrui, H.; Meguro, H. ^{1}H NMR Analyses of Rotameric Distribution of C5-C6 bonds of D-Glucopyranoses in Solution. *J. Carbohydr. Chem.* **1988**, *7*, 239−250.
(65) Brochier-Salon, M.-C.; Morin, C. Conformational analysis of 6-deoxy-6-iodo-D-glucose in aqueous solution. *Magn. Reson. Chem.* **2000**, *38*, 1041−1042.
(66) Hori, H.; Nishida, Y.; Ohrui, H.; Meguro, H. Conformational Analysis of Hydroxymethyl Group of D-Mannose Derivatives Using (6S)- and (6R)-(6−2H1)-D-Mannose. *J. Carbohydr. Chem.* **1990**, *9*, 601−618.
(67) Rao, V. S. R.; Qasba, P. K.; Balaji, P. V.; Chandrasekaran, R. Conformation of Disaccharides. In *Conformation of Carbohydrates*; Rao, V. S. R., Ed.; Harwood Academic: The Netherlands, 1998; pp 91−130.
(68) Shefter, E.; Trueblood, K. N. The crystal and molecular structure of D(+)-barium uridine-5′-phosphate. *Acta Crystallogr.* **1965**, *18*, 1067−1077.
(69) Wolfe, S. Gauche effect. Stereochemical consequences of adjacent electron pairs and polar bonds. *Acc. Chem. Res.* **1972**, *5*, 102−111.
(70) Schmid, N.; Eichenberger, A.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843−856.
(71) Huang, W.; Lin, Z.; van Gunsteren, W. F. Validation of the GROMOS 54A7 Force Field with Respect to β-Peptide Folding. *J. Chem. Theory Comput.* **2011**, *7*, 1237−1243.

5 Discussão Geral

5.1 Caracterização conformacional de glicoproteínas por DM

A correta descrição da conformação e dinâmica de uma glicoproteína através de DM, possibilitando a inferência de propriedades bioquímicas e funcionais, passa por uma série de pré-requisitos. Um dos principais é um campo de força apropriado, capaz de descrever a estrutura e plasticidade da porção proteica e da porção sacarídica, bem como a ligação entre elas e a mútua influência que uma exerce sobre a outra. Nesse sentido, o GROMOS 43A1, inicialmente testado na reprodução de dados conformacionais derivados de RMN (Pol-Fachin *et al*., 2009; Fernandes *et al*., 2010), vem se mostrando também adequado em produzir dados em concordância com variáveis experimentais, que incluem a influência da glicosilação sobre a afinidade (em nossos estudos, qualitativamente relacionada à energia de interação) de heparina à AT. Da mesma forma, esse conjunto de parâmetros, utilizado até então nos estudos do grupo junto ao pacote de simulação GROMACS, oferece um baixo custo computacional e consequente agilidade na obtenção de amostragens grandes o suficiente para descrever as propriedades conformacionais dos sistemas de interesse. Isso, que tem sido feito na escala de tempo de 100 ns, se configura como um dos pré-requisitos supracitados e que, aliados a estruturas adequadas de partida para as simulações, se somam para a obtenção de *ensembles* conformacionais apropriados para os sistemas em estudo.

Tal como previamente discutido, apesar de sua importância para vários eventos biológicos (Varki, 1993; Dwek, 1996), e da ampla distribuição dessa classe de biomoléculas na natureza (Spiro *et al*., 2002), ainda são poucos os estudos envolvendo simulação computacional de glicoproteínas na literatura. No entanto, desde agosto de 2009, quando teve início o desenvolvimento dos trabalhos da presente Tese, ocorreu um aumento significativo no número de artigos publicados avaliando dinamicamente sua estrutura e conformação (Tabela 6), contendo principalmente glicanas N- e O-ligadas. Esses trabalhos empregaram campos de força da série AMBER, nas versões GLYCAM04 (Woods *et al*., 1995) *e* GLYCAM06 (Kirschner *et al*., 2008) para a porção sacarídica, e o campo de força GROMOS 43A1 nos estudos desenvolvidos junto ao Grupo de Bioinformática Estrutural. Esse crescimento é indicativo de um maior interesse da comunidade científica, nos

últimos tempos, em estudos de sistemas biomoleculares levando em consideração todas as variáveis a que determinada molécula está sujeita, incluindo glicosilação. Da mesma forma, é também uma provável consequência da maior oferta de poder computacional, obtida através de simulação por placas de vídeo, tendo sido implementada com bastante sucesso junto ao pacote de simulação molecular AMBER (Salomon-Ferrer *et al.*, 2012), de campo de força de mesmo nome, e sugere um crescimento ainda mais acentuado desse tipo de estudo nos próximos tempos.

Tabela 6: Estudos envolvendo DM de glicoproteínas desde 2009

Glicoproteína / Glicopeptídeo	Glicosilação	Campo de Força	Referência
Protectina (CD59)	N-	GROMOS	Pol-Fachin *et al.*, 2009
Domínio EGF-like de fVII	O-	GROMOS	Pol-Fachin *et al.*, 2009
Domínio de adesão – CD2	N-	GROMOS	Pol-Fachin *et al.*, 2009
α-hCG	N-	GROMOS	Pol-Fachin *et al.*, 2009
Glicopept. anticongelamento	C-	AMBER-GLYCAM	Tam *et al.*, 2009
Domínio I-*like* de β1-integrina	N-	AMBER-GLYCAM	Pan & Song, 2010
Glicopept. anticongelamento	O-	AMBER-GLYCAM	Corzana *et al.*, 2010
Ciclooxigenase murina	N-	GROMOS	Fernandes *et al.*, 2010
Ciclooxigenase ovina	N-	GROMOS	Fernandes *et al.*, 2010
Glicopept. de mucina	O-	AMBER-GLYCAM	Corzana *et al.*, 2011
Receptor α1-nicotínico	N-	AMBER-GLYCAM	Dimitropoulous *et al.*, 2011
γ-Glutamil Transferase	N-	AMBER-GLYCAM	West *et al.*, 2011
Antitrombina III	N-	GROMOS	Resultados, trabalho II
Glicopept. de extensinas	O-	GROMOS	Seção 9, trabalho VI
CSF114(Glc)	N-	AMBER-GLYCAM	Guardiani *et al.*, 2012
Peptídeo *Pars Intercerebralis*	O-	AMBER-GLYCAM	Kaushik *et al.*, 2012
Hemaglutinina	N-	AMBER-GLYCAM	Chen *et al.*, 2012
Autotaxina	N-	AMBER-GLYCAM	Koyama *et al.*, 2012
Domínio de adesão – CD2	N-	AMBER-GLYCAM	Wang *et al.*, 2012
Art v 1	O-	GROMOS	Seção 9, trabalho III
IgG2a murino	N-	AMBER-GLYCAM	Wang *et al.*, 2013
NETNES	N- / P-	GROMOS	Chiodi & Verli, 2013
Aril sulfatase A	N-	GROMOS	Seção 9, trabalho IV
Trombina	N-	GROMOS	Resultados, trabalho III

5.1.1 Glicoproteínas N-ligadas

A síntese e modificação de oligossacarídeos durante a N-glicosilação, como discutido, são amplamente estudadas e compreendidas (Helenius & Aebi, 2001), na qual seu processamento sempre mantém o *core pentassacaride* em sua estrutura (Figura 13). Por causa dessa característica, inerente a oligossacarídeos N-ligados a proteínas, os protocolos experimentais de purificação e determinação de sua estrutura bidimensional estão mais bem estabelecidos, envolvendo principalmente espectrometria de massas (Leymarie & Zaia, 2012). Da mesma forma, a obtenção de modelos para glicoproteínas contendo essas árvores sacarídicas é facilitada, através do portal *Glycosciences* (Lütteke *et al*., 2006), que oferece a ferramenta *Glyprot* de N-glicosilação *in silico* de proteínas (Bohne-Lang & von der Lieth, 2005), que possui um banco de dados de estruturas oligossacarídicas conhecidas e preditas.

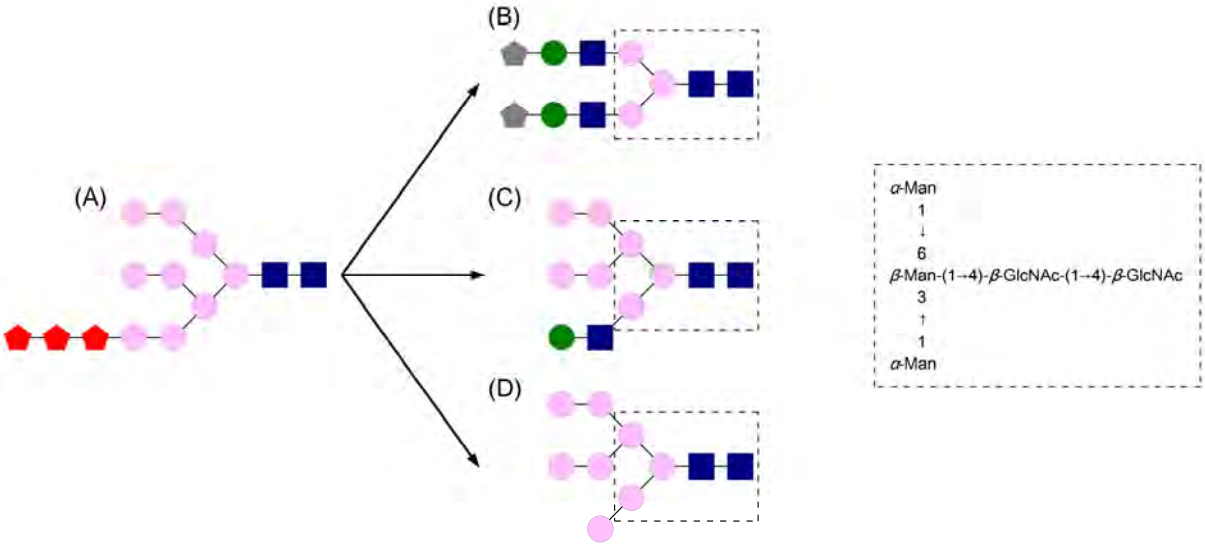


Figura 13: Representação do (A) tetradecassacarídeo precursor e de (B-D) três tipos de glicanas a que essa estrutura pode ser convertida através de processamento no RE e no complexo de Golgi. As formas azuis representam GlcNAc; as rosas, Man; as vermelhas, Glc; as verdes, Gal; e as cinzas, NeuNAc. Em linhas tracejadas, é apresentada a estrutura do *core pentassacaride*. Adaptado de Pol-Fachin, 2009.

5.1.2 Glicoproteínas O-, P-, C- e S-ligadas

Ao contrário da N-glicanas, as cadeias sacarídicas associadas aos demais tipos de glicosilação não possuem uma estrutura conservada, variando desde resíduos isolados até pequenos oligossacarídeos. Ainda assim, apenas à exceção

da S-ligação glicosídica, descoberta recentemente (Stepper *et al.*, 2011), e que envolve resíduos de Cys, todas as demais já foram estudadas por DM (Tabela 6). Não há dados conformacionais, obtidos a partir de técnicas experimentais, nem para P-ligação glicosídica, que envolve resíduos de Ser ou Thr fosforilados (Mehta *et al.*, 1996; Haynes, 1998), nem para C-ligação glicosídica, que foi descrita entre Man e Trp (Hofsteenge *et al.*, 1994; de Beer *et al.*, 1995). Da mesma forma, os dados disponíveis para a S-ligação glicosídica, no código PDB 2KUY, foram estimados computacionalmente (Venugopal *et al.*, 2011). A O-ligação glicosídica, por sua vez, já foi estudada tanto contida em glicoproteínas e glicopeptídeos completos quanto isoladamente em solução (Tabela 4). Ainda assim, há uma ampla variedade de ligações entre aminoácidos e monossacarídeos a serem estudadas (Spiro *et al.*, 2002).

Não há qualquer servidor disponível para ligação de cadeias oligossacarídicas C-, P-, S- ou O-ligadas, em oposição ao que ocorre N-glicanas. Dessa forma, a estrutura inicial de uma glicoproteína contendo quaisquer dessas glicosilações precisa ser construída à mão, em que a conformação da porção sacarídica, bem como a ligação entre aminoácido e monossacarídeo, deve ser obtida através das alternativas mostradas nos Resultados, trabalho I desta Tese. Sendo assim, um novo banco de dados de glicosilação *in silico* de proteínas, abrangendo também C-, P-, S- ou O-glicanas, apresenta um grande potencial científico para a obtenção de modelos iniciais para glicoproteínas C-, P-, S- ou O-ligadas.

5.2 Modulação da cascata de coagulação por heparina

A principal ação da heparina sobre a cascata de coagulação sanguínea reside na ativação de AT sobre suas serino proteases alvo. Ainda assim, a ação do complexo AT-heparina é constantemente ampliada a outros sistemas, por exemplo, através da identificação recente da formação de um complexo entre o zimogênio fX e AT, na presença de heparina (Nakatomi *et al.*, 2012). No entanto, já se identificou que a modulação da hemostasia, de uma forma geral, não requer, tal como AT, a sequência do pentassacarídeo mínimo, presente no GAG (Bouças *et al.*, 2012). Por outro lado, tal polissacarídeo e heparan sulfato continuam sendo o “padrão ouro” para comparação em estudos envolvendo distúrbios tromboembólicos, e vêm sendo cada vez mais reconhecidos como biomoléculas cruciais para uma série de outros processos biológicos, incluindo crescimento e diferenciação, resposta imune e

invasão de patógenos (Capila & Linhardt, 2002; Coombe & Kett, 2005; Raman *et al.*, 2005; Gandhi & Mancera, 2008).

Nesse sentido, tal polissacarídeo não está apenas associado às biomoléculas estudadas na presente Tese, tendo outros alvos tanto relacionados à hemostasia quanto não (Gandhi & Mancera, 2008). Vários complexos entre proteínas e o GAG já foram estudados na literatura, estando boa parte deles disponíveis no PDB (Gandhi & Mancera, 2008). Adicionalmente, já foram propostas sequências consenso para sítios de ligação à heparina em proteínas, compostas por xBBxBx ou xBBBxxBx, onde *B* são resíduos de Lys ou Arg (e raramente His) e *x* são preferencialmente Asn, Ser, Ala, Gly, Ile, Leu ou Tyr (Cardin & Weintraub, 1989). Uma terceira sequência consenso, xBBBxxBBBxxBBx, foi proposta baseada em uma proteína da coagulação, o fator Von Willebrand, e posteriormente encontrada em fatores não relacionados diretamente com a hemostasia. Nesse sentido, a interação de heparina com seus alvos foi observada, na presente Tese, majoritariamente com Lys ou Arg, identificados como resíduos-chave nos estudos citados acima, o que amplia a validação dos resultados aqui mostrados.

5.2.1 AT

Relacionados à hemostasia, e de interesse para esta Tese, AT, tal como discutido ao longo desta Tese, cofator 2 de heparina, inibidor de C1 (do inglês *C1-inhibitor*) e o inibidor de proteases dependente de proteína Z (do inglês *protein Z-dependent protease inhibitor*) são quatro serpinas que são moduladas por heparina (Tollefsen *et al.*, 1982; Caldwell *et al.*, 1999), e alguns dos principais alvos desse polissacarídeo quando de seu uso na clínica médica. Dentre as serpinas supracitadas, somente a interação de AT com heparina ou compostos sintéticos semelhantes já foi elucidada tridimensionalmente (McCoy *et al.*, 2003; Li *et al.*, 2004) e avaliada de forma dinâmica, em solução, a nível molecular (Verli & Guimarães, 2005; seção 4, Resultados, trabalho II). Esses estudos, inclusive aqueles realizados ao longo desta Tese, visam ao melhor entendimento das bases moleculares da interação entre a serpina e heparina, oferecendo um embasamento estrutural para o desenho racional de novos agentes bioativos que mimetizem sua ação. Sendo assim, um dado composto, a fim de disparar a cascata de modificações na estrutura tridimensional em AT, levando-a a ativação conformacional, deve presumivelmente ter interações semelhantes às de heparina com a estrutura da serpina. Como

consequência, a interação desses compostos com AT através de resíduos componentes das hélices A, D e P viabilizaria a transmissão desse sinal à β-folha A, de acordo com os resultados obtidos e mostrados na seção de Resultados, trabalho II desta Tese. Nesse sentido, vários polissacarídeos sulfatados derivados de heparina, de diferentes tamanhos, já foram desenvolvidos e são empregados em tromboprofilaxia e na terapia anticoagulante com sucesso (Harenberg & Wehling, 2008). Por outro lado, apesar de preencher boa parte dessas prerrogativas, a interação de outros agentes leva apenas a uma ativação parcial da serpina (Navarro-Fernández *et al.*, 2012). Nesse contexto, o estudo da interação entre esses compostos, incapazes de ativar AT de forma plena, por DM pode oferecer novos e importantes *insights* sobre o processo de ativação conformacional da serpina. Adicionalmente, é importante ressaltar que o uso de novos compostos que visam interagir com AT, e que sejam desenvolvidos com elevada especificidade pela serpina, dificilmente afetará as proteases da cascata de coagulação, uma vez que os sítios de ligação a ânions, ou moléculas com carga negativa – como heparina, na serpina e em trombina foram considerados diferentes (Mosier *et al.*, 2012).

5.2.2 Proteases fIIa e fXa

Nesse sentido, tal como citado no trabalho III da presente Tese, inibidores diretos de fIIa e fXa já foram amplamente estudados na literatura, no entanto, estes visam o sítio ativo da enzima ou um exosítio de ligação diferente daquele onde heparina se liga (por exemplo, Fraternali *et al.*, 1998; De Filippis *et al.*, 2002). Apenas recentemente a busca por novos agentes visando a inibição via exosítio de ligação a heparina tem sido alvo de inibidores diretos, tendo sido realizados em enzimas como a própria trombina (Fernández *et al.*, 2013) e em fXIa (Karuturi *et al.*, 2013). Em ambos os casos supracitados, a modulação da atividade enzimática foi obtida através de compostos sulfatados, não relacionados com a heparina. Da mesma forma, tal como realizado para AT, e em um dos trabalhos supracitados (Fernández *et al.*, 2013), o estudo das bases moleculares da interação desses compostos com as serino proteases deve oferecer bases estruturais para o desenho racional de novos agentes bioativos, que levem à inibição alostérica da enzima. Para isso, considerando que apenas a estrutura tridimensional do complexo trombina-heparina está disponível no PDB, os resíduos importantes para a interação das enzimas com o polissacarídeo, e consequente modulação por ele, devem ser

conhecidos. Assim como para fIIa e fXa (Sheehan & Sadler, 1994; Rezaie, 2000), explorados no trabalho III desta Tese, os aminoácidos importantes para a interação dos fatores IXa (Misenheimer & Sheehan, 2010) e XIa (Yang *et al.*, 2009) da coagulação foram identificados recentemente.

