



Rui Miguel Alegria Borralho

Licenciado em Bioquímica

Reconhecimento molecular por RMN da Lectina-galactose presente em Macrófagos

Dissertação para obtenção do Grau de Mestre em
Química Bioorgânica

Orientador: Doutora Filipa Marcelo, Investigadora Auxiliar, Faculdade de
Ciências e Tecnologias – Universidade Nova de Lisboa

Júri:

Presidente de Júri: Professora Doutora Paula Cristina do Sério Branco

Arguente: Doutor Aldino José Martins Viegas

Vogal: Doutora Filipa Margarida Barradas de Morais Marcelo

Rui Miguel Alegria Borralho

Licenciado em Bioquímica

**Reconhecimento molecular por RMN da
Lectina-galactose presente em Macrófagos**

Dissertação para obtenção do Grau de Mestre em
Química Bioorgânica

Orientador: Doutora Filipa Marcelo, Investigadora Auxiliar, Faculdade de
Ciências e Tecnologias – Universidade Nova de Lisboa

Júri:

Presidente de Júri: Professora Doutora Paula Cristina do Sério Branco

Arguente: Doutor Aldino José Martins Viegas

Vogal: Doutora Filipa Margarida Barradas de Moraes Marcelo



Setembro, 2019

Reconhecimento molecular por RMN da Lectina-galactose presente em Macrófagos

Copyright © Rui Miguel Alegria Borralho, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Agradecimentos

Gostaria de começar por agradecer à minha orientadora Doutora Filipa Marcelo por me ter aceite e por ter tornado possível a realização deste trabalho. Agradeço-lhe também toda a sua dedicação, orientação e conhecimento transmitido durante o tempo da execução da tese.

Ao Prof. Doutor Eurico Cabrita e ao grupo de investigação *(Bio)Molecular Structure and Interactions by NMR*, por me fornecerem todas as condições necessárias para a realização deste trabalho e pelos ensinamentos que transmitiram ao longo do ano sobre a técnica de RMN.

À doutoranda Ana Diniz, por me ter guiado e acompanhado durante todo o tempo da tese, sempre com uma grande disponibilidade para me ajudar e ensinar.

À doutora Helena Coelho, por toda a ajuda para a realização da tese, desde à parte prática ao tratamento e discussão dos resultados obtidos.

Ao doutor Aldino Viegas, por todo o conhecimento transmitido e pela sua grande disponibilidade para ajudar em tudo o que fosse necessário para a realização deste trabalho, nomeadamente nas experiências de relaxação.

A todos os membros e colegas do grupo de investigação *(Bio)Molecular Structure and Interactions by NMR*, que me acompanharam ao longo deste ano de trabalho e que tornaram o laboratório num local agradável e de boa disposição.

Ao Prof. Doutor Francisco Corzana da Universidad de La Rioja em Logroño, que realizou e disponibilizou todo o trabalho de dinâmica molecular apresentado na tese.

Quero agradecer o apoio da minha família e mostrar a minha gratidão, pois sem eles não teria sido possível ter esta oportunidade de completar esta etapa da minha vida.

E por último, um grande abraço aos meus amigos que me acompanharam nesta jornada, de onde guardo grandes memórias.

Resumo

A glicosilação aberrante é uma característica típica do cancro e parece conferir vantagens às células cancerígenas, tais como o aumento da sua capacidade de crescimento e da sua habilidade em suprimir a resposta imunitária. A lectina-galactose humana presente em macrófagos (MGL) é uma proteína expressa exclusivamente em células do sistema imunitário que tem grande afinidade para resíduos de *N*-acetilgalactosamina (GalNAc). O resíduo GalNAc é raramente exposto nos glicoconjugados de células humanas, contudo a sua expressão é aumentada em diferentes glicoconjugados presentes em células cancerígenas. A interação entre glicoconjugados, contendo o resíduo GalNAc, expressos no cancro, e a proteína MGL, parece estar envolvida em fenómenos de supressão da resposta imunitária.