5.2.3 Complexos ternários AT-heparina-enzimas

Adicionalmente aos dados obtidos sobre a interação entre heparina e os fatores da coagulação, realizados por ocasião desta Tese, simulações por DM contendo os complexos ternários completos estão em andamento, a fim de corroborar os resultados até então observados, e obter novos *insights* para o melhor entendimento da inibição de fIIa e fXa por AT. Nessas simulações, apenas β-AT está sendo avaliada juntamente com trombina glicosilada e fXa. O emprego desta glicoforma foi baseado nos resultados obtidos no trabalho II da seção de Resultados desta Tese, uma vez que a presença da glicana ligada à Asn135 em α-AT diminuiu a energia de interação proteína-carboidrato, e causou um dobramento na cadeia do polissacarídeo na direção do RCL. A energia de interação menos intensa pode ser relacionada com a menor afinidade de α-AT pelo polissacarídeo, e o dobramento da cadeia pode, inclusive, prejudicar o reconhecimento entre as serino proteases e a serpina na glicoforma α. Baseando-se nesses resultados, justifica-se o uso de β-AT nos estudos dos complexos ternários completos, e esta glicoforma se reforça como a principal inibidora das enzimas da cascata de coagulação no plasma.

Com relação à fIIa e fXa, os complexos contendo uma heparina de cadeia longa foram ajustados de forma a que as enzimas interagissem com o polissacarídeo tal como na estrutura cristalográfica contido no código PDB 1TB6. Apesar disso, a partir dos estudos das proteases isoladas, a energia de interação entre a heparina e os resíduos componentes do exosítio 2 dessas enzimas foi mais intensa nas simulações por DM em que a geometria da tríade catalítica se desorganizou. Isso sugere que a orientação observada na superfície das enzimas para heparina é aquela que leva à sua modulação alostérica, observada em condições fisiológicas. Considerando tais orientações, observadas para fIIa glicosilado e fXa na ausência de cálcio, para trombina glicosilada (Figura 14A) uma cadeia polissacarídica longa teria a melhor orientação para interagir concomitantemente com a serpina e a protease, formando, assim, o complexo em mecanismo de ponte. Adicionalmente, para fXa sem cálcio, a montagem desse

complexo seria a mais difícil (Figura 14B), considerando o grande dobramento que a cadeia do GAG faria, o que está de acordo com a maior susceptibilidade dessa enzima ao mecanismos de ativação conformacional mediado por heparina. Por outro lado, a presença de cálcio no sistema com fXa aproxima a orientação observada para o polissacarídeo na superfície da proteína daquela observada para fIIa (Figura 14C), o que pode explicar, à nível atômico, por que, na presença de cálcio, o mecanismo de ponte é observado para fXa.

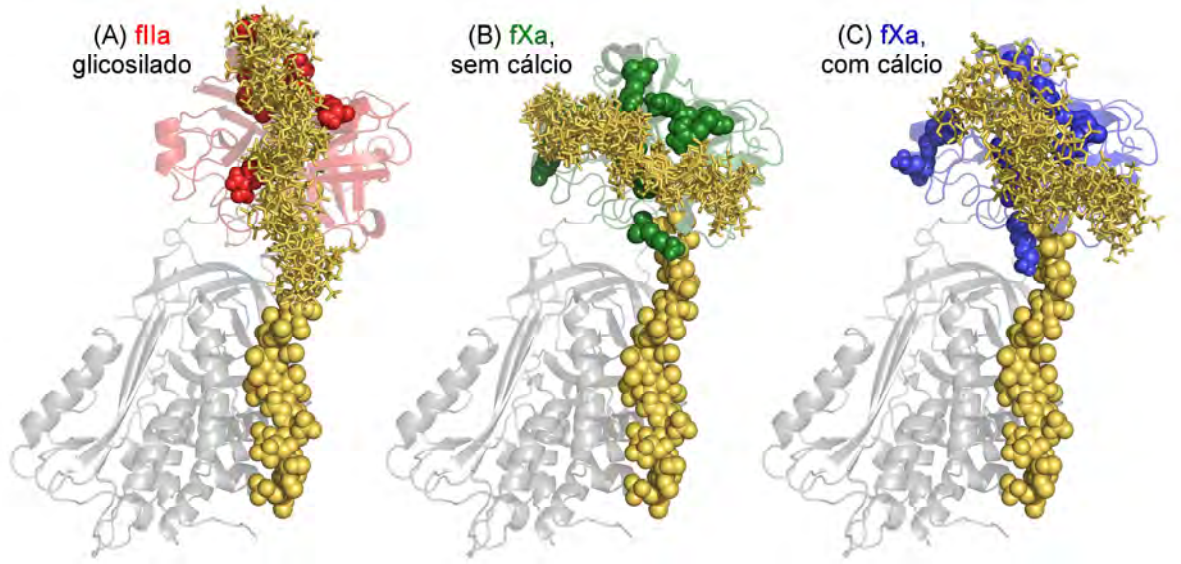


Figura 14: Orientação da cadeia de heparina durante as simulações por DM de fIIa e fXa, em comparação com a estrutura cristalográfica sob código PDB 1TB6.

De uma forma geral, considerando a inibição das serino proteases fIIa e fXa da cascata de coagulação por AT e heparina, os resultados obtidos a partir da interação do polissacarídeo com as três proteínas isoladamente, inseridos na presente Tese, oferecem novos elementos para o melhor entendimento desse processo, e podem ser utilizados, tal como discutido, como base estrutural para o desenvolvimento de novos agentes anticoagulates. Não obstante, a metodologia empregada pode ser extrapolada para outras proteases da cascata de coagulação e/ou serpinas moduladas por heparina envolvidas em outros sistemas fisiológicos, tais como sistema complemento e inflamação.

5.3 Desenvolvimento e aprimoramento de campos de força

O desenvolvimento e contínuo aprimoramento de campos de força para DM molecular são essenciais na obtenção de dados e resultados de forma mais ágil e precisa, independentemente do tipo de biomolécula. No contexto dos campos de força da série GROMOS, os parâmetros para proteínas foram recentemente aperfeiçoados para as versões 54A7 (Schmid *et al*., 2011) e 54A8 (Reif *et al*., 2013), após sucessivas atualizações, desde o 43A1 (Scott *et al*., 1999), passando pelo 45A3 (Schuler *et al*., 2001) e pelo 53A6 (Oostenbrink *et al*., 2004). De interesse para a presente Tese, o desenvolvimento de campos de força para carboidratos vem crescendo gradativamente nos últimos anos, e foi foco no último triênio após a identificação de deficiências nos parâmetros (Lins & Hünenberger, 2005) distribuídos junto aos pacotes de simulação GROMOS e GROMACS. Dentre as modificações propostas (Autieri *et al*., 2010; Hansen & Hünenberger, 2011), aquelas que mantêm maior compatibilidade com a filosofia GROMOS, com os parâmetros pré-existentes para outras classes de biomoléculas e com ambos pacotes de simulação supracitados são aqueles desenvolvidos por ocasião da presente Tese, inseridos junto ao campo de força intitulado GROMOS 53A6GLYC.

Em comparação com o que vem sendo empregado nos estudos do nosso grupo de pesquisas até então, ou seja, os parâmetros do campo de força GROMOS 43A1 e as cargas atômicas de Löwdin, na base HF/6-31G**, o novo campo de força GROMOS 53A6GLYC apresenta dois novos termos torsionais para cada diedro contendo C-C-C-O em monossacarídeos, e a ausência de diedros impróprios mantendo fixa a conformação da piranose em estudo. Dentre os dados disponíveis para compará-los, sabe-se que a conformação de dissacarídeos, estudados no trabalho IV da seção de Resultados desta Tese utilizando GROMOS 53A6GLYC e em estudos prévios com GROMOS 43A1, é descrita adequadamente por ambos os conjuntos de parâmetros, em relação com dados experimentais (Figura 15, estrela). Da mesma forma, dados preliminares sobre a geometria das ligações glicosídicas de N-glicanas, da ligação entre aminoácidos e monossacarídeos, e da influência da glicosilação sobre a estrutura protéica, obtidos a partir da parametrização do GROMOS 53A6GLYC para glicoproteínas, são comparáveis àqueles obtidos utilizando GROMOS 43A1. Nesse sentido, os dois conjuntos de parâmetros podem ser considerados equivalentes para a descrição conformacional de carboidratos e

glicoconjugados, como glicoproteínas. Por outro lado, o campo de força GROMOS 53A6GLYC ainda não pode ser utilizado para estudar polissacarídeos sulfatados, uma vez que não há parâmetros disponíveis para os grupamentos sulfato e sulfonamida. Sendo assim, uma vez que novos trabalhos avaliem outras propriedades inerentes a carboidratos, como sua flexibilidade, solvatação e interação com moléculas alvo, e mais grupamentos exocíclicos, como sulfato e sulfonamida, sejam parametrizados, poderá ser feita uma comparação mais ampla entre esses campos de força.

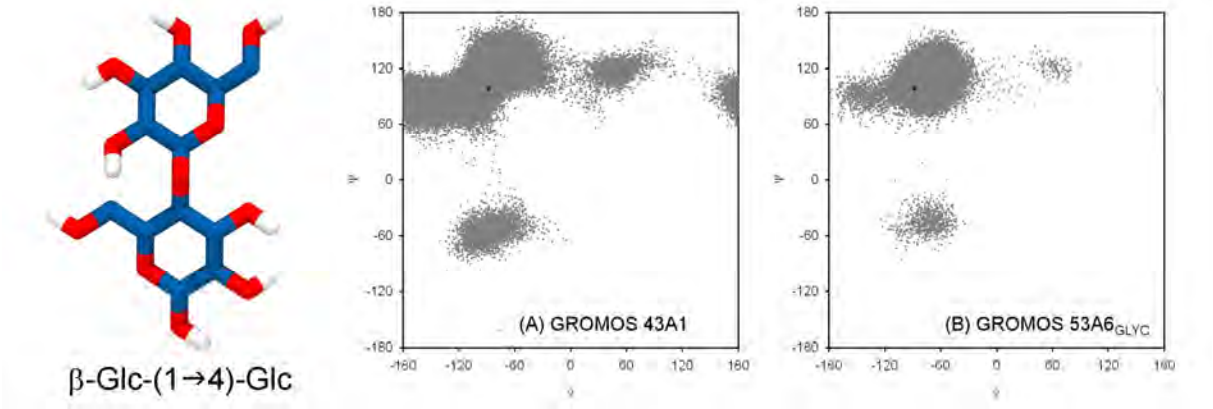


Figura 15: Perfil conformacional em solução de β-Glc-(1→4)-Glc (celobiose), obtido a partir da utilização dos campos de força GROMOS 43A1 e 53A6GLYC.

A qualidade dos conjuntos de parâmetros disponíveis para simulações computacionais deve ser avaliada, tal como supracitado e já discutido nesta Tese, através da comparação com dados experimentais. Nesse sentido, tanto o GROMOS 43A1, somado às cargas atômicas de Löwdin, quanto o GROMOS 53A6GLYC apresentam grande potencial, uma vez que reproduzem adequadamente dados derivados de RMN. A distribuição desses parâmetros junto a pacotes de simulação compilados com algoritmos rápidos, e disponíveis gratuitamente, facilita sua utilização e consequente citação na literatura, o que também amplia e divulga o trabalho dos grupos de pesquisa envolvidos na sua parametrização. Somando tudo, as simulações por DM de carboidratos livres, complexados às suas proteínas alvo, ou compondo glicoconjugados se reforçam como uma estratégia promissora para descrever e prever a estrutura, conformação, dinâmica e outras propriedades inerentes a esse tipo de biomolécula.

6 Conclusões

A partir dos objetivos traçados, o presente trabalho permitiu:

- Revisar as metodologias e protocolos empregados no estudo de glicoproteínas, oferecendo uma síntese e uma fonte de consulta sobre o assunto, que foi, ao longo deste trabalho de Tese, e poderá novamente ser empregada como base para trabalhos envolvendo essa classe de biomoléculas;
- Propor uma hipótese capaz de explicar as diferentes afinidades de α- e β-AT por heparina, a partir da identificação da interferência da glicana ligada à Asn135 de AT na interação da serpina com heparina, bem como o comportamento conformacional e dinâmico dessas glicoformas em solução;
- Construir um modelo para a via de transmissão de sinal na estrutura de AT, associada a sua ativação conformacional, desde a interação de heparina com as α-hélices A, D e P até a folha-β A e o RCL;
- Avaliar os efeitos da glicosilação sobre a conformação e dinâmica de trombina;
- Sugerir, a partir da dinâmica de heparina na superfície das proteases da cascata de coagulação estudadas, porque trombina e fXa possuem susceptibilidades distintas aos mecanismos de ação da heparina;
- Aprimorar o conjunto de parâmetros para carboidratos do GROMOS 45A4, dando origem a um novo campo de força para carboidratos, GROMOS 53A6GLYC, capaz de reproduzir qualitativamente a conformação dos aneis piranosídicos dos monossacarídeos avaliados.

Os resultados obtidos conferem uma visão estrutural, e em nível molecular, de dados experimentais prévios acerca dos complexos ternários envolvendo AT-heparina-proteases da cascata de coagulação sanguínea. Adicionalmente, confirmam o potencial de ferramentas de MM no estudo de sistemas biológicos contendo carboidratos, inclusive oferecendo um novo conjunto de parâmetros para o estudo das propriedades conformacionais de oligossacarídeos.

7 Perspectivas

Considerando os protocolos de simulação computacional empregados no presente trabalho, especialmente os relacionados ao estudo de glicoproteínas, as seguintes perspectivas podem ser traçadas:

- Ampliar os estudos de glicoproteínas utilizando campos de força da série GROMOS;
- Concluir as análises relacionadas aos complexos ternários envolvendo AT-heparina-proteases da cascata de coagulação sanguínea, a fim de confirmar as observações feitas nos estudos de DM do presente trabalho;
- Caracterização da interação de heparina com fIXa, protease da cascata de coagulação que também é inibida por AT, bem como a dinâmica do complexo AT-heparina-fIXa;
- Expandir o campo de força GROMOS 53A6$_{GLYC}$, disponível apenas para aldohexopiranoses até a finalização da presente Tese, para carboidratos contidos em glicoproteínas e GAGs.

8 Referências Bibliográficas

Referências citadas no texto e nos artigos publicados

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-(β-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.; SHEEHAN, J. K.: Predicting the molecular shape of polysaccharides from dynamic interactions with water. *Glycobiology*, **2003**, *13*, 255-264.

ALZAHRANI, S. H.; AJJAN, R. A.: Coagulation and fibrinolysis in diabetes. *Diab. Vasc. Dis. Res.*, **2010**, *7*, 260-273.

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.

ÅQVIST, J.; MEDINA, C.; SAMUELSSON, J. E.: A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.*, **1994**, *7*, 385-391.

ÅQVIST, J.; LUZHKOV, V. B.; BRANDSDAL, B. O.: Ligand binding affinities from MD simulations. *Acc. Chem. Res.*, **2002**, *35*, 358-365.

ARNOLD, K.; BORDOLI, L.; KOPP, J.; SCHWEDE, T.: The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, **2006**, *22*, 195-201.

AROCAS, V.; TURK, B.; BOCK, S. C.; OLSON, S. T.; BJÖRK, I.: The region of antithrombin interacting with full-length heparin chains outside the high-affinity pentasaccharide sequence extends to Lys136 but not to Lys139. *Biochemistry*, **2000**, *39*, 8512-8518.

AROCAS, V.; BOCK, S. C.; RAJA, S.; OLSON, S. T.; BJÖRK, I.: Lysine 114 of antithrombin is of crucial importance for the affinity and kinetics of heparin pentasaccharide binding. *J. Biol. Chem.*, **2001**, *276*, 43809-43817.

AUTIERI, E.; SEGA, M.; PEDERIVA, F.; GUELLA, G.: Puckering free energy of pyranoses: a NMR and metadynamics-umbrella sampling investigation. *J. Chem. Phys.*, **2010**, *133*, 095104.

AUTIERI, E.; SEGA, M.; PEDERIVA, F.; GUELLA, G.: Erratum: “Puckering free energy of pyranoses: a NMR and metadynamicsumbrella sampling investigation” [J. Chem. Phys. 133, 095104 (2010)]. *J. Chem. Phys.*, **2011**, *134*, 149901.

AUTIN, L.; MITEVA, M.A.; LEE, W.H.; MERTENS, K.; RADTKE, K.P.; VILLOUTREIX, B.O.: Molecular models of the procoagulant Factor VIIIa–Factor IXa complex. *J. Thromb. Haemost.*, **2005**, *3*, 2044-2056.

BAILEY, D.; RENOUF, D. V.; LARGE, D. G.; WARREN, C. D.; HOUNSELL, E. F.: Conformational studies of the glycopeptide Ac-Tyr-[Man_5GlcNAc-β-(1→4)GlcNAc-β-(1→Nδ)]-Asn-Leu-Thr-Ser-OBz and the constituent peptide and oligosaccharide. *Carbohydr. Res.*, **2000**, *324*, 242-254.

BAKER, D.; SALI, A.: Protein structure prediction and structural genomics. *Science*, **2001**, *294*, 93-96.

BARDUCCI, A.; BONOMI, M.; PARRINELLO, M.: Metadynamics. *WIREs Comput. Mol. Sci.*, **2011**, *1*, 826-843.

BARNETT, C. B.; NAIDOO, K. J.: Ring Puckering: A Metric for Evaluating the Accuracy of AM1, PM3, PM3CARB-1, and SCC-DFTB Carbohydrate QM/MM Simulations. *J. Phys. Chem. B*, **2010**, *114*, 17142-17154.

BATTA, G.; KOVER, K. E.; GERVAY, J.; HORNYAK, M.; ROBERTS, G. M.: Temperature dependence of molecular conformation, dynamics, and chemical shift anisotropy of α,α-trehalose in D2O by NMR relaxation. *J. Am. Chem. Soc.*, **1997**, *119*, 1336-1345.

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.

BELZAR, K. J.; ZHOU, A.; CARRELL, R. W.; GETTINS, P. G. W.; HUNTINGTON, J. A.: Helix D elongation and allosteric activation of antithrombin. *J. Biol. Chem.*, **2002**, *277*, 8551-8558.

BENKERT, P.; BIASINI, M.; SCHWEDE, T.: Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, **2011**, *27*, 343-350.

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.; HENRICK, K.; NAKAMURA, H.; MARKLEY, J.: Building meaningful models of glycoproteins. *Nat. Struct. Mol. Biol.*, **2007**, *14*, 354-355.

BEUSEN, D. D.; SHANDS, E. F. B.: Systematic search strategies in conformational analysis. *Drug Discov. Today*, **1996**, *1*, 429-437.

BJÖRK, I.; YLINENJÄRVI, K.; OLSON, S. T.; HERMENTIN, P.; CONRADT, H. S.; ZETTLMEISSL, G.: Decreased affinity of recombinant antithrombin for heparin due to increased glycosylation. *Biochem. J.*, **1992**, *286*, 793-800.

BLANCHARD, V.; CHEVALIER, F.; IMBERTY, A.; LEEFLANG, B. R.; BASAPPA; SUGAHARA, K.; KAMERLING, J. P.: Conformational studies on five octasaccharides isolated from chondroitin sulfate using NMR spectroscopy and molecular modeling. *Biochemistry*, **2007**, *46*, 1167-1175.

BODE, W.; MAYR, I.; BAUMANN, U.; HUBER, R.; STONE, S. R.; HOFSTEENGE, J.: The refined 1.9 A crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.*, **1989**, *8*, 3467-3475.

BODE, W.; TURK, D.; KARSHIKOV, A.: The refined 1.9-A X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human alpha-thrombin: structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships. *Protein Sci.*, **1992**, *1*, 426-471.

BOHNE-LANG, A.; VON DER LIETH, C. -W.: GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res.*, **2005**, *33*, W214-W219.

BONOMI, M.; BRANDUARDI, D.; BUSSI, G.; CAMILLONI, C.; PROVASI, D.; RAITIERI, P.; DONADIO, D.; MARINELLI, F.; PIETRUCCI, F.; BROGLIA, R. A.; PARRINELLO, M.: PLUMED: a portable plugin for free energy calculations with molecular dynamics. *Comput. Phys. Commun.*, **2009**, *180*, 1961-1972.

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.

BOUÇAS, R. I.; JARROUGE-BOUÇAS, T. R.; LIMA, M. A.; TRINDADE, E. S.; MORAES, F. A.; CAVALHEIRO, R. P.; TERSARIOL, I. L.; HOPPENSTEAD, D.; FAREED, J.; NADER, H. B.: Glycosaminoglycan backbone is not required for the modulation of hemostasis: effect of different heparin derivatives and non-glycosaminoglycan analogs. *Matrix Biol.*, **2012**, *31*, 308-316.

BRENNAN, S. O.; GEORGE, P. M.; JORDAN, R. E.: Physiological variant of antithrombin-III lacks carbohydrate sidechain at Asn 135. *FEBS Lett.*, **1987**, *219*, 431-436.

BRIGO, A.; LEE, K. W.; IURCU MUSTATA, G.; BRIGGS, J. M.: Comparison of multiple molecular dynamics trajectories calculated for the drug-resistant HIV-1 Integrase T66I/M154I catalytic domain. *Biophys. J.*, **2005**, *88*, 3072-3082.

BROCHIER-SALON, M. -C.; MORIN, C.: Conformational analysis of 6-deoxy-6-iodo-D-glucose in aqueous solution. *Magn. Reson. Chem.*, **2000**, *38*, 1041-1042.

BROZE, G. J. JR.; WARREN, L. A.; NOVOTNY, W. F.; HIGUCHI, D. A.; GIRARD, J. J.; MILETICH, J. P.: The lipoprotein-associated coagulation inhibitor that inhibits the factor VII-tissue factor complex also inhibits factor Xa: insight into its possible mechanism of action. *Blood*, **1988**, *71*, 335-343.

BUBB, W.: NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concepts Magn. Reson. Part A*, **2003**, *19A*, 1-13.

CALDWELL, E. E.; ANDREASEN, A. M.; BLIETZ, M. A.; SERRAHN, J. N.; VANDERNOOT, V.; PARK, Y.; YU, G.; LINHARDT, R. J.; WEILER, J. M.: Heparin binding and augmentation of C1 inhibitor activity. *Arch. Biochem. Biophys.*, **1999**, *361*, 215-222.

CAPILA, I.; LINHARDT, R. J.: Heparin-protein interactions. *Angew Chem. Int. Ed. Engl.*, **2002**, *41*, 390-412.

CARDIN, A. D.; WEINTRAUB, H. J.: Molecular modeling of protein-glycosaminoglycan interactions. *Arterioscler. Thromb. Vasc. Biol.*, **1989**, *9*, 21-32.

CASE, D. A.; CHEATHAM, T. E. 3RD; DARDEN, T.; GOHLKE, H.; LUO, R.; MERZ, K. M. JR.; ONUFRIEV, A.; SIMMERLING, C.; WANG, B.; WOODS, R. J.: The Amber Biomolecular Simulation Programs. *J. Comput. Chem.*, **2005**, *26*, 1668-1688.

CASTRO, M. O.; POMIN, V. H.; SANTOS, L. L.; VILELA-SILVA, A.-C. E. S.; HIROHASHI, N.; POL-FACHIN, L.; VERLI, H.; MOURÃO, P. A. S.: A Unique 2-Sulfated β-Galactan from the Egg Jelly of the Sea Urchin *Glyptocidaris crenularis*. Conformation flexibility versus induction of the sperm acrosome reaction. *J. Biol. Chem.*, **2009**, *284*, 18790-18800.

CAVASOTTO, C. N.; PHATAK, S. S.: Homology modeling in drug discovery: current trends and applications. *Drug Discov. Today*, **2009**, *14*, 676-683.

CHEETHAM, N. W. H.; DASGUPTA, P.; BALL, G. E.: NMR and modelling studies of disaccharide conformation. *Carbohydr. Res.*, **2003**, *338*, 955-962.

CHEN, W.; SUN, S.; LI, Z.: Two glycosylation sites in H5N1 Influenza virus hemagglutinin that affect binding preference by computer-based analysis. *PLoS ONE*, **2012**, *7*, e38794.

CHIODI, C. G.; VERLI, H.: Structural characterization of NETNES glycopeptide from *Trypanosoma cruzi*. *Carbohydr. Res.*, **2013**, *373*, 28-34.

CHOAY, J.; PETITOU, M.; LORMEAU, J. -C.; SINAY, 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.

CHURCH, T.; CARMICHAEL, I.; SERIANNI, A. S.: Two-bond ^{13}C-^{13}C spin-coupling constants in carbohydrates: effect of structure on coupling magnitude and sign. *Carbohydr. Res.*, **1996**, *280*, 177-186.

COOMBE, D. R.; KETT, W. C.: Heparan sulfate-protein interactions: therapeutic potential through structure-function insights. *Cell. Mol. Life Sci.*, **2005**, *62*, 410-424.

COMPLEX CARBOHYDRATE RESEARCH CENTER, University of Georgia, Athens, GA. Woods Group (2005-2010). GLYCAM http://www.glycam.com (Accessed April 28th, 2010).

CORZANA, F.; BUSTO, J. H.; JIMÉNEZ-OSÉS, G.; ASENSIO, J. L.; JIMÉNEZ-BARBERO, J.; PEREGRINA, J. M.; AVENOZA, A.: New Insights into α-GalNAc–Ser Motif: Influence of Hydrogen Bonding versus Solvent Interactions on the Preferred Conformation. *J. Am. Chem. Soc.*, **2006a**, *128*, 14640-14648.

CORZANA, F.; BUSTO, J. H.; ENGELSEN, S. B.; JIMÉNEZ-BARBERO, J.; ASENSIO, J. L.; PEREGRINA, J. M.; AVENOZA, A.: Effect of β-O-Glucosylation on L-Ser and L-Thr Diamides: A Bias toward α-Helical Conformations. *Chemistry*, **2006b**, *12*, 7864-7871.

CORZANA, F.; BUSTO, J. H.; JIMÉNEZ-OSÉS, G.; GARCÍA DE LUIS, M.; ASENSIO,J. L.; JIMÉNEZ-BARBERO, J.; PEREGRINA, J. M.; AVENOZA, A.: Serine versus Threonine Glycosylation: the methyl group causes a drastic alteration on the carbohydrate orientation and on the surrounding water shell. *J. Am. Chem. Soc.*, **2007**, *129*, 9458-9467.

CORZANA, F.; BUSTO, J. H.; JIMÉNEZ-OSÉS, G.; GARCÍA DE LUIS, M.; FERNÁNDEZ-TEJADA, A.; RODRÍGUEZ, F.; JIMÉNEZ-BARBERO, J.; AVENOZA, A.; PEREGRINA, J. M.: Dynamics and Hydration Properties of Small Antifreeze-Like Glycopeptides Containing Non-Natural Amino Acids. *Eur. J. Org. Chem.*, **2010**, 3525-3532.

CORZANA, F.; BUSTO, J. H.; MARCELO, F.; GARCÍA DE LUIS, M.; ASENSIO, J. L.; MARTÍN-SANTAMARÍA, S.; JIMÉNEZ-BARBERO, J.; AVENOZA, A.; PEREGRINA, J. M.: Engineering O-glycosylation points in non-extended peptides: implications for the molecular recognition of short tumor-associated glycopeptides. *Chemistry*, **2011**, *17*, 3105-3110.

CREMER, D.; POPLE, J. A.: A General definition of ring puckering coordinates. J. *Am. Chem. Soc.*, **1975**, *97*, 1354-1358.

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.

DAHLBÄCK, B.: Blood coagulation. *Lancet*, **2000**, *355*, 1627-1632.

DAHLBÄCK, B.; VILLOUTREIX, B. O.: Regulation of blood coagulation by the protein C anticoagulant pathway. Novel insights into structure-function relationships and molecular recognition. *Arterioscler. Thromb. Vasc. Biol.*, **2005**, 25, 1311-1320.

DAMM, W.; FRONTERA, A.; TIRADO-RIVES, J.; JORGENSEN, W.: OPLS all-atom force field for carbohydrates. *J. Comput. Chem.*, **1997**, *18*, 1955-1970.

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.

DA SILVA, C. O.; NASCIMENTO, M. A.: Ab initio conformational maps for disaccharides in gas phase and aqueous solution. *Carbohydr. Res.*, **2004**, *339*, 113-122.

DE BEER, T.; VLIEGENTHART, 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_s Is α-Mannopyranose. *Biochemistry*, **1995**, *34*, 11785-11789.

DE FILIPPIS, V.; COLOMBO, G.; RUSSO, I.; SPADARI, B.; FONTANA, A.: Probing the hirudin-thrombin interaction by incorporation of noncoded amino acids and molecular dynamics simulation. *Biochemistry*, **2002**, *41*, 13556-13569.

DE GROOT, B. L.; GRUBMÜLLER, H.: Water Permeation Across Biological Membranes: Mechanism and Dynamics of Aquaporin-1 and GlpF. *Science*, **2001**, *294*, 2353-2357.

DELA CRUZ, R. G.; JAIRAJPURI, M. A.; BOCK, S. C.: Disruption of a tight cluster surrounding tyrosine 131 in the native conformation of antithrombin III activates it for factor Xa inhibition. *J. Biol. Chem.*, **2006**, *281*, 31668-31676.

DELANO, W. L.: The PyMOL Molecular Graphics System, Schrodinger Sales Center, Portland, OR, USA, 2002. http://www.pymol.org

DESAI, U.; SWANSON, R.; BOCK, S. C.; BJÖRK, I.; OLSON, S. T.: Role of arginine 129 in heparin binding and activation of antithrombin. *J. Biol. Chem.*, **2000**, *275*, 18976-18984.

DEMELBAUER, U. M.; PLEMATL, A.; JOSIC, D.; ALLMAIER, G.; RIZZI, A.: On the variation of glycosylation in human plasma derived antithrombin. *J. Chromatogr. A*, **2005**, *1080*, 15-21.

DIMITROPOULOS, N.; PAPAKYRIAKOU, A.; DALKAS, G. A.; CHASAPIS, C. T.; POULAS, K.; SPYROULIAS, G. A.: A computational investigation on the role of glycosylation in the binding of alpha1 nicotinic acetylcholine receptor with two alpha-neurotoxins. *Proteins*, **2011**, *79*, 142-152.

DUNDAS, J.; OUYANG, Z.; TSENG, J.; BINKOWSKI, A.; TURPAZ, Y.; LIANG, J.: CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.*, **2006**, *34*, W116-W118.

DWEK, R. A.: Glycobiology: toward understanding the function of sugars. *Chem. Rev.*, **1996**, *96*, 683-720.

EARGLE, J.; LUTHEY-SCHULTEN, Z.: NetworkView: 3D display and analysis of protein-RNA interaction networks. *Bioinformatics*, **2012**, *28*, 3000-3001.

ELKLUND, R.; WIDMALM, G.: Molecular dynamics simulations of an oligosaccharide using a force field modified for carbohydrates. *Carbohydr. Res.*, **2003**, *338*, 393-398.

ENGLEBIENNE, P.; MOITESSIER, N.: Docking ligands into flexible and solvated macromolecules. 4. are popular scoring functions accurate for this class of proteins? *J. Chem. Inf. Model.*, **2009**, *49*, 1568-1580.

ERBEL, P. J. A.; KARIMI-NEJAD, Y.; VAN KUIK, J. A.; BOELENS, R.; KAMERLING, J. P.; VLIEGENTHART, J. F. G.: Effects of the N-Linked Glycans on the 3D Structure of the Free α-Subunit of Human Chorionic Gonadotropin. *Biochemistry*, **2000**, *39*, 6012-6021.

ERSDAL-BADJU, E.; LU, A.; ZUO, Y.; PICARD, V.; BOCK, S. C.: Identification of the antithrombin III heparin binding site. *J. Biol. Chem.*, **1997**, *272*, 19393-19400.

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.

FEDOROV, M. V.; GOODMAN, J. M.; NERUCK, D.; SCHUMM, S.: Self assembly of trehalose molecules on a lysozyme surface: the broken glass hypothesis. *Phys. Chem. Chem. Phys.*, **2011**, *13*, 2294-2299.

FERNANDES, C. L.; SACHETT, L. G.; POL-FACHIN, L.; VERLI, H.: GROMOS96 43a1 performance in predicting oligosaccharide conformational ensembles within glycoproteins. *Carbohydr. Res.*, **2010**, *345*, 663-671.

FERNÁNDEZ-TEJADA, A.; CORZANA, F.; BUSTO, J. H.; JIMÉNEZ-OSÉS, G.; JIMÉNEZ-BARBERO, J.; AVENOZA, A.; PEREGRINA, J. M.: Insights into the geometrical features underlying β-O-GlcNAc glycosylation: water pockets drastically modulate the interactions between the carbohydrate and the peptide backbone. *Chemistry*, **2009**, *15*, 7297-7301.

FERNÁNDEZ, P. V.; QUINTANA, I.; CEREZO, A. S.; CARAMELO, J. J.; POL-FACHIN, L.; VERLI, H.; ESTEVEZ, J. M.; CIANCIA, M.: Anticoagulant activity of a unique sulfated pyranosic (1→3)-β-L-arabinan through direct interaction with thrombin. *J. Biol. Chem.*, **2013**, *288*, 223-233.

FERRO, D. R.; PROVASOLI, A.; RAGAZZI, M.; TORRI, G.; CASU, B.; GATTI, G.; JACQUINET, J. C.; SINAY, P.; PETITOU, M.; CHOAY, J.: Evidence for conformational equilibrium of the sulfated

L-iduronate residue in heparin and in synthetic heparin mono- and oligosaccharides: NMR and force-field studies. *J. Am. Chem. Soc.*, **1986**, *108*, 6773-6778.

FINCHER, G. B.: Exploring the evolution of (1,3;1,4)-β-D-glucans in plant cell walls: comparative genomics can help! *Curr. Opin. Plant Biol.*, **2009**, *12*, 140-147.

FISCHER, H.; OLIVEIRA NETO, M.; NAPOLITANO, H. B.; CRAIEVICH, A. F.; POLIKARPOV, I.: The molecular weight of proteins in solution can be determined from a single SAXS measurement on a relative scale. *J. Appl. Cryst.*, **2010**, *43*, 101-109.

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.

FRANCA, E. F.; LINS, R. D.; FREITAS, L. C. G.; STRAATSMA, T. P.: Characterization of chitin and chitosan molecular structure in aqueous solution. *J. Chem. Theory Comput.*, **2008**, *4*, 2141-2149.

FRANCA, E. F.; FREITAS, L. C. G.; LINS, R. D.: Chitosan molecular structure as a function of N-acetylation. *Biopolymers*, **2011**, *95*, 448-460.

FRANK, M.; BOHNE-LANG, A.; WETTER, T.; VON DER LIETH, C. -W.: Rapid generation of a representative ensemble of N-glycan conformations. *In Silico Biol.*, **2002**, *2*, 427-439.

FRANZÉN, L. E.; SVENSSON, S.; LARM, O.: Structural studies on the carbohydrate portion of human antithrombin III. *J. Biol. Chem.*, **1980**, *255*, 5090-5093.

FRATERNALI, F.; DO, Q.-T.; DOAN, B.-T.; ATKINSON, R. A.; PALMAS, P.; SKLENAR, V.; SAFAR, P.; WILDGOOSE, P.; STROP, P.; SAUDEK, V.: Mapping the active site of factor Xa by selective inhibitors: an NMR and MD study. *Proteins*, **1998**, *30*, 265-274.

FREBELIUS, S.; ISAKSSON, S.; SWEDENBORG, J.: Thrombin inhibition by antithrombin III on the subendothelium is explained by the isoform ATβ. *Arterioscler. Thromb. Vasc. Biol.*, **1996**, *16*, 1292-1297.

FREDENBURGH, J. C.; STAFFORD, A. R.; WEITZ, J. I.: Evidence for allosteric linkage between exosites 1 and 2 of thrombin, *J. Biol. Chem.*, **1997**, *272*, 25493-25499.

FRENCH, A. D.; KELTERER, A. -M.; JOHNSON, G. P.; DOWD, M. K.; CRAMER, C. J.: HF/6-31G* energy surfaces for disaccharide analogs. *J. Comput. Chem.*, **2001**, *22*, 65-78.