Neste contexto, e numa primeira fase, este trabalho consistiu no estudo das interações entre o domínio de reconhecimento de hidratos de carbono da MGL (MGL-CRD) e o antigénio Tn (α -GalNAc/Thr), bem como vários glicopéptidos derivados da glicoproteína da mucina-1 (MUC1) glicosilados com Tn em diferentes locais da sequência peptídica (Tn-T3 MUC1, Tn-T15 MUC1 e Tn-T3, T15 MUC1). Para tal, foram realizadas experiências de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD na ausência e na presença dos diferentes ligandos. Através da análise da perturbação do desvio químico (CSP) verificou-se que a sequência peptídica modula o processo de reconhecimento molecular da MGL. Em particular, verificou-se que a sequência peptídica influi de forma relevante na apresentação do antigénio Tn à proteína MGL-CRD.

Numa segunda parte do trabalho, a dinâmica da cadeia principal do domínio MGL-CRD, na sua forma livre e na presença de distintos ligandos foi analisada tendo por base técnicas de relaxação de ^{15}N HetNOE, R_1 e R_2 . Foram detetadas diferenças consideráveis em R_2 entre a forma livre e complexada da proteína, em alguns aminoácidos das regiões pertencentes aos *loops* 5 e 6 e a folha β 3 do domínio MGL-CRD. Contudo, entre os vários complexos não foram observadas diferenças significativas. A registar apenas uma diminuição do valor de R_2 para os resíduos D270 e C296, em relação ao complexo de referência MGL-CRD/ α -MeGalNAc, aquando da interação do domínio MGL-CRD com os ligandos mais complexos. Por fim, o tempo de correlação (τ_c) determinado para cada um dos complexos mostrou que o MGL-CRD existe como um monómero em solução.

Palavras-chave: MGL-CRD, Glicanos, Interações lectina-glicanos, RMN, Dinâmica molecular, Glicosilação aberrante

Abstract

Aberrant glycosylation is a typical feature of cancer and appears to confer advantages to the cancer cells, such as an increase of their growth capability and their ability to suppress the anti-tumor immune response. The human macrophage galactose-type lectin (MGL), expressed on cells of the immune system, has high affinity for *N*-acetylgalactosamine (GalNAc) residues. The GalNAc moiety is rarely found in glycoconjugates in human cells. However, it is overexpressed in different glycoconjugates present on cancer cells. It has been shown that the interaction between GalNAc-containing glycoconjugates, overexpressed in tumors, and the protein MGL, induces anti-tumor immune suppression.

In this context, in the first part of this work, we investigated the interactions between the carbohydrate recognition domain of MGL (MGL-CRD) and the Tn-antigen (α -GalNAc/Thr). Mucin-1 (MUC1) derived glycopeptides, glycosylated with Tn antigen located at distinct sites of the peptidic sequence (Tn-T3 MUC1, Tn-T15 MUC1 e Tn-T3, T15 MUC1) were studied. For that purpose, ^1H , ^{15}N -HSQC experiments of MGL-CRD in absence and in the presence of the ligands were acquired. Through the analysis of the chemical shift perturbation (CSP), it was demonstrated that the peptide sequence of MUC1 modulates Tn presentation and strongly influences the molecular recognition of the Tn antigen by MGL lectin. In the second part of the work, the molecular dynamics of the MGL-CRD backbone sequence, in free and bound to distinct ligands, were investigated by ^{15}N relaxation techniques, namely HetNOE, R_1 , and R_2 . Specifically, differences in the R_2 values, between MGL-CRD free and complexed state, of certain amino acids located at loops 5, 6 and sheet β 3 were detected. However, no significant variations, amongst the distinct type of complexes, were identified. Only the residues D270 and C296 of MGL-CRD show a decrease in their R_2 value, in the presence of more complex ligands in comparison with the reference MGL-CRD/ α -MeGalNAc complex. Finally, the correlation time (τ_c) for MGL-CRD in free and bound states was determined, and clearly demonstrate that MGL-CRD is monomer in solution.

Keywords: MGL-CRD, Carbohydrates, Lectin-carbohydrates interactions, NMR, Aberrant glycosylation

Índice de matérias

Agradecimentos	V
Resumo	VI
Abstract	VIII
Índice de matérias	X
Índice de figuras	XIV
Abreviaturas	XXIII
1-Introdução	1
1.1-Sistema Imunitário	1
1.2 – Sistema Imunitário Inato	1
1.3 – Sistema Imunitário Adquirido	2
1.4 – Reconhecimento de células pelo sistema imunitário	3
1.5- A função das glicoproteínas no sistema imunitário	3
1.6- Glicosilação, cancro e sistema imunitário	4
1.7- As Lectinas e o seu papel no Sistema Imunitário	6
1.8- Lectina-galactose presente em Macrófagos (MGL)	6
1.8.1- Principais características estruturais da MGL humana	7
1.9- Interações lectina-glicano	9
1.10- O RMN no estudo de complexos proteína-ligando	10
1.10.1- Estudo de um complexo proteína-ligando do ponto de vista da proteína:¹H,¹⁵N-HSQC	11
1.11- Estudo da dinâmica no contexto do reconhecimento molecular	13
1.12- Estudos prévios realizados sobre a MGL	16
3-Métodos e Procedimento Experimental	26