FRISCH, M. J.; TRUCKS, G. W.; SCHLEGEL, H. B.; SCUSERIA, G. E.; ROBB, M. A.; CHEESEMAN, J. R.; ZAKRZEWSKI, V. G.; STRATMANN, R. E.; BURANT, J. C.; DAPPRICH, S.; MILLAM, J. M.; DANIELS, A. D.; KUDIN, K. N.; STRAIN, M. C.; FARKAS, O.; TOMASI, J.; BARONE, V.; COSSI, M.; CAMMI, R.; MENNUCCI, B.; POMELLI, C.; ADAMO, C.; CLIFFORD, S.; OCHTERSKI, J.; PETERSSON, G. A.; AYALA, P. Y.; CUI, Q.; MOROKUMA, K.; REGA, N.; SALVADOR, P.; DANNENBERG, J. J.; MALICK, D. K.; RABUCK, A. D.; RAGHAVACHARI, K.; FORESMAN, J. B.; CIOSLOWSKI, J.; ORTIZ, J. V.; BABOUL, A. G.; STEFANOV, B. B.; LIU, G.; LIASHENKO, A.; PISKORZ, P.; KOMAROMI, I.; GOMPERTS, R.; MARTIN, R. L.; FOX, D. J.; KEITH, T.; LAHAM, A.; PENG, C. Y.; NANAYAKKARA, A.; CHALLACOMBE, M.; GILL, P. M. W.; JOHNSON, B.; CHEN, W.; WONG, M. W.; ANDRES, J. L.; GONZALEZ, C.;

GORDON, H.; REPLOGLE, E. S.; POPLE, J. A.: Gaussian 98, Revision A.11.4.; Gaussian, Inc.: Pittsburgh, PA, 2002.

FUGLESTAD, B.; GASPER, P. M.; TONELLI, M.; MCCAMMON, J. A.; MARKWICK, P. R. L.; KOMIVES, E. A.: The dynamic structure of thrombin in solution. *Biophys. J.*, **2012**, *103*, 79-88.

FURIE, B.; FURIE, B. C.: The molecular basis of blood coagulation. *Cell*, **1988**, *33*, 505-518.

GALLWITZ, M.; ENOKSSON, M.; THORPE, M.; HELLMAN, L.: The extended cleavage specificity of human thrombin. *PLoS ONE*, **2012**, *7*, e31756.

GANDHI, N. S.; MANCERA, R. L.: The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.*, **2008**, *72*, 455-482.

GANDHI, N. S.; MANCERA, R. L.: Free energy calculations of glycosaminoglycan–protein interactions. *Glycobiology*, **2009**, *19*, 1103-1115.

GANDHI, N. S.; MANCERA, R. L.: Can current force fields reproduce ring puckering in 2-O-sulfo-alpha-L-iduronic acid? A molecular dynamics simulation study. *Carbohydr. Res.*, **2010**, *345*, 689–695.

GARONE, L.; EDMUNDS, T.; HANSON, E.; BERNASCONI, R.; HUNTINGTON, J. A.; MEAGHER, J. L.; FAN, B.; GETTINS, P. G. W.: Antithrombin-heparin affinity reduced by fucosylation of carbohydrate at asparagine 155. *Biochemistry*, **1996**, *35*, 8881-8889.

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.

GETTINS, P. G. W.: Serpin structure, mechanism, and function. *Chem. Rev.*, **2002**, *102*, 4751-4803.

GIRARD, T. J.; WARREN, L. A.; NOVOTNY, W. F.; LIKERT, K. M.; BROWN, S. G.; MILETICH, J. P.; BROZE, G. J. JR.: Functional significance of the Kunitz-type inhibitory domains of lipoprotein-associated coagulation inhibitor. *Nature*, **1989**, *338*, 518-520.

GOHLKE, H.; KIEL, C.; CASE, D. A.: Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.*, **2003**, *330*, 891-913.

GONZÁLEZ-OUTEIRIÑO, J.; KADIRVELRAJ, R.; WOODS, R. J.: Structural elucidation of type III group B Streptococcus capsular polysaccharide using molecular dynamics simulations: the role of sialic acid. *Carbohydr. Res.*, **2005**, *340*, 1007-1018.

GRAY, J. J.; MOUGHON, S. E.; KORTEMME, T.; SCHUELER-FURMAN, O.; MISURA, K. M.; MOROZOV, A. V.; BAKER, D.: Protein-protein docking predictions for the CAPRI experiment. *Proteins*, **2003**, *52*, 118-122.

GREEN, D.: Coagulation cascade. *Hemodial. Int.*, **2006**, *10* (Issue Suppl. s1), S2-S4.

GREEN, D. F.: Optimized parameters for continuum solvation calculations with carbohydrates. *J. Phys. Chem. B*, **2008**, *112*, 5238-5249.

GREENFIELD, N. J.: Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.*, **2006**, *1*, 2876-2890.

GUARDIANI, C.; SIGNORINI, G. F.; LIVI, R.; PAPINI, A. M.; PROCACCI, P.: Conformational landscape of N-glycosylated peptides detecting autoantibodies in multiple sclerosis, revealed by Hamiltonian replica exchange. *J. Phys. Chem. B*, **2012**, *116*, 5458-5467.

GUERRINI, M.; GUGLIERI, S.; CASU, B.; TORRI, G.; MOURIER, P.; BOUDIER, C.; VISKOV, C.: Antithrombin-binding octasaccharides and role of extensions of the active pentasaccharide sequence in the specificity and strength of interaction. Evidence for very high affinity induced by an unusual glucuronic acid residue. *J. Biol. Chem.*, **2008**, *283*, 26662-26675.

GUIMARÃES, C. R. W.; MATHIOWETZ, A. M.: Addressing limitations with the MMGB/SA scoring procedure using the watermap method and free energy perturbation calculations. *J. Chem. Inf. Model.*, **2010**, *50*, 547-559.

GUPTA, R.; BIRCH, H.; RAPACKI, K.; BRUNAK, S.; HANSEN, J. E.: O-GLYCBASE version 4.0: a revised databse of O-glycosylated proteins. *Nucleic Acids Res.*, **1999**, *27*, 370-372.

GUVENCH, O.; GREENE, S. N.; KAMATH, G.; BRADY, J. W.; VENABLE, R. M.; PASTOR, R. W.; MACKERELL, A. D. J.: Additive empirical force field for hexopyranose monosaccharides. *J. Comput. Chem.*, **2008**, *29*, 2543-2564.

GUVENCH, O.; HATCHER, E. R.; VENABLE, R. M.; PASTOR, R. W.; MACKERELL, A. D. J.: CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses. *J. Chem. Theory Comput.*, **2009**, *5*, 2353-2370.

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.

HAASNOOT, C. A. G.: Conformational analysis of six-membered rings in solution: ring puckering coordinates derived from vicinal NMR proton-proton coupling constants. *J. Am. Chem. Soc.*, **1993**, *115*, 1460-1468.

HACKENG, T. M.; MAURISSEN, L. F. A.; CASTOLDI, E.; ROSING, J.: Regulation of TFPI function by protein S. *J. Thromb. Haemost.*, **2009**, *7* (Suppl. 1), 165-168.

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, H. S.; HÜNENBERGER, P. H.: A Reoptimized GROMOS force field for hexopyranose-based carbohydrates accounting for the relative free energies of ring conformers, anomers, epimers, hydroxymethyl rotamers, and glycosidic linkage conformers. *J. Comput. Chem.*, **2011**, *32*, 998-1032.

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.

HARENBERG, J.; WEHLING, M.: Current and future prospects for anticoagulant therapy: inhibitors of factor Xa and factor IIa. *Semin. Thromb. Hemostasis*, **2008**, *34*, 39-57.

HARRIS, J.L.; BACKES, B.J.; LEONETTI, F.; MAHRUS, S.; ELLMAN, J.A.; CRAIK, C.S.: Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. *Proc. Natl. Acad. Sci. U.S.A.*, **2000**, *97*, 7754-7759.

HAYES, M. L.; SERIANNI, A. S.; BARKER, R.: Methyl β-lactoside: 600-MHz 1H- and 75-MHz 13C-n.m.r. studies of 2H- and 13C-enriched compounds. *Carbohydr. Res.*, **1982**, *100*, 87-100.

HAYNES, P. A.: Phosphoglycosylation: a new structural class of glycosylation? *Glycobiology*, **1998**, *8*, 1-5.

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.

HESS, B.; KUTZNER, C.; VAN DER SPOEL, D.; LINDAHL, E.: GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, **2009**, *4*, 435-447.

HILL, A. D.; REILLY, P. J.: A Gibbs Free energy correlation for automated docking of carbohydrates. *J. Comput. Chem.*, **2008**, *29*, 1131-1141.

HOFFMAN, M.; MONROE III, D. M.: A cell-based model of hemostasis. *Thromb. Haemost.*, **2001**, *85*, 958-965.

HOFSTEENGE, J.; MÜLLER, D. R.; DE BEER, T.; LÖFFLER, A.; RICHTER, W. J.; VLIEGENTHART, J. F.: New type of linkage between a carbohydrate and a protein: C-glycosylation of a specific tryptophan residue in human RNase U_s. *Biochemistry*, **1994**, *33*, 13524-13530.

HONG, X.; HOPFINGER, A. J.: Construction, molecular modeling, and simulation of Mycobacterium tuberculosis cell walls. *Biomacromolecules*, **2004**, *5*, 1052-1065.

HOOVER, W. G.: Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A*, **1985**, *31*, 1695-1697.

HORI, H.; NISHIDA, Y.; OHRUI, H.; MEGURO, H.: Conformational Analysis of Hydroxymethyl Group of D-Mannose Derivatives Using (6S)- and (6R)-(6–2H1)-D-Mannose. *J. Carbohydr. Chem.*, **1990**, *9*, 601i618.

HORTA, B. A. C.; PERIC-HASSLER, L.; HÜNENBERGER, P. H.: Interaction of the disaccharides trehalose and gentiobiose with lipid bilayers: A comparative molecular dynamics study. *J. Mol. Graphics Modell.*, **2010**, *29*, 331-346.

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.

HUANG, W.; LIN, Z.; VAN GUNSTEREN, W. F.: Validation of the GROMOS 54A7 Force Field with Respect to β-Peptide Folding. *J. Chem. Theory Comput.*, **2011**, *7*, 1237-1243.

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.

HUNTINGTON, J. A.; READ, R. J.; CARRELL, R. W.: Structure of a serpin-protease complex shows inhibition by deformation. *Nature*, **2000a**, *407*, 923-926.

HUNTINGTON, J. A.; MCCOY, A.; BELZAR, K. J.; PEI, X. Y.; GETTINS, P. G. W.; CARRELL, R. W.: The conformational activation of antithrombin. A 285-Å structure of a fluorescein derivative reveals an electrostatic link between the hinge and heparin binding regions. *J. Biol. Chem.*, **2000b**, *275*, 15377-15383.

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.; 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.; PERÉZ, 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.

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.

JACKSON R. L.; BUSCH S. J.; CARDIN A. D.: Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. *Physiol. Rev.*, **1991**, *71*, 481-539.

JAIRAJPURI, M. A.; LU, A.; DESAI, U. R.; OLSON, S. T.; BJÖRK, I.; BOCK, S. C.: Antithrombin III phenylalanines 122 and 121 contribute to its high affinity for heparin and its conformational activation. *J. Biol. Chem.*, **2003**, *278*, 15941-15950.

JIN, L.; ABRAHAMS, J. P.; SKINNER, R.; PETITOU, M.; PIKE, R. N.; CARRELL, R. W.: The anticoagulant activation of antithrombin by heparin. *Proc. Natl. Acad. Sci. U.S.A.*, **1997**, *94*, 14683-14688.

JOHNSON, D. J. D.; HUNTINGTON, J. A.: Crystal structure of antithrombin in a heparin-bound intermediate state. *Biochemistry*, **2003**, *42*, 8712-8719.

JOHNSON, D. J. D.; LANGDOWN, L.; LI, W.; LUIS, S. A.; BAGLIN, T. A.; HUNTINGTON, J. A.: Crystal structure of monomeric native antithrombin reveals a novel reactive center loop conformation. *J. Biol. Chem.*, **2006**, *281*, 35478-35486.

JOHNSON, D. J. D.; LI, W.; ADAMS, T. E.; HUNTINGTON, J. A.: Antithrombin–S195A factor Xa-heparin structure reveals the allosteric mechanism of antithrombin activation. *EMBO J.*, **2006**, *25*, 2029-2037.

JORGENSEN, W. L.; CHANDRASEKHAR, J.; MADURA, J. D.; IMPEY, R. W.; KLEIN, M. L.: Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **1983**, *79*, 926-935.

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.

KABSCH, W.; SANDER, C.: Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **1983**, *22*, 2577-2637.

KAMATA, K.; KAWAMOTO, H.; HONMA, T.; IWAMA, T.; KIM, S. H.: Structural basis for chemical inhibition of human blood coagulation factor Xa. *Proc. Natl. Acad. Sci. U. S. A.*, **1998**, *95*, 6630-6635.

KARPLUS, M.; PETSKO, G. A.: Molecular dynamics simulations in biology. *Nature*, **1990**, *347*, 631-639.

KARPLUS, M.; MCCAMMON, J. A.: Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.*, **2002**, *9*, 646-652.

KARUTURI, R.; AL-HORANI, R. A.; MEHTA, S. C.; GAILANI, D.; DESAI, U. R.: Discovery of allosteric modulators of dactor XIa by targeting hydrophobic domains adjacent to its heparin-binding site. *J. Med. Chem.*, **2013**, *56*, 2415-2428.

KAUSHIK, S.; MOHANTY, D.; SUROLIA, A.: Molecular dynamics simulations on *pars intercerebralis* major peptide-C (PMP-C) reveal the role of glycosylation and disulfide bonds in its enhanced structural stability and function. *J. Biomol. Struct. Dyn.*, **2012**, *29*, 905-920.

KERZMANN, A.; NEUMANN, D.; KOHLBACHER, O.: SLICK - scoring and energy functions for protein-carbohydrate interactions. *J. Chem. Inf. Model.*, **2006**, *46*, 1635-1642.

KHAN, S.; GOR, J.; MULLOY, B.; PERKINS, S. J.: Semi-rigid solution structures of heparin by constrained X-ray scattering modelling: new insight into heparin-protein complexes. *J. Mol. Biol.*, **2010**, *395*, 504-521.

KIFLI, N; HTAR, T. T.; DE CLERCQ, E.; BALZARINI, J.; SIMONS, C.: Novel bicyclic sugar modified nucleosides: synthesis, conformational analysis and antiviral evaluation. *Bioorg. Med. Chem.*, **2004**, *12*, 3247-3257.

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.

KOLLER, A. N.; SCHWALBE, H.; GOHLKE, H.: Starting structure dependence of NMR order parameters derived from MD simulations: implications for judging force-field quality. *Biophys. J.*, **2008**, *95*, L04-L06.

KONY, D.; DAMM, W.; STOLL, S.; VAN GUNSTEREN, W. F.: An improved OPLS-AA force field for carbohydrates. *J. Comput. Chem.*, **2002**, *23*, 1416-1429.

KOYAMA, M.; NISHIMASU, H.; ISHITANI, R.; NUREKI, O.: Molecular dynamics simulation of autotaxin: roles of the nuclease-like domain and the glycan modification. *J. Phys. Chem. B*, **2012**, *116*, 11798-11808.

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.

KUKURUZINSKA, M. A.; LENNON, K.: Protein N-Glycosylation: molecular genetics and functional significance. *Crit. Rev. Oral Biol. Med.*, **1998**, *9*, 415-448.

KUSCHE, M.; TORRI, G.; CASU, B.; LINDAHL, U.: Biosynthesis of heparin. Availability of glucosaminyl 3-O-sulfation sites. *J. Biol. Chem.*, **1990**, *265*, 7292-7300.

KUTTEL, M.; BRADY, J. W.; NAIDOO, K. J.: Carbohydrate solution simulations: Producing a force field with experimentally consistent primary alcohol rotational frequencies and populations. *J. Comput. Chem.*, **2002**, *23*, 1236-1243.

LAIO, A.; PARRINELLO, M.: Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.*, **2002**, *99*, 12562-12566.

LAIO, A.; GERVASIO, F. L.: Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.*, **2008**, *71*, 126601.

LANDSTRÖM, J.; WIDMALM, G.: Glycan flexibility: insights into nanosecond dynamics from a microsecond molecular dynamics simulation explaining an unusual nuclear Overhauser effect. *Carbohydr. Res.*, **2010**, *345*, 330-333.

LANE, D. L.; DENTON, J.; FLYNN, A. M.; THUNBERG, L.; LINDAHL, U.: Anticoagulant activities of heparin oligosaccharides and their neutralization by platelet factor 4. *Biochem. J.*, **1984**, *218*, 725-732.

LANGDOWN, J.; JOHNSON, D. J.; BAGLIN, T. P.; HUNTINGTON, J. A.: Allosteric activation of antithrombin critically depends upon hinge region extension. *J. Biol. Chem.*, **2004**, *279*, 47288-47297.

LANGDOWN, L.; BELZAR, K. J.; SAVORY, W. J.; BAGLIN, T. A.; HUNTINGTON, J. A.: The critical role of hinge-region expulsion in the induced-fit heparin binding mechanism of antithrombin. *J. Mol. Biol.*, **2009**, *386*, 1278-1289.

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.

LAURENT, T. C.; TENGBLAD, A.; THUNBERG, L.; HÖÖK, M.; LINDAHL, U.: The molecular-weight-dependence of the anti-coagulant activity of heparin. *Biochem. J.*, **1978**, *175*, 691-701.

LAWSON, J. H.; KALAFATIS, M.; STRAM, S.; MANN, K. G.: A model for the tissue factor pathway to thrombin. I. An empirical study. *J. Biol. Chem.*, **1994**, *269*, 23357-23366.

LECHTENBERG, B. C.; JOHNSON, D. J. D.; FREUND, S. M. V.; HUNTINGTON, J. A.: NMR resonance assignments of thrombin reveal the conformational and dynamic effects of ligation. *Proc. Natl. Acad. Sci. U. S. A.*, **2010**, *107*, 14087-14092.

LECHTENBERG, B. C.; FREUND, S. M. V.; HUNTINGTON, J. A.: An ensemble view of thrombin allostery. *Biol. Chem.*, **2012**, *393*, 889-898.

LEE, C.J.; LIN, P.; CHANDRASEKARAN, V.; DUKE, R.E.; EVERSE, S.J.; PERERA, L.; PEDERSEN, L.G.: Proposed structural models of human factor Va and prothrombinase. *J. Thromb. Haemost.*, **2008**, *6*, 83-89.

LEMKUL, J. A.; ALLEN, W. J.; BEVAN, D. R.: Practical considerations for building GROMOS-compatible small-molecule topologies. *J. Chem. Inf. Model.*, **2010**, *50*, 2221-2235

LEYMARIE, N.; ZAIA, J.: Effective use of mass spectrometry for glycan and glycopeptide structural analysis. *Anal. Chem.*, **2012**, *84*, 3040-3048.

LI, W.; JOHNSON, D. J. D.; ESMON, C. T.; HUNTINGTON, J. A.: Structure of the antithrombin–thrombin–heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nat. Struct. Mol. Biol.*, **2004**, *11*, 857-862.

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.

LI, Y. -H.; KUO, C. -H.; SHI, G. -Y.; WU, H. -L.: The role of thrombomodulin lectin-like domain in inflammation. *J. Biomed. Sci.*, **2012**, *19*:34, 1-8.

LIMA, L. M.; BECKER, C. F.; GIESEL, G. M.; MARQUES, A. F.; CARGNELUTTI, M. T.; DE OLIVEIRA NETO, M.; MONTEIRO, R. Q.; VERLI, H.; POLIKARPOV, I.: Structural and thermodynamic analysis of thrombin:suramin interaction in solution and crystal phases. *Biochim. Biophys. Acta*, **2009**, *1794*, 873-881.

LINS, R. D.; HÜNENBERGER, P. H.: A New GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Comput. Chem.*, **2005**, *26*, 1400-1412.

LIWO, A.; KHALILI, M.; CZAPLEWSKI, C.; KALINOWSKI, S.; OLDZIEJ, S.; WACHUCIK, K.; SCHERAGA, H. A.: Modification and Optimization of the United-Residue (UNRES) potential energy function for canonical simulations. I. temperature dependence of the effective energy function and tests of the optimization method with single training proteins. *J. Phys. Chem. B*, **2007**, *111*, 260-285.

LOMMERSE, J. P.; KROON-BATENBURG, L. M.; KAMERLING, J. P.; VLIEGENTHART, J. F.: Conformational analysis of the Xylose-containing N-Glycan of pineapple stem bromelain as part of the intact glycoprotein. *Biochemistry*, **1995**, *34*, 8196-8206.

LONGAS, M. O.; FERGUSON, W. S.; FINLAY, T. H.: Studies on the interaction of heparin with thrombin, antithrombin, and other plasma proteins. *Arch. Biochem. Biophys.*, **1980**, *200*, 595-602.

LOPEZ, C. A.; RZEPIELA, A. J.; DE VRIES, A. H.; DIJKHUIZEN, L.; HÜNENBERGER, P. H.; MARRINK, S. J.: Martini coarse-grained force field: extension to carbohydrates. *J. Chem. Theory Comput.*, **2009**, *5*, 3195-3210.

LOWE, J. B.; MARTH, J. D.: A genetic approach to mammalian glycan function. *Annu. Rev. Biochem.*, **2003**, *72*, 643-691.

LU, Z.; HU, H.; YANG, W.; MARSZALEK, P. E.: Simulating force-induced conformational transitions in polysaccharides with the SMD replica exchange method. *Biophys. J.*, **2006**, *91*, L57-L59.

LÜTHY, R.; BOWIE, J. U.; EISENBERG, D.: Assessment of protein models with three-dimensional profiles. *Nature*, **1992**, *356*, 83-85.

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.; 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.

LYMAN, E.; ZUCKERMAN, D. M.: Ensemble-Based convergence analysis of biomolecular trajectories. *Biophys. J.*, **2006**, *91*, 164-172.

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^{RD}; 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.

MALM, K.; ARNLJOTS, B.; DAHLBÄCK, B.: Human activated protein C variants in a rat model of arterial thrombosis. *Thromb. J.*, **2008**, *6*:16, 1-9.

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.

MARTIN-PASTOR, M.; BUSH, C. A.: New strategy for the conformational analysis of carbohydrates based on NOE and ^{13}C NMR coupling constants. Application to the flexible polysaccharide of *Streptococcus mitis* J22. *Biochemistry*, **1999**, *38*, 8045-8055.

MCCAMMON, J. A.; GELIN, B. R.; KARPLUS, M.: Dynamics of folded proteins. *Nature*, **1977**, *267*, 585-590.

MCCOY, A. J.; PEI, X. Y.; SKINNER, R.; ABRAHAMS, J. P.; CARRELL, R. W.: Structure of β-antithrombin and the effect of glycosylation on antithrombin's heparin affinity and activity. *J. Mol. Biol.*, **2003**, *326*, 823-833.

MCNAUGHT, A. D.: International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology - Joint Commission on Biochemical Nomenclature -

Nomenclature of carbohydrates - Recommendations 1996. *Pure Appl. Chem.*, **1996**, *68*, 1919-2008.

MEAGHER, J. L.; OLSON, S. T.; GETTINS, P. G. W.: Critical role of the linker region between helix D and strand 2A in heparin activation of antithrombin. *J. Biol. Chem.*, **2000**, *275*, 2698-2704.

MEHTA, D. P.; ICHIKAWA, M.; SALIMATH, P. V.; ETCHISON, J. R.; HAAK, R.; MANZI, A.; FREEZE, H. H.: A lysosomal cysteine proteinase from *Dictyostelium discoideum* contains N-acetylglucosamine-1-phosphate bound to serine but not mannose-6-phosphate on N-linked oligosaccharides. *J. Biol. Chem.*, **1996**, *271*, 10897-10903.

MERTENS, K.; BERTINA., R. M.: Pathways in the activation of human coagulation factor X. *Biochem. J.*, **1980**, *185*, 647-658.

MEZEI, M.: The finite difference thermodynamic integration, tested on calculating the hydration free energy difference between acetone and dimethylamine in water. *J. Chem. Phys.*, **1987**, *86*, 7084-7088.

MILTON, M. J.; HARRIS, M. A.; PROBERT, M. A.; FIELD, R. A.; HOMANS, S. W.: New conformational constraints in isotopically (13C) enriched oligosaccharides. *Glycobiology*, **1998**, *8*, 147-153.

MISENHEIMER, T. M.; SHEEHAN, J. P.: The regulation of factor IXa by supersulfated low molecular weight heparin. *Biochemistry*, **2010**, *49*, 9997-10005.

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.

MONTICELLI, L.; KANDASAMY, S. K.; PERIOLE, X.; LARSON, R. G.; TIELEMAN, D. P.; MARRINK, S. J.: The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.*, **2008**, *4*, 819-834.

MOSIER, P. D.; KRISHNASAMY, C.; KELLOGG, G. E.; DESAI, U. R.: On the specificity of heparin/heparan sulfate binding to proteins. Anion-binding sites on antithrombin and thrombin are fundamentally different. *PLoS One*, **2012**, *7*, e48632.

MUKHOPADHYAY, C.: Molecular Dynamics Simulation of Glycoprotein-Glycans of Immunoglobulin G and Immunoglobulin M. *Biopolymers*, **1998**, *45*, 177-190.

MULINARI, F.; BECKER-RITT, A. B.; DEMARTINI, D. R.; LIGABUE-BRAUN, R.; STANISÇUASKI, F.; VERLI, H.; FRAGOSO, R. R.; SCHROEDER, E. K.; CARLINI, C. R.; GROSSI-DE-SÁ, M. F.: Characterization of JBURE-IIb isoform of Canavalia ensiformis (L.) DC urease. *Biochim. Biophys. Acta*, **2011**, *1814*, 1758-1768.

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.

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.; 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.

NAKATOMI, Y.; TSUJI, M.; GOKUDAN, S.; HANADA-DATEKI, T.; NAKASHIMA, T.; MIYAZAKI, H.; HAMAMOTO, T.; NAKAGAKI, T.; TOMOKIYO, K.: Stable complex formation between serine protease inhibitor and zymogen: coagulation factor X cleaves the Arg393-Ser394 bond in a reactive centre loop of antithrombin in the presence of heparin. *J. Biochem.*, **2012**, *152*, 463-470.

NAVARRO-FERNÁNDEZ, J.; PÉREZ-SÁNCHEZ, H.; MARTÍNEZ-MARTÍNEZ, I.; MELICIANI, I.; GUERRERO, J. A.; VICENTE, V.; CORRAL, J.; WENZEL, W.: In silico discovery of a compound with nanomolar affinity to antithrombin causing partial activation and increased heparin affinity. *J. Med. Chem.*, **2012**, *55*, 6403-6412.

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.

NILSSON, B.; HORNE, M. K. R.; GRALNICK, H. R.: The carbohydrate of human thrombin: structural analysis of glycoprotein oligosaccharides by mass spectrometry. *Arch. Biochem. Biophys.*, **1983**, *224*, 127-133.

NISHIDA, Y.; OHRUI, H.; MEGURO, H.: 1H-NMR Studies of (6R)-and (6S)-Deuterated D-Hexoses: Assignment of the Preferred Rotamers about C5-C6 Bond of D-Glucose and D-Galactose Derivatives in Solutions. *Tetrahedron Lett.*, **1984**, *25*, 1575-1578.

NISHIDA, Y.; HORI, H.; OHRUI, H.; MEGURO, H.: 1H NMR Analyses of Rotameric Distribution of C5-C6 bonds of D-Glucopyranoses in Solution. *J. Carbohydr. Chem.*, **1988**, *7*, 239-250.

NOSÉ, S.; KLEIN, M. L.: Constant pressure molecular dynamics for molecular systems. *Mol. Phys.*, **1983**, *50*, 1055-1076.

NOSÉ, S.: A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, **1984**, *52*, 255-268.

O´BRIEN, L. A.; STAFFORD, A. R.; FREDENBURGH, J. C.; WEITZ, J. I.: Glycosaminoglycans bind factor Xa in a Ca2+-dependent fashion and modulate its catalytic activity. *Biochemistry*, **2003**, *42*, 13091-13098.

OHRUI, H.; NISHIDA, Y.; HIGUCHI, H.; HORI, H.; MEGURO, H.: The preferred rotamer about the C5-C6 bond of D-galactopyranoses and the stereochemistry of dehydrogenation by D-galactose oxidase. *Can. J. Chem.*, **1987**, *65*, 1145-1153.

OLSON, S. T.; SRINIVASAN, K. R.; BJÖRK, I.; SHORE, J. D.: Binding of high affinity heparin to antithrombin III. Stopped flow kinetic studies of the binding interaction. *J. Biol. Chem.*, **1981**, *256*, 11073-11079.

OLSON, S. T.; BJÖRK, I.; SHEFFER, R.; CRAIG, P. A.; SHORE, J. D.; CHOAY, J.: Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. *J. Biol. Chem.*, **1992**, *267*, 12528-12538.

OLSON, S. T.; FRANCES-CHMURA, A. M.; SWANSON, R.; BJÖRK, I.; ZETTLMEISSL, G.: Effect of individual carbohydrate chains of recombinant antithrombin on heparin affinity and on the generation of glycoforms differing in heparin affinity. *Arch. Biochem. Biophys.*, **1997**, *341*, 212-221.

OLSON, S. T.; SWANSON, R.; RAUB-SEGALL, E.; BEDSTED, T.; SADRI, M.; PETITOU, M.; HERAULT, J. -P.; HERBERT, J. -M.; BJÖRK, I.: Accelerating ability of synthetic oligosaccharides on antithrombin inhibition of proteinases of the clotting and fibrinolytic systems. Comparison with heparin and low-molecular-weight heparin. *Thromb. Haemost.*, **2004**, *92*, 929-939.

OLSON, S. T.; RICHARD, B.; IZAGUIRRE, G.; SCHEDIN-WEISS, S.; GETTINS, P. G. W.: Molecular mechanisms of antithrombineheparin regulation of blood clotting proteinases. A paradigm for understanding proteinase regulation by serpin family protein proteinase inhibitors. *Biochimie*, **2010**, *92*, 1587-1596.

OOSTA, G. M.; GARDNER, W. T.; BEELER, D. L.; ROSENBERG, R. D.: Multiple functional domains of the heparin molecule. *Proc. Natl. Acad. Sci. U.S.A.*, **1981**, *78*, 829-833.

OOSTENBRINK, C.; VILLA, A.; MARK, A. E.; VAN GUNSTEREN, W. F.: A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.*, **2004**, *25*, 1656-1676.

OSCHATZ, C.; MAAS, C.; LECHER, B.; JANSEN, T.; BJÖRKQVIST, J.; TRADLER, T.; SEDLMEIER, R.; BURFEIND, P.; CICHON, S.; HAMMERSCHMIDT, S.; MÜLLER-ESTERL, W.; WUILLEMIN, W.A.; NILSSON, G.; RENNÉ, T.: Mast cells increase vascular permeability by heparin-initiated bradykinin formation in vivo. *Immunity*, **2011**, *34*, 258-268.

PAN, D.; SONG, Y.: Role of altered sialylation of the I-like domain of β1 integrin in the binding of fibronectin to β1 integrin: thermodynamics and conformational analyses. *Biophys. J.*, **2010**, *99*, 208-217.

PARFONDY, A.; CYR, N.; PERLIN, A. S.: ^{13}C-^{1}H inter-residue, coupling in disaccharides, and the orientations of glycosidic bonds. *Carbohydr. Res.*, **1977**, *59*, 299-309.

PARRINELLO, M.; RAHMAN, A.: Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, **1981**, *52*, 7182-7190.

PEREIRA, C. S.; KONY, D.; BARON, R.; MÜLER, M.; VAN GUNSTEREN, W. F.; HÜNENBERGER, P. H.: Conformational and dynamical properties of disaccharides in water: a molecular dynamics study. *Biophys. J.*, **2006**, *90*, 4337-4344.

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.; TARAVEL, F.; VERGELATI, C.: Experimental evidences of solvent induced conformational changes in maltose. *Nouv. J. Chim.*, **1985**, *4*, 561-564.

PÉREZ, S.; KOUWIJZER, M.; MAZEAU, K.; ENGELSEN, S. B.: Modeling polysaccharides: Present status and challenges. *J. Mol. Graph.*, **1996**, *14*, 307-321.

PÉREZ, S.; IMBERTY, A.; ENGELSEN, S. B.; GRUZA, J.; MAZEAU, K.; JIMENEZ-BARBERO, J.; POVEDA, A.; ESPINOSA, J.-F.; VAN EYCK, B. P.; JOHNSON, G.; FRENCH, A. D.; KOUWIJZER, M. L. C. E.; GROOTENUIS, P. D. J.; BERNARDI, A.; RAIMONDI, L.; SENDEROWITZ, H.; DURIER, V.; VERGOTEN, G.; RASMUSSEN, K.: A comparison and chemometric analysis of several molecular mechanics force fields and parameter sets applied to carbohydrates. *Carbohydr. Res.*, **1998**, *314*, 141-155.

PÉREZ, S.; KOUWIJZER, M.: In Carbohydrates: Structures, Dynamics and Syntheses; Finch, P., Ed.; Kluwer Academic Press: Dordrecht, The Netherlands, 1999; pp 258-293.

[5]PERIC-HASSLER, L.; HANSEN, H. S.; BARON, R.; HÜNENBERGER, P. H.: Conformational properties of glucose-based disaccharides investigated using molecular dynamics simulations with local elevation umbrella sampling. *Carbohydr. Res.*, **2010**, *345*, 1781-1801.

PETERSEN, B. O.; VINOGRADOV, E.; KAY, W.; WÜRTZ, P.; NYBERG, N. T.; DUUS, J. Ø.; SØRENSEN, O. W.: H2BC: a new technique for NMR analysis of complex carbohydrates. *Carbohydr. Res.*, **2006**, *341*, 550-556.

PETERSON, C. B.; BLACKBURN, M. N.: Isolation and characterization of an antithrombin III variant with reduced carbohydrate content and enhanced heparin binding. *J. Biol. Chem.*, **1985**, *260*, 610-615.

PETRERA, N. S.; STAFFORD, A. R.; LESLIE, B. A.; KRETZ, C. A.; FREDENBURGH, J. C.; WEITZ, J. I.: Long range communication between exosites 1 and 2 modulates thrombin function. *J. Biol. Chem.*, **2009**, *284*, 25620-25629.

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.

PICARD, V.; ERSDAL-BADJU, E.; BOCK, S. C.: Partial glycosylation of antithrombin III asparagine-135 is caused by the serine in the third position of its N-glycosylation consensus sequence and is responsible for production of the b-antithrombin III isoform with enhanced heparin affinity. *Biochemistry*, **1995**, *34*, 8433-8440.

PILOBELLO, K. T.; MAHAL, L. K.: Deciphering the glycocode: the complexity and analytical challenge of glycomics. *Curr. Opin. Chem. Biol.*, **2007**, *11*, 300-305.

PLEMATL, A.; DEMELBAUER, U. M.; JOSIC, D.; RIZZI, A.: Determination of the site-specific and isoform-specific glycosylation in human plasmaderived antithrombin by IEF and capillary HPLC-ESI-MS/MS. *Proteomics*, **2005**, *5*, 4025-4033.

[5]A partir de uma desatenção no processo de revisão do Trabalho IV da presente Tese, as referências 27 e 28 de tal artigo são as mesmas. Sendo assim, a referência aqui destacada corresponde à de número 28 no trabalho.

POL-FACHIN, L.; VERLI, H.: Depiction of the forces participating in the 2-*O*-sulfo-α-L-iduronic acid conformational preference in heparin sequences in aqueous solutions. *Carbohydr. Res.*, **2008**, *343*, 1435-1445.

POL-FACHIN, L.; FERNANDES, C. L.; VERLI, H.: GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations. *Carbohydr. Res.*, **2009**, *344*, 491-500.

POL-FACHIN, L.: Descrição conformacional de carboidratos e glicoproteínas: validação de protocolo baseado em dinâmica molecular e implicações funcionais. *Dissertação de Mestrado*. Universidade Federal do Rio Grande do Sul, Brasil, **2009**.

POL-FACHIN, L.; BECKER, C. F.; GUIMARÃES, J. A.; VERLI, H.: Effects of glycosylation on heparin binding and antithrombin activation by heparin. *Proteins*, **2011**, *79*, 2735-2745.

POL-FACHIN, L.; VERLI, H.: Assessment of glycoproteins dynamics from computer simulations. *Mini-Rev. Org. Chem.*, **2011**, *8*, 229-238.

POTEMPA, J.; KORZUS, E.; TRAVIS, J.: The serpin superfamily of proteinase inhibitors: structure, function, and regulation. *J. Biol. Chem.*, **1994**, *269*, 15957-15960.

POVEDA, A.; VICENT, C.; PENADES, S.; JIMENEZ-BARBERO, J.: NMR experiments for the detection of NOEs and scalar coupling constants between equivalent protons in trehalose-containing molecules. *Carbohydr. Res.*, **1997**, *301*, 5-10.

PRICE, D. J.; BROOKS III, C. L.: Modern protein force fields behave comparably in molecular dynamics simulations. *J. Comput. Chem.*, **2002**, *23*, 1045-1057.

PRICE, N. J.; PINHEIRO, C.; SOARES, C. M.; ASHFORD, D. A.; RICARDO, C. P.; JACKSON, P. A.: A biochemical and molecular characterization of LEP1, an extension peroxidase from lupin. *J. Biol. Chem.*, **2003**, *278*, 41389-41399.

PRYZDIAL, E. G. G.; KESSLER, G. E.: Kinetics of blood coagulation factor Xaα: autoproteolytic conversion to factor Xaβ: effect on inhibition by antithrombin, prothrombinase assembly, and enzyme activity. *J. Biol. Chem.*, **1996**, *28*, 16621-16626.

RAMAN, R.; SASISEKHARAN, V.; SASISEKHARAN, R.: Structural insights into biological roles of protein-glycosaminoglycan interactions. *Chem. Biol.*, **2005**, *12*, 267-277.

RAMAN, E. P.; GUVENCH, O.; MACKERELL, A. D. J.: CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J. Phys. Chem. B*, **2010**, *114*, 12981-12994.

RAND, M. D.; LOCK, J. B.; VAN´T VEER, C.; GAFFNEY, D. P.; MANN, K. G.: Blood clotting in minimally altered whole blood. *Blood*, **1996**, *88*, 3432-3445.

RAO, V. S. R.; QASBA, P. K.; BALAJI, P. V.; CHANDRASEKARAN, R.: Conformation of Disaccharides. In Conformation of Carbohydrates; Rao, V. S. R., Ed.; Harwood Academic: The Netherlands, 1998; pp 91-130.

RASCHE, H.: Haemostasis and thrombosis: an overview. *Eur. Heart J. Suppl.*, **2001**, *3*, Q3-Q7.

RAU, J. C.; BEAULIEU, L. M.; HUNTINGTON, J. A.; CHURCH, F. C.: Serpins in thrombosis, hemostasis and fibrinolysis. *J. Thromb. Haemost.*, **2007**, *5*, 102-115.

RECACHA, R.; COSTANZO, M. J.; MARYANOFF, B. E.; CARSON, M.; DELUCAS, L.; CHATTOPADHYAY, D.: Structure of human alpha-thrombin complexed with RWJ-51438 at 1.7 A: unusual perturbation of the 60A-60I insertion loop. *Acta Crystallogr. D Biol. Crystallogr.*, **2000**, *56*, 1395-1400.

REEVES, R. E.: The shape of pyranose rings. *J. Am. Chem. Soc.*, **1950**, *72*, 1499-1506.

REIF, M. M.; WINGER, M.; OOSTENBRINK, C.: Testing of the GROMOS force-field parameter set 54A8: structural properties of electrolyte solutions, lipid bilayers, and proteins. *J. Chem. Theory Comput.*, **2013**, *9*, 1247-1264.

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.

REZAIE, A. R.: Calcium enhances heparin catalysis of the antithrombin-factor Xa reaction by a template mechanism. Evidence that calcium alleviates GLA domain antagonism of heparin binding to factor Xa. *J. Biol. Chem.*, **1998**, *273*, 16824-16827.

REZAIE, A. R.: Identification of basic residues in the heparin-binding exosite of factor Xa critical for heparin and factor Va binding. *J. Biol. Chem.*, **2000**, *275*, 3320-3327.

REZAIE, A. R.; OLSON, S. T.: Calcium enhances heparin catalysis of the antithrombin-factor Xa reaction by promoting the assembly of an intermediate heparin-antithrombin-factor Xa bridging complex. Demonstration by rapid kinetics studies. *Biochemistry*, **2000**, *39*, 12083-12090.

RICCI, C. G.; DE ANDRADE, A. S.; MOTTIN, M.; NETZ, P. A.: Molecular dynamics of DNA: comparison of force fields and terminal nucleotide definitions. *J. Phys. Chem. B*, **2010**, *114*, 9882-9893.

ROSENFELD, L.; DANISHEFSKY, I.: Effects of enzymatic deglycosylation on the biological activities of human thrombin and antithrombin. *Arch. Biochem. Biophys.*, **1984**, *229*, 359-367.

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.

SALOMON-FERRER, R.; CASE, D. A.; WALKER, R. C.: An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.*, **2012**, doi: 10.1002/wcms.1121

SAMPAIO, L.O.; TERSARIOL, I. L. S.; LOPES, C. C.; BOUÇAS, R. I.; NASCIMENTO, F. D.; ROCHA, H. A. O.; NADER, H. B.: Heparins and heparan sulfates. Structure, distribution and protein interactions. In "*Insights into Carbohydrate Structure and Biological Function*", 2006, Hugo Verli (editor), Kerala: Transworld Research Network, 1-24.

SANBONMATSU, K. Y.; TUNG, C.-S.: Large-scale simulations of the ribosome: a new landmark in computational biology. *J. Phys.: Conf. Ser.*, **2006**, *46*, 334-342.

SANCHEZ, R.; PIEPER, U.; MIRKOVIC, N.; DE BAKKER, P. I. W.; WITTENSTEIN, E.; SALI, A.: ModBase, a database of annotated comparative protein structure models. *Nucleic Acids Res.*, **2000**, *28*, 250-253.

SASISEKHARAN R.; VENKATARAMAN G.: Heparin and heparan sulfate: biosynthesis, structure and function. *Curr. Opin. Chem. Biol.*, **2000**, *4*, 626-631.

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.

SCHEDIN-WEISS, S.; DESAI, U. R.; BOCK, S. C.; GETTINS, P. G. W.; OLSON, S. T.; BJÖRK, I.: Importance of lysine 125 for heparin binding and activation of antithrombin. *Biochemistry*, **2002**, *41*, 4779-4788.

SCHEDIN-WEISS, S.; RICHARD, B.; OLSON, S. T.: Kinetic evidence that allosteric activation of antithrombin by heparin is mediated by two sequential conformational changes. *Arch. Biochem. Biophys.*, **2010**, *504*, 169-176.

SCHMIDT, M. W.; BALDRIDGE, 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.

SCHMID, N.; EICHENBERGER, A.; CHOUTKO, A.; RINIKER, S.; WINGER, M.; MARK, A. E.; VAN GUNSTEREN, W. F.: Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.*, **2011**, *40*, 843-856.

SCHNUPF, U.; WILLETT, J. L.; BOSMA, W. B.; MOMANY, F. A.: DFT studies of the disaccharide, α-maltose: relaxed isopotential maps. *Carbohydr. Res.*, **2007**, *342*, 2270-2285.

SCHULER, L. D.; VAN GUNSTEREN, W. F.: On the Choice of Dihedral Angle Potential Energy Functions for n-Alkanes. *Mol. Simul.*, **2000**, *25*, 301-319.

SCHULER, L. D.; DAURA, X.; VAN GUNSTEREN, W. F.: An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase. *J. Comput. Chem.*, **2001**, *22*, 1205-1218.

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.

SCOTT, W. R. P.; HÜNENBERGER, P. H.; TIRONI, I. G.; MARK, A. E.; BILLETER, S. R.; FENNEN, J.; TORDA, A. E.; HUBER, T.; KRUEGER, P.; VAN GUNSTEREN, W. F.: The GROMOS biomolecular simulation program. *J. Phys. Chem. A*, **1999**, *103*, 3596-3607.

SEARS, P.; WONG, C. -H.: Enzyme action in glycoprotein synthesis. *Cell. Mol. Life Sci.*, **1998**, *54*, 223-252.

SEGERS, K.; DAHLBÄCK, B.; NICOLAES, G. A. F.: Coagulation factor V and thrombophilia: Background and mechanisms. *Thromb. Haemost.*, **2007**, *98*, 530-542.

SHASHKOV, A. S.; LIPKIND, G. M.; KOCHETKOV, N. K.: Nuclear overhauser effects for methyl β-maltoside and the conformational states of maltose in aqueous solution. *Carbohydr. Res.*, **1986**, *147*, 175-182.

SHAW, D. E.; MARAGAKIS, P.; LINDORFF-LARSEN, K.; PIANA, S.; DROR, R. O.; EASTWOOD, M. P.; BANK, J. A.; JUMPER, J. M.; SALMON, J. K.; SHAN, Y.; WRIGGERS, W.: Atomic-level characterization of the structural dynamics of proteins. *Science*, **2010**, *330*, 341-346.

SHEEHAN, J. P.; SADLER, J. E.: Molecular mapping of the heparin-binding exosite of thrombin. *Proc. Natl. Acad. Sci. U. S. A.*, **1994**, *91*, 5518-5522.

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.

SILVERMAN, G. A.; BIRD, P. I.; CARRELL, R. W.; CHURCH, F. C.; COUGHLIN, P. B.; GETTINS, P. G. W.; IRVING, J. A.; LOMAS, D. A.; LUKE, C. J.; MOYER, R. W.; PEMBERTON, P. A.; REMOLD-O'DONNELL, E.; SALVESEN, G. S.; TRAVIS, J.; WHISSTOCK, J. C.: The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, mechanism of inhibition, novel functions, and a revised nomenclature. *J. Biol. Chem.*, **2001**, *276*, 33293-33296.

SKINNER, R.; ABRAHAMS, J. P.; WHISSTOCK, J. C.; LESK, A. M.; CARRELL, R. W.; WARDELL, M. R.: The 2.6 A structure of antithrombin indicates a conformational change at the heparin binding site. *J. Mol. Biol.*, **1997**, *266*, 601-609.

SKROPETA, D.: The effect of individual N-glycans on enzyme activity. *Bioorg. Med. Chem.*, **2009**, *17*, 2645-2653.

SOUSA DA SILVA, A. W.; VRANKEN, W. F.; LAUE, E. D.: ACPYPE – AnteChamber PYthon Parser interfacE. Manuscript to be submitted.

SPIRO, R. G.: Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, **2002**, *12*, 43R-56R.

SPIWOK, V.; KRÁLOVÁ, B.; TVAROSKA, I.: Modelling of β-D-glucopyranose ring distortion in different force fields: a metadynamics study. *Carbohydr. Res.*, **2010**, *345*, 530-537.

STANCA-KAPOSTA, E. C.; GAMBLIN, D. P.; COCINERO, E. J.; FREY, J.; KROEMER, R. T.; FAIRBANKS, A. J.; DAVIS, B. G.; SIMONS, J. P.: Solvent interactions and conformational choice in a core n-glycan segment: gas phase conformation of the central, branching trimannose unit and its singly hydrated complex. *J. Am. Chem. Soc.*, **2008**, *130*, 10691-10696.

STENUTZ, R.; WEINTRAUB, A.; WIDMALM, G.: The structures of Escherichia coli O-polysaccharide antigens. *FEMS Microbiol. Rev.*, **2006**, *30*, 382-403.

STEPPER, J.; SHASTRI, S.; LOO, T. S.; PRESTON, J. C.; NOVAK, P.; MAN, P.; MOORE, C. H.; HAVLÍCEK, V.; PATCHETT, M. L.; NORRIS, G. E.: Cysteine S-glycosylation, a new post-translational modification found in glycopeptides bacteriocins. *FEBS Lett.*, **2011**, *585*, 645-650.

STORTZ, C. A.; CEREZO, A. S.: Disaccharide conformational maps: 3D contours or 2D plots? *Carbohydr. Res.*, **2002**, *337*, 1861-1871.

SUNNERHAGEN, M.; OLAH, G. A.; STENFLO, J.; FORSEN, S.; DRAKENBERG, T.; TREWHELLA, J.: The relative orientation of Gla and EGF domains in coagulation factor X is altered by Ca2+ binding to the first EGF domain. A combined NMR-small angle X-ray scattering study. *Biochemistry*, **1996**, *35*, 11547-11559.

TAM, R. Y.; ROWLEY, C. N.; PETROV, I.; ZHANG, T.; AFAGH, N. A.; WOO, T. K.; BEN, R. N.: Solution conformation of C-linked antifreeze glycoprotein analogues and modulation of ice recrystallization. *J. Am. Chem. Soc.*, **2009**, *131*, 15745-15753.

TAMS, J. W.; VIND, J.; WELINDER, K. G.: Adapting protein solubility by glycosylation. N-Glycosylation mutants of Coprinus cinereus peroxidase in salt and organic solutions. *Biochim. Biophys. Acta*, **1999**, *1432*, 214-221.

THANKA CHRISTLET, T. H.; VELURAJA, K.: Database Analysis of O-Glycosylation Sites in Proteins. *Biophys. J.*, **2001**, *80*, 952-960.

THIBAUDEAU, C.; STENUTZ, R.; HERTZ, B.; KLEPACH, T.; ZHAO, S.; WU, Q.; CARMICHAEL, I.; SERIANNI, A. S.: Correlated C-C and C-O bond conformations in saccharide hydroxymethyl groups: parametrization and application of redundant 1H-1H, 13C-1H, and 13C-13C NMR Jcouplings. *J. Am. Chem. Soc.*, **2004**, *126*, 15668-15685.

TIRONI, I. G.; SPERB, R.; SMITH, P. E.; VAN GUNSTEREN, W. F.: A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.*, **1995**, *102*, 5451-5459.

TOLLEFSEN, D. M.; MAJERUS, D. W.; BLANK, M. K.: Heparin cofactor II. Purification and properties of a heparin-dependent inhibitor of thrombin in human plasma. *J. Biol. Chem.*, **1982**, *257*, 2162-2169.

TURK, B.; BRIEDITIS, I.; BOCK, S. C.; OLSON, S. T.; BJÖRK, I.: The oligosaccharide side chain on Asn-135 of a-antithrombin, absent in b-antithrombin, decreases the heparin affinity of the inhibitor by affecting the heparin-induced conformational change. *Biochemistry*, **1997**, *36*, 6682-6691.

TURKO, I. V.; FAN, B.; GETTINS, P. G. W.: Carbohydrate isoforms of antithrombin variant N135Q with different heparin affinities. *FEBS Lett.*, **1993**, *335*, 9-12.

TURNBULL, J. E.; FIELD, R. A.: Emerging glycomics technologies. *Nat. Chem. Biol.*, **2007**, *3*, 74-77.

TVAROSKA, I.; PÉREZ, S.: Conformational-energy calculations for oligosaccharides: a comparison of methods and a strategy of calculation. *Carbohydr. Res.*, **1986**, *149*, 389-410.

VAN BOECKEL, C. A. A.; GROOTENHUIS, P. D. J.; VISSER, A.: A mechanism for heparin-induced potentiation of antithrombin III. *Nat. Struct. Biol.*, **1994**, *1*, 423-425.

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.; HUENENBERGER, 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, Zurique, 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.

VENUGOPAL, H.; EDWARDS, P. J.; SCHWALBE, M.; CLARIDGE, J. K.; LIBICH, D. S.; STEPPER, J.; LOO, T.; PATCHETT, M. L.; NORRIS, G. E.; PASCAL, S. M.: Structural, dynamic, and

chemical characterization of a novel S-glycosylated bacteriocin. *Biochemistry*, **2011**, *50*, 2748-2755.

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.; CALAZANS, A.; BRINDEIRO, R.; TANURI, A.; GUIMARÃES, J. A.: Molecular dynamics analysis of HIV-1 matrix protein: clarifying differences between crystallographic and solution structures. *J. Mol. Graph. Model.*, **2007**, *26*, 62-68.

VINE, A. K.: Recent advances in haemostasis and thrombosis. *Retina*, **2009**, *29*, 1-7.

VITKUP, D.; MELAMUD, E.; MOULT, J.; SANDER, C.: Completeness in structural genomics. *Nat. Struct. Biol.*, **2001**, *8*, 559-566.

WANG, X. Y.; JI, C. G.; ZHANG, J. Z.: Exploring the molecular mechanism of stabilization of the adhesion domains of human CD2 by N-glycosylation. *J. Phys. Chem. B*, **2012**, *116*, 11570-11577.

WANG, X.; KUMAR, S.; BUCK, P. M.; SINGH, S. K.: Impact of deglycosylation and thermal stress on conformational stability of a full length murine IgG2a monoclonal antibody: observations from molecular dynamics simulations. *Proteins*, **2013**, *81*, 443-460.

WEST, M. B.; WICKHAM, S.; QUINALTY, L. M.; PAVLOVICZ, R. E.; LI, C.; HANIGAN, M. H.: Autocatalytic cleavage of human γ-glutamyl transpeptidase is highly dependent on N-glycosylation at asparagine 95. *J. Biol. Chem.*, **2011**, *286*, 28876-28888.

WHISSTOCK, J. C.; PIKE, R. N.; JIN, L.; SKINNER, R.; PEI, X. Y.; CARRELL, R. W.; LESK, A. M.: Conformational changes in serpins. II. The mechanism of activation of antithrombin by heparin. *J. Mol. Biol.*, **2000**, *301*, 1287-1305.

WHISSTOCK, J. C.; BOTTOMLEY, S. P.: Serpins' mystery solved. *Nature*, **2008**, *455*, 1189-1190.

WILDGOOSE, P.; KISIEL, W.: Activation of human factor VII by factors IXa and Xa on human bladder carcinoma cells. *Blood*, **1989**, *73*, 1888-1895.

WILSON, I. B. H.; GAVEL, Y.; VON HEIJNE, G.: Amino acid distributions around O-linked glycosylation sites. *Biochem. J.*, **1991**, *275*, 529-534.

WITMER, M. R.; HATTON, M. W.: Antithrombin III-beta associates more readily than antithrombin III-alpha with uninjured and de-endothelialized aortic wall in vitro and in vivo. *Arterioscler. Thromb.*, **1991**, *11*, 530-539.

WOLFE, S.: Gauche effect. Stereochemical consequences of adjacent electron pairs and polar bonds. *Acc. Chem. Res.*, **1972**, *5*, 102-111.

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.

WOODS, R. J.; PATHIASERIL, A.; WORMALD, M. R.; EDGE, C. J.; DWEK, R. A.: The high degree of internal flexibility observed for an oligomannose oligosaccharide does not alter the overall topology of the molecule. *Eur. J. Biochem.*, **1998**, *258*, 372-386.

WYSS, D. F.; CHOI, J. S.; LI, J.; KNOPPERS, M. H.; WILLIS, K. J.; ARULANANDAM, A. R. N.; SMOLYAR, A.; REINHERZ, E. L.; WAGNER, G.: Conformation and Function of the N-linked Glycan in the Adhesion Domain of Human CD2. *Science*, **1995**, *269*, 1273-1278.

WYSS, D. F.; WAGNER, G.: The structural role of sugars in glycoproteins. *Curr. Opin. Biotechnol.*, **1996**, *7*, 409-416.

YANG, L.; SUN, M.-F.; GAILANI, D.; REZAIE, A. R.: Characterization of a heparin-binding site on the catalytic domain of factor XIa: mechanism of heparin acceleration of factor XIa inhibition by the serpins antithrombin and C1-inhibitor. *Biochemistry*, **2009**, *48*, 1517-1524.

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.

YOUSFI, N.; SEKKAL-RAHAL, M.; SAYEDE, A.; SPRINGBORG, M.: Relaxed energetic maps of κ-Carrabiose: A DFT study. *J. Comput. Chem.*, **2010**, *31*, 1312-1320.

YUNG, S.; CHAN, T. K.: Glycosaminoglycans and proteoglycans: overlooked entities? *Perit. Dial. Int.*, **2007**, *27* (Suppl. 2), S104-S109.

ZAGROVIC, B.; PANDE, V. S.: How does averaging affect protein structure comparison on the ensemble level? *Biophys. J.*, **2004**, *87*, 2240-2246.

ZHU, X.; BORCHERS, C.; BIENSTOCK, R. J.; TOMER, K. B.: Mass spectrometric characterization of the glycosylation pattern of HIV-gp120 expressed in CHO cells. *Biochemistry*, **2000**, *39*, 11194-11204.

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.

ZWANZIG, R. W.: High temperature equation of state by a perturbation method. I. nonpolar gases. *J. Chem. Phys.*, **1954**, *22*, 1420-1426.

9 Anexos

Durante a realização desta Tese, um projeto iniciado durante meu período como aluno de Mestrado, em colaboração com o Prof. Rodrigo Vassoler Serrato (Universidade Federal do Paraná), foi concluído (trabalho I). Da mesma forma, um projeto realizado em supervisão a um então aluno de Iniciação Científica, Conrado Pedebos (aluno de Mestrado na ocasião da conclusão da presente Tese), também foi concluído (trabalho II). O protocolo para simulação de glicoproteínas, desenvolvido durante meu período como aluno de Mestrado, empregado para AT e trombina na presente Tese, foi aplicado a outras glicoproteínas (trabalhos II, IV e V), incluindo um projeto (trabalho VI) em colaboração com o grupo do Prof. Jose Manuel Estevez (Instituto de Fisiología, Biología Molecular y Neurociências – CONICET, Universidad de Buenos Aires, Argentina). Adicionalmente, outros três projetos foram concluídos, na linha de polissacarídeos sulfatados, em colaboração (trabalho VII) com o grupo da Prof. Marina Ciancia (Departamento de Química Orgánica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina) e em colaboração (trabalhos VIII e IX) com o grupo da Prof. Helena Nader (Departamento de Bioquímica, Universidade Federal de São Paulo, Escola Paulista de Medicina, Brasil). Destes resultados, surgiram trabalhos publicados, incluídos a seguir, juntamente com os *residue topology parameters* (RTP) para as dezesseis aldohexopiranoses estudadas no trabalho IV da presente Tese. Essas topologias seguem a estrutura geral do campo de força GROMOS 53A6, com três potenciais de torção adicionais, a serem adicionados no arquivo de *individual topology parameters* (ITP) de ligações covalentes (*ffbonded.itp*), conforme segue:

```
#define gd_42   180.000       6.66          1
; parâmetro em substituição ao gd_3
; em resíduos de Gal, Gul, Ido e Tal
;
#define gd_43     0.000       5.88          1
;    parâmetro em substituição ao gd_17
; de diedros C – C – C – O em carboidratos
;
#define gd_44     0.000       7.67          3
;    parâmetro em substituição ao gd_17
; de diedros C – C – C – O em carboidratos
```

9.1 Trabalho I

Carbohydr. Res., **2010**, *345*, 1922-1931

Solution conformation and dynamics of exopolysaccharides from *Burkholderia* species

Laercio Pol-Fachin, Rodrigo Vassoler Serrato, Hugo Verli

RESUMO

Exopolysaccharides (EPSs) from *Burkholderia* genus are proposed to be involved in pathological conditions in humans, as cystic fibrosis and septicemia, as well as in the stability of soil aggregates. Hence, considering that the conformational and dynamical aspects of such EPSs may influence their biological activity, the current work employs a series of molecular dynamics simulations on di-, oligo- and polysaccharide fragments of three EPSs, from *B. caribensis*, *B. cepacia*, and *B. pseudomallei*, with previously determined NOE data, in order obtain a conformational description of such EPSs at the atomic level. As the obtained results show good agreement with the experimental data, pointing to the adequacy of the employed methodology to accurately describe the dynamics of polysaccharides, the strategy was also employed to predict the conformational behavior of an additional compound, from *B. tropica*, for which NOE signals are not available. Taking into account the potential importance of EPSs on the interaction of *Burkholderia* bacteria with distinct environments, it may be expected that a greater understanding of their structural aspects may contribute to controlling their pathological roles and potential agricultural applications.

9.2 Trabalho II

J. Nat. Prod., **2012**, *75*, 1196-1200

Unrestrained conformational characterization of *Stenocereus eruca* saponins in aqueous and nonaqueous solvents

Conrado Pedebos, **Laércio Pol-Fachin**, Hugo Verli

RESUMO

Saponins are secondary metabolites that have a plethora of biological activities. However, the absence of knowledge of their 3D structures is a major drawback for structural-based strategies in medicinal chemistry. To address this problem, the current work presents structural models of *Stenocereus eruca* saponins, named erucasaponin A and stellatoside B. These compounds were constructed on the basis of a combination of unrestrained molecular dynamics (MD) simulations and NOESY data, in both pyridine and water. The models obtained in this way offer a robust description of the saponin dynamics in solution and support the use of submicrosecond MD simulations in describing and predicting glycoconjugate conformations.

9.3 Trabalho III

Glycobiology, **2012**, *22*, 817-825

Structural glycobiology of the major allergen of *Artemisia vulgaris* pollen, Art v 1: O-glycosylation influence on the protein dynamics and allergenicity

Laércio Pol-Fachin, Hugo Verli

RESUMO

Art v 1 is the major allergen of mugwort (*Artemisia vulgaris*) pollen. It is formed by an N-terminal globular defensin-like part and a C-terminal proline-rich domain. As the structure and dynamics of Art v 1 have been mostly described for its recombinant, non-glycosylated form, which does not occur in normal plant physiology, the present work intends to obtain a 3D model for Art v 1 native *O*-glycosylation structure and evaluate the influence of such glycans over the protein dynamics and allergenicity through molecular dynamics simulations in triplicates. Structural insights into the mutual recognition of Art v 1 protein and carbohydrate moieties recognition by antibodies were obtained, in which glycan chains remained close to the previously identified epitopes in the defensin-*like* domain, thus pointing to potential interferences with antibodies recognition. To our knowledge, this is the first structural report of an entire furanose-containing glycoprotein. As well, together to the previously determined NMR structures, the obtained results contribute in the comprehension of the effect of glycosylation over both proline-rich and defensin-*like* domains, providing an atomic representation of such alterations.

9.4 Trabalho IV

J. Biomol. Struct. Dyn., **2013**, *accepted,* doi: 10.1080/07391102.2013.780982

Effects of glycosylation and pH conditions in the dynamics of human arylsulfatase A

Madza Yasodara Farias Virgens, **Laércio Pol-Fachin**, Hugo Verli, Maria Luiza Saraiva-Pereira

RESUMO

Arylsulfatase A (ARSA) is a lysosomal sulfatase that catalyzes the hydrolysis of cerebroside sulfate. Its deficiency results in Metachromatic Leukodystrophy, whereas a minor condition called ARSA pseudodeficiency occurs in healthy individuals, which has been associated with the substitution of the glycosylated Asn350 by a Ser. In the present work, we have investigated ARSA dynamics employing molecular dynamics simulations in response to (1) different pH's, as it has been recently identified in cytoplasmatic medium, and (2) glycan occupancies, including its normal glycosylation state, presenting three high mannose-type oligosaccharides. Accordingly, four systems were studied considering ARSA under different conditions: (1) non-glycosylated at pH~7 ($ARSA_{pH7}$); (2) non-glycosylated at pH~5 ($ARSA_{pH5}$); (3) triple glycosylated at pH~5 ($ARSA_{glyc,pH5}$); and (4) ARSA-N350S mutant at pH~5 ($ARSA_{N350S,pH5}$). Lowering pH and increasing glycosylation was found to reduce the flexibility of the enzyme. In addition, at acidic pH, the glycosylated system presented a higher secondary conformational stability when compared to its non-glycosylated counterpart, supporting experimental findings on triple glycosylation as the essential state of ARSA. The N350S mutant exhibited a consistent degree of unfolding, which may be related to its *in vitro* reduced activity. Finally, the obtained data are discussed in the search for structural evidences able to contribute to the understanding of biological activity of ARSA and molecular etiology of ARSA pseudodeficiency, as determined by ARSA-N350S in absence of polyadenilation defect

9.5 Trabalho V

In "Strategies for the Determination of Carbohydrates Structure and Conformation", **2010**, Hugo Verli (editor), Kerala: Transworld Research Network, 103-131

Molecular modeling on the characterization of glycoproteins conformation

Laércio Pol-Fachin, Hugo Verli

RESUMO

Glycosylation is one of the most frequent and variable modification of proteins, which can occur from eubacteria to eukaryotes, depending on specific consensus sequences and cellular factors. It has been proposed that more than half of the known proteins can be potentially glycosylated, while their added glycans greatly vary in terms of size and composition. Furthermore, such carbohydrate moieties can be attached to proteins through several amino acid residues, as well as involving several different monosaccharides. The most useful methods to assess glycoproteins structure and conformation are NMR spectroscopy and X-ray crystallography, both presenting advantages and limitations. In this context, molecular modeling techniques emerge as complementary methodologies to help explaining experimentally observed properties, as well as to predict unknown features of the studied systems. In this review, we bring together some of the main molecular modeling efforts and results to the comprehension of glycoproteins conformation and structural organization. Among such topics, there are included force field parameters for MD simulations, databases of oligosaccharide structures and web-based tools for constructing three-dimensional glycoproteins, as well as aspects of protein-carbohydrate glycosidic linkages, glycans conformation and effects of glycosylation over protein structure, obtained, mainly, by quantum or molecular mechanics calculations. As a general feature, several aspects of the conformation and function of glycoproteins achievable by such methods are in accordance to previous experimental data, which reinforce their consistency. However, a lot still remains to be done, especially due to the vast amount of glycoproteins not yet studied, as well as the unknown roles of its carbohydrate moieties in biological phenomena.

9.6 Trabalho VI

Science, **2011**, *332*, 1401-1403

O-Glycosylated Cell Wall Proteins Are Essential in Root Hair Growth

Silvia M. Velasquez, Martiniano M. Ricardi, Javier Gloazzo Dorosz, Paula V. Fernandez, Alejandro D. Nadra, **Laércio Pol-Fachin**, Jack Egelund, Sascha Gille, Jesper Harholt, Marina Ciancia, Hugo Verli, Markus Pauly, Antony Bacic, Carl Erik Olsen, Peter Ulvskov, Bent Larsen Petersen, Chris Somerville, Norberto D. Iusem, Jose M. Estevez

RESUMO

Root hairs are single cells that develop by tip growth and are specialized in the absorption of nutrients. Their cell walls are composed of polysaccharides and hydroxyproline-rich glycoproteins (HRGPs) that include extensins (EXTs) and arabinogalactan-proteins (AGPs). Proline hydroxylation, an early posttranslational modification of HRGPs that is catalyzed by prolyl 4-hydroxylases (P4Hs), defines the subsequent O-glycosylation sites in EXTs (which are mainly arabinosylated) and AGPs (which are mainly arabinogalactosylated). We explored the biological function of P4Hs, arabinosyltransferases, and EXTs in root hair cell growth. Biochemical inhibition or genetic disruption resulted in the blockage of polarized growth in root hairs and reduced arabinosylation of EXTs. Our results demonstrate that correct O-glycosylation on EXTs is essential for cell-wall self-assembly and, hence, root hair elongation in *Arabidopsis thaliana*.

9.7 Trabalho VII

J. Biol. Chem., **2013**, *288*, 223-233

Anticoagulant activity of a unique sulfated pyranosic (1→3)-β-L-arabinan through direct interaction with thrombin

Paula V. Fernández, Irene Quintana, Alberto S. Cerezo, Julio J. Caramelo, **Laércio Pol-Fachin**, Hugo Verli, José M. Estevez, Marina Ciancia

RESUMO

A highly sulfated 3-linked β-arabinan (Ab1) with arabinose in the pyranose form was obtained from green seaweed *C. vermilara* (Bryopsidales). It comprised major amounts of units sulfated on C-2 and C-4, and constitutes the first polysaccharide of this type isolated in the pure form and fully characterized. Ab1 showed anticoagulant activity by global coagulation tests. Less sulfated arabinans obtained from the same seaweed, have less or no activity. Ab1 exerts its activity through direct and indirect (antithrombin- and heparin cofactor II-mediated) inhibition of thrombin. Direct thrombin inhibition was studied in detail. By native PAGE it was possible to detect formation of a complex between Ab1 and human thrombin (HT). Ab1 binding to HT was measured by fluorescence spectroscopy. CD spectra of the Ab1-complex suggested that ligand binding induced a small conformational change on HT. Ab1-thrombin interactions were studied by molecular dynamic simulations using the persulfated octasaccharide as model compound. Most carbohydrate-protein contacts would occur by interaction of sulfate groups with basic amino acid residues in the surface of the enzyme, being more than 60% of them performed by the exosite 2-composing residues. In these interactions the sulfate groups on C-2 showed to interact more intensely with thrombin structure. In contrast, the disulfated oligosaccharide does not promote major conformational modifications at the catalytic site when complexed to exosite 1. These results show that this novel pyranosic sulfated arabinan Ab1 exerts its anticoagulant activity by a mechanism different to those found previously for other sulfated polysaccharides and glycosaminoglycans.

9.8 Trabalho VIII

PLoS ONE, **2013**, *8*, e70880

Insights into the N-Sulfation mechanism: molecular dynamics simulations of the N-sulfotransferase domain of Ndst1 and mutants

Tarsis F. Gesteira, **Laércio Pol-Fachin**, Vivien Jane Coulson-Thomas, Marcelo A. Lima, Hugo Verli, Helena B. Nader

RESUMO

Sulfation patterns along glycosaminoglycan (GAG) chains dictate their functional role. The N-deacetylase N-sulfotransferase family (NDST) catalyzes the initial downstream modification of heparan sulfate and heparin chains by removing acetyl groups from subsets of N-acetylglucosamine units and, subsequently, sulfating the residual free amino groups. These enzymes transfer the sulfuryl group from 3′-phosphoadenosine-5′-phosphosulfate (PAPS), yielding sulfated sugar chains and 3′-phosphoadenosine-5′-phosphate (PAP). For the N-sulfotransferase domain of NDST1, Lys833 has been implicated to play a role in holding the substrate glycan moiety close to the PAPS cofactor. Additionally, Lys833 together with His716 interact with the sulfonate group, stabilizing the transition state. Such a role seems to be shared by Lys614 through donation of a proton to the bridging oxygen of the cofactor, thereby acting as a catalytic acid. However, the relevance of these boundary residues at the hydrophobic cleft is still unclear. Moreover, whether Lys833, His716 and Lys614 play a role in both glycan recognition and glycan sulfation remains elusive. In this study we evaluate the contribution of NDST mutants (Lys833, His716 and Lys614) to dynamical effects during sulfate transfer using comprehensive combined docking and essential dynamics. In addition, the binding location of the glycan moiety, PAPS and PAP within the active site of NDST1 throughout the sulfate transfer were determined by intermediate state analysis. Furthermore, NDST1 mutants unveiled Lys833 as vital for both the glycan binding and subsequent N-sulfotransferase activity of NDST1.

9.9 Trabalho IX

Molecular BioSystems, **2014**, *10*, 54-64

On the catalytic mechanism of polysaccharide lyases: evidence of His and Tyr involvement in heparin lysis by heparinase I and the role of Ca^{2+}

Carolina R. Córdula, Marcelo A. Lima, Samuel K. Shinjo, Tarsis F. Gesteira, **Laércio Pol-Fachin**, Vivien J. Coulson-Thomas, Hugo Verli, Edwin A. Yates, Timothy R. Rudd, Maria A. S. Pinhal, Leny Toma, Carl P. Dietrich, Helena B. Nader, Ivarne L. S. Tersariol

RESUMO

The structurally diverse polysaccharide lyase enzymes are distributed from plants to animals but share common catalytic mechanisms. One, heparinase I (F. heparinum), is employed in the production of the major anticoagulant drug, low molecular weight heparin, and is a mainstay of cell surface proteoglycan analysis. We demonstrate that heparinase I specificity and efficiency depend on the cationic form of the substrate. Ca^{2+}–heparin, in which α-L-iduronate-2-O-sulfate residues adopt $^1C_4$ conformation preferentially, is a substrate, while Na+–heparin is an inhibitor. His and Tyr residues are identified in the catalytic step and a model based on molecular dynamics and docking is proposed, in which deprotonated His203 initiates β-elimination by abstracting the C5 proton of the α-L-iduonate-2-O-sulfate residue in the substrate, and protonated Tyr357 provides the donor to the hexosamine leaving group.

9.10 Topologias para carboidratos – GROMOS 53A6GLYC

α-Allose

```
[ ALLA ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C4    O3    C2    C3     gi_2
   C5    O4    C3    C4     gi_2
   C1    C3    O2    C2     gi_2
   C1    O5    O1    C2     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    O5     gd_5
   O6    C6    C5    O5     gd_37
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_6
   O5    C1    O1   HO1     gd_28
 [ exclusions ]
  HO1    O5
```

β-Allose

```
[ ALLB ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C4    O3    C2    C3     gi_2
   C5    O4    C3    C4     gi_2
   C1    C3    O2    C2     gi_2
   C2    O5    O1    C1     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    O5     gd_5
   O6    C6    C5    O5     gd_37
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_2
   O5    C1    O1   HO1     gd_32
 [ exclusions ]
  HO1    O5
```

α-Altrose

```
[ ALTA ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C4    O3    C2    C3     gi_2
   C5    O4    C3    C4     gi_2
   C2    C3    O2    C1     gi_2
   C1    O5    O1    C2     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    O5     gd_5
   O6    C6    C5    O5     gd_37
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_6
   O5    C1    O1   HO1     gd_28
 [ exclusions ]
  HO1    O5
```

β-Altrose

```
[ ALTB ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C4    O3    C2    C3     gi_2
   C5    O4    C3    C4     gi_2
   C2    C3    O2    C1     gi_2
   C2    O5    O1    C1     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    O5     gd_5
   O6    C6    C5    O5     gd_37
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_2
   O5    C1    O1   HO1     gd_32
 [ exclusions ]
  HO1    O5
```

α-Galactose

```
[ GALA ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C3    O3    C2    C4     gi_2
   C4    O4    C3    C5     gi_2
   C1    C3    O2    C2     gi_2
   C1    O5    O1    C2     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    C4     gd_1
   O6    C6    C5    O5     gd_42
   O6    C6    C5    O5     gd_35
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_6
   O5    C1    O1   HO1     gd_28
 [ exclusions ]
  HO1    O5
```

β-Galactose

```
[ GALB ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C3    O3    C2    C4     gi_2
   C4    O4    C3    C5     gi_2
   C1    C3    O2    C2     gi_2
   C2    O5    O1    C1     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    C4     gd_1
   O6    C6    C5    O5     gd_42
   O6    C6    C5    O5     gd_35
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_2
   O5    C1    O1   HO1     gd_32
 [ exclusions ]
  HO1    O5
```

α-Glicose

```
[ GLCA ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C3    O3    C2    C4     gi_2
   C5    O4    C3    C4     gi_2
   C1    C3    O2    C2     gi_2
   C1    O5    O1    C2     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    O5     gd_5
   O6    C6    C5    O5     gd_37
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_6
   O5    C1    O1   HO1     gd_28
 [ exclusions ]
  HO1    O5
```

β-Glicose

```
[ GLCB ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C3    O3    C2    C4     gi_2
   C5    O4    C3    C4     gi_2
   C1    C3    O2    C2     gi_2
   C2    O5    O1    C1     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    O5     gd_5
   O6    C6    C5    O5     gd_37
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_2
   O5    C1    O1   HO1     gd_32
 [ exclusions ]
  HO1    O5
```

α-Gulose

```
[ GULA ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C4    O3    C2    C3     gi_2
   C4    O4    C3    C5     gi_2
   C1    C3    O2    C2     gi_2
   C1    O5    O1    C2     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    C4     gd_1
   O6    C6    C5    O5     gd_42
   O6    C6    C5    O5     gd_35
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_6
   O5    C1    O1   HO1     gd_28
 [ exclusions ]
  HO1    O5
```

β-Gulose

```
[ GULB ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C4    O3    C2    C3     gi_2
   C4    O4    C3    C5     gi_2
   C1    C3    O2    C2     gi_2
   C2    O5    O1    C1     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    C4     gd_1
   O6    C6    C5    O5     gd_42
   O6    C6    C5    O5     gd_35
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_2
   O5    C1    O1   HO1     gd_32
 [ exclusions ]
  HO1    O5
```

α-Idose

```
[ IDOA ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C4    O3    C2    C3     gi_2
   C4    O4    C3    C5     gi_2
   C2    C3    O2    C1     gi_2
   C1    O5    O1    C2     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    C4     gd_1
   O6    C6    C5    O5     gd_42
   O6    C6    C5    O5     gd_35
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_6
   O5    C1    O1   HO1     gd_28
 [ exclusions ]
  HO1    O5
```

β-Idose

```
[ IDOB ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C4    O3    C2    C3     gi_2
   C4    O4    C3    C5     gi_2
   C2    C3    O2    C1     gi_2
   C2    O5    O1    C1     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    C4     gd_1
   O6    C6    C5    O5     gd_42
   O6    C6    C5    O5     gd_35
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_2
   O5    C1    O1   HO1     gd_32
 [ exclusions ]
  HO1    O5
```

α-Manose

```
[ MANA ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C3    O3    C2    C4     gi_2
   C5    O4    C3    C4     gi_2
   C2    C3    O2    C1     gi_2
   C1    O5    O1    C2     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    O5     gd_5
   O6    C6    C5    O5     gd_37
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_6
   O5    C1    O1   HO1     gd_28
 [ exclusions ]
  HO1    O5
```

β-Manose

```
[ MANB ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C3    O3    C2    C4     gi_2
   C5    O4    C3    C4     gi_2
   C2    C3    O2    C1     gi_2
   C2    O5    O1    C1     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    O5     gd_5
   O6    C6    C5    O5     gd_37
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_2
   O5    C1    O1   HO1     gd_32
 [ exclusions ]
  HO1    O5
```

α-Talose

```
[ TALA ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C3    O3    C2    C4     gi_2
   C4    O4    C3    C5     gi_2
   C2    C3    O2    C1     gi_2
   C1    O5    O1    C2     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    C4     gd_1
   O6    C6    C5    O5     gd_42
   O6    C6    C5    O5     gd_35
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_6
   O5    C1    O1   HO1     gd_28
 [ exclusions ]
  HO1    O5
```

β-Talose

```
[ TALB ]
 [ atoms ]
   C4   CH1     0.23200     0
   O4    OA    -0.64200     0
  HO4     H     0.41000     0
   C3   CH1     0.23200     1
   O3    OA    -0.64200     1
  HO3     H     0.41000     1
   C2   CH1     0.23200     2
   O2    OA    -0.64200     2
  HO2     H     0.41000     2
   C6   CH2     0.23200     3
   O6    OA    -0.64200     3
  HO6     H     0.41000     3
   C5   CH1     0.37600     4
   O5    OA    -0.48000     4
   C1   CH1     0.23200     4
   O1    OA    -0.53800     4
  HO1     H     0.41000     4
 [ bonds ]
   C4    O4    gb_20
   O4   HO4    gb_1
   C4    C3    gb_26
   C4    C5    gb_26
   C3    O3    gb_20
   C3    C2    gb_26
   O3   HO3    gb_1
   C2    O2    gb_20
   C2    C1    gb_26
   O2   HO2    gb_1
   C6    O6    gb_20
   C6    C5    gb_26
   O6   HO6    gb_1
   C5    O5    gb_20
   O5    C1    gb_20
   C1    O1    gb_20
   O1   HO1    gb_1
 [ angles ]
  HO4    O4    C4     ga_12
   O4    C4    C3     ga_9
   O4    C4    C5     ga_9
   C3    C4    C5     ga_8
   C4    C3    O3     ga_9
   C4    C3    C2     ga_8
   O3    C3    C2     ga_9
   C3    O3   HO3     ga_12
   C3    C2    O2     ga_9
   C3    C2    C1     ga_8
   O2    C2    C1     ga_9
   C2    O2   HO2     ga_12
   O6    C6    C5     ga_9
   C6    O6   HO6     ga_12
   C4    C5    C6     ga_8
   C4    C5    O5     ga_9
   C6    C5    O5     ga_9
   C5    O5    C1     ga_10
   C2    C1    O5     ga_9
   C2    C1    O1     ga_9
   O5    C1    O1     ga_9
   C1    O1   HO1     ga_12
 [ impropers ]
   C4    C6    O5    C5     gi_2
   C3    O3    C2    C4     gi_2
   C4    O4    C3    C5     gi_2
   C2    C3    O2    C1     gi_2
   C2    O5    O1    C1     gi_2
 [ dihedrals ]
  HO4    O4    C4    C3     gd_30
   O4    C4    C3    O3     gd_18
   O4    C4    C3    C2     gd_43
   O4    C4    C3    C2     gd_44
   C5    C4    C3    O3     gd_43
   C5    C4    C3    O3     gd_44
   C5    C4    C3    C2     gd_34
   O4    C4    C5    C6     gd_43
   O4    C4    C5    C6     gd_44
   C3    C4    C5    C6     gd_34
   C3    C4    C5    O5     gd_43
   C3    C4    C5    O5     gd_44
   C2    C3    O3   HO3     gd_30
   C4    C3    C2    O2     gd_43
   C4    C3    C2    O2     gd_44
   C4    C3    C2    C1     gd_34
   O3    C3    C2    O2     gd_18
   O3    C3    C2    C1     gd_43
   O3    C3    C2    C1     gd_44
   C1    C2    O2   HO2     gd_30
   C3    C2    C1    O5     gd_43
   C3    C2    C1    O5     gd_44
   C3    C2    C1    O1     gd_43
   C3    C2    C1    O1     gd_44
   O2    C2    C1    O1     gd_18
   C5    C6    O6   HO6     gd_30
   O6    C6    C5    C4     gd_1
   O6    C6    C5    O5     gd_42
   O6    C6    C5    O5     gd_35
   C4    C5    O5    C1     gd_29
   C5    O5    C1    C2     gd_29
   O5    C1    O1   HO1     gd_2
   O5    C1    O1   HO1     gd_32
 [ exclusions ]
  HO1    O5
```

10 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.

Mestrado em Biologia Celular e Molecular pelo Programa de Pós-Graduação em Biologia Celular e Molecular da pela Universidade Federal do Rio Grande do Sul, de 2008/1 a 2009/2. *Descrição conformacional de carboidratos e glicoproteínas: validação de protocolo baseado em dinâmica molecular e implicações funcionais*. Orientador: Prof. Hugo Verli.

II. Prêmios recebidos

Science Direct – Elsevier, 2010: 23º lugar na "Top25 Hottest Articles, January to March 2010, Journal of Molecular Graphics and Modelling", com o artigo “Pol-Fachin, L., Fraga, C. A. M., Barreiro, E. J. & Verli, H.: Characterization of the conformational ensemble from bioactive N-acylhydrazone derivatives. *J. Mol. Graph. Mod.*, 2010, *28*, 446-454”.

42ª Reunião Anual da SBBq, 2013: Vencedor do 17º Prêmio Jovem Talento em Ciências da Vida, promovido pela Sociedade Brasileira de Bioquímica e Biologia Molecular (SBBq) juntamente com a GE Healthcare Life Sciences.

III. 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 struture 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 Lindoia, 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 Lindoia, 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 Lindoia, SP, Brasil.

Pedebos, C., **Pol-Fachin, L.** & Verli, H. **2010** Caracterização conformacional de saponinas. *Resumos do XXII Salão de Iniciação Científica da Universidade Federal do Rio Grande do Sul*. Universidade Federal do Rio Grande do Sul, XXII Salão de Iniciação Científica e XIX Feira de Iniciação Científica da Universidade Federal do Rio Grande do Sul, 18 a 22 de outubro de 2010, Porto Alegre, RS, Brasil.

Virgens, M. Y. F., **Pol-Fachin, L.**, Verli, H. & Saraiva-Pereira, M. L. **2010** The effects of glycosylation and pH conditions in Arylsulfatase A. *Programas e Resumos da XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, N-37. XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 18 a 21 de maio de 2010, Foz do Iguaçú, PR, Brasil.

Siebert, M., Bock, H., Giugliani, R., **Pol-Fachin, L.**, Verli, H. & Saraiva-Pereira, M. L. **2010** Structural Analysis of Glucocerebrosidase and Molecular Analysis of the GBA Gene. *Programas e Resumos da XXXIX Reunião Anual da Sociedade Brasileira de*

Bioquímica e Biologia Molecular, N-25. XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 18 a 21 de maio de 2010, Foz do Iguaçú, PR, Brasil.

Pedebos, C., **Pol-Fachin, L.** & Verli, H. **2010** Depiction of saponins conformational ensemble. *Programas e Resumos da XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, M-21. XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 18 a 21 de maio de 2010, Foz do Iguaçú, PR, Brasil.

Pol-Fachin, L. & Verli, H. **2010** Effect of glycosylation on antithrombin conformation and heparin binding. *Programas e Resumos da XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, N-29. XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 18 a 21 de maio de 2010, Foz do Iguaçú, PR, Brasil.

Pedebos, C., Teixeira, C. V., **Pol-Fachin, L.** & Verli, H. **2011** Caracterização conformacional da saponina QS-21 e as implicações na formação micelar. *Resumos do XXIII Salão de Iniciação Científica da Universidade Federal do Rio Grande do Sul*. Universidade Federal do Rio Grande do Sul, XXIII Salão de Iniciação Científica e XX Feira de Iniciação Científica da Universidade Federal do Rio Grande do Sul, 3 a 7 de outubro de 2011, Porto Alegre, RS, Brasil.

Pedebos, C., **Pol-Fachin, L.** & Verli, H. **2011** Structure and conformation of QS-21 saponin: Implications for micelle formation. *Programas e Resumos da XL Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, C-4. XL Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, de 30 de abril a 3 de maio de 2011, Foz do Iguaçú, PR, Brasil.

Pol-Fachin, L. & Verli, H. **2011** Structural insights into coagulation cascade modulation by heparin. *Programas e Resumos da XL Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, C-21. XL Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, de 30 de abril a 3 de maio de 2011, Foz do Iguaçú, PR, Brasil.

Pedebos, C., **Pol-Fachin, L.** & Verli, H. **2012** Atomic Model for Micelles Composed by the Saponin QS-21. *Programas e Resumos da XLI Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, C-12. XLI Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, de 19 a 22 de maio de 2012, Foz do Iguaçú, PR, Brasil.

Pol-Fachin, L. & Verli, H. **2012** Atomic-level details on thrombin and fXa allosteric inhibition by heparin. *Programas e Resumos da XLI Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, C-17. XLI Reunião Anual da Sociedade Brasileira

de Bioquímica e Biologia Molecular, de 19 a 22 de maio de 2012, Foz do Iguaçú, PR, Brasil.

John, E. O., **Pol-Fachin, L.** & Verli, H. **2013** Structural biology of proline-rich substrates recognition by *A. thaliana* prolyl-4-hydroxylases. *Programas e Resumos da XLII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, C-20. XLII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, de 18 a 21 de maio de 2013, Foz do Iguaçú, PR, Brasil.

Alves, W. J. C., Macedo, B., Torres, P., Gonçalves, K. M., Oliveira Junior, R. S., **Pol-Fachin, L.**, Verli, H., Linden, R., Cordeiro, Y. & Pascutti, P. **2013** Conformational ensembles of glycosylated prion protein: a strategy for mapping interactions with physiological and therapeutic ligands. *Programas e Resumos da XLII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular*, C-18. XLII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, de 18 a 21 de maio de 2013, Foz do Iguaçú, PR, 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 Lindoia, 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 Lindoia, SP, Brasil.

Woicickoski, C., **Pol-Fachin, L.**, Fraga, C. A. M., Barreiro, E. J. & Verli, H. **2008** Description of the conformational ensemble of N-acylhidrazone 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 Lindoia, SP, Brasil.

IV. Publicações em periódicos especializados:

a. Internacionais:

Pol-Fachin, L. & Verli, H. **2008** Depiction of the forces participating in the 2-O-sulfo-α-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 β-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.

Pol-Fachin, L., Serrato, R. V. & Verli, H. **2010** Solution conformation and dynamics of exopolysaccharides from Burkholderia species. *Carbohydr. Res.*, *345*, 1922-1931.

Velasquez, S. M., Ricardi, M. M., Dorosz, J. G., Fernandez, P. V., Nadra, A. D., **Pol-Fachin, L.**, Egelund, J., Gille, S., Harholt, J., Ciancia, M., Verli, H., Pauly, M., Bacic, A., Olsen, C. E., Ulvskov, P., Petersen, B. L., Somerville, C., Iusem, N. D. & Estevez, J. M. **2011** O-Glycosylated Cell Wall Proteins Are Essential in Root Hair Growth. *Science*, *332*, 1401-1403.

Pol-Fachin, L. & Verli, H. **2011** Assessment of Glycoproteins Dynamics from Computer Simulations. *Mini Rev. Org. Chem.*, *8*, 229-238.

Pol-Fachin, L., Becker, C. F., Guimarães, J. A. & Verli, H. **2011** Effects of glycosylation on heparin binding and antithrombin activation by heparin. *Proteins*, *79*, 2735-2745.

Pol-Fachin, L. & Verli, H. **2012** Structural glycobiology of the major allergen of Artemisia vulgaris pollen, Art v 1: O-glycosylation influence on the protein dynamics and allergenicity. *Glycobiology*, *22*, 817-825.

Pedebos, C., **Pol-Fachin, L.** & Verli, H. **2012** Unrestrained conformational characterization of *Stenocereus eruca* saponins in aqueous and nonaqueous solvents. *J. Nat. Prod.*, *75*, 1196-1200.

Pol-Fachin, L., Rusu, V. H., Verli, H. & Lins, R. D. **2012** GROMOS 53A6$_{GLYC}$, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Chem. Theory Comput.*, *8*, 4681-4690.

Fernández, P. V., Quintana, I., Cerezo, A. S., Caramelo, J. J., **Pol-Fachin, L.**, Verli, H., Estevez, J. M. & Ciancia, M. **2013** Anticoagulant activity of a unique sulfated pyranosic (1→3)-β-L-arabinan through direct interaction with thrombin. *J. Biol. Chem.*, *288*, 223-233.

Gesteira, T. F., **Pol-Fachin, L.**, Coulson-Thomas, V. J.. Lima, M. A., Verli, H. & Nader, H. B. **2013** Insights into the N-sulfation mechanism: molecular dynamics simulations of the N-sulfotransferase domain of Ndst1 and mutants. *PLoS ONE*, *8*, e70880.

Pol-Fachin, L. & Verli, H. **2014** Structural glycobiology of heparin dynamics on the exosite 2 of coagulation cascade proteases: Implications for glycosaminoglycans antithrombotic activity. *Glycobiology*, *24*, 97-105.

Córdula, C. R., Lima, M. A., Shinjo, S. K., **Pol-Fachin, L.**, Coulson-Thomas, V. J., Verli, H., Yates, E. A., Rudd, T. R., Pinhal, M. A. S., Toma, L., Dietrich, C. P., Nader, H. B. & Tersariol, I. L. S. **2014** On the catalytic mechanism of polysaccharide lyases: evidence of His and Tyr involvement in heparin lysis by heparinase I and the role of Ca^{2+}. *Molecular BioSystems*, *10*, 54-64.

V. 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.

Conrado Pedebos. *Caracterização conformacional de saponinas em piridina*. Centro de Biotecnologia, UFRGS, de julho de 2008 a dezembro de 2011.

b. Em andamento:

Elisa Beatriz de Oliveira John. *Caracterização da interação entre as enzimas prolil-hidroxilases P4H5, P4H2 e P4H13 e extensinas de Arabidopsis thaliana*. Centro de Biotecnologia, UFRGS, a partir de julho de 2012.

VI. 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 de março de 2008 a julho de 2009.

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