3.1-Expressão e Purificação do domínio MGL-CRD uniformemente marcada em ^{15}N	26
3.1.1-Produção de células de <i>E. coli</i> . competentes.....	26
3.1.2-Transformação de células de <i>E. coli</i> competentes.....	26
3.1.3-Expressão e purificação da MGL-CRD isotopicamente marcada com ^{15}N	27
3.2- Espectroscopia de Ressonância Magnética Nuclear	28
3.2.1- Interação da MGL-CRD com os diferentes glicopéptidos	28
3.2.2- Experiências de dinâmica da MGL-CRD: Estudos de Heteronuclear NOE, R_1 e R_2	29
3.2.3- Análise $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD unido a diferentes ligandos: Determinação da perturbação do desvio químico (CSP).....	30
3.2.4-Análise da dinâmica molecular do MGL-CRD unido a diferentes ligandos: HetNOE, R_1 e R_2	31
3.2.5- Determinação do tempo de correlação a partir da razão R_2/R_1	32
4- Análise e Discussão de Resultados	34
4.1.-Expressão e Purificação da MGL-CRD isotopicamente marcada com ^{15}N	34
4.2-Estudos de interação da MGL-CRD com glicopéptidos cancerígenos por RMN.....	38
4.2.1- MGL-CRD complexado com Tn-Thr.....	39
4.2.2.- MGL-CRD complexado com os monoglicopéptidos MUC1 Tn-T3 e MUC1 Tn-T15.....	44
4.2.3. MGL-CRD complexado com o péptido MUC1 duplamente glicosilado Tn-T3, T15.....	50
4.3.- Estudos de dinâmica molecular do domínio MGL-CRD	53

4.3.1 Dados de relaxação globais e cálculo do tempo de correlação	55
4.3.2 Dados de relaxação MGL-CRD vs MGL-CRD/1.....	58
4.3.3 Dados de relaxação do MGL-CRD/1 vs MGL-CRD/2-5	61
5-Conclusões e perspectivas futuras	66
6-Bibliografia	69
7-Anexos	73
7.1-Sequência Peptídica e vetor de expressão da hMGL-CRD	73
7.2-Soluções utilizadas na expressão da MGL-CRD uniformemente marcada com ¹⁵ N	74
7.4- Dados da MGL-CRD calculados pela ferramenta ExpASy.....	77
7.5- Espectros de ¹ H, ¹⁵ N-HSQC da MGL-CRD na ausência e na presença dos monoglicopéptidos Tn-T3 e Tn-T15 estudados.....	78
7.6- CSPs da MGL-CRD na presença dos monoglicopéptidos Tn-T3 MUC1 e Tn-T15 MUC1.....	79
7.7- Espectro de ¹ H, ¹⁵ N-HSQC do MGL-CRD na presença do diglicopéptido Tn-T3, T15 MUC1	80
7.8- Dados de RMN das experiências de HetNOE da MGL-CRD livre e na presença dos ligandos estudados	81
7.9- Valores de R ₁ obtidos para a MGL-CRD na forma livre e na presença dos ligandos estudados	87
7.10- Valores de R ₂ obtidos para a MGL-CRD na forma livre e na presença dos ligandos estudados	93
7.11- Valores de R ₂ /R ₁ calculados para cada um dos resíduos na presença dos ligandos em estudo.....	99
7.12- Histogramas dos valores de relaxação HetNOE, R ₁ e R ₂ da MGL-CRD na forma livre e na presença dos vários ligandos em estudo	103

Índice de figuras

Figura 1: Diversas estruturas de hidratos de carbono são reconhecidas como antigénios cancerígenos presentes em células tumorais, devido ao fenómeno de glicosilação aberrante das mesmas. Na imagem em cima encontram-se representados vários antigénios cancerígenos, entre eles os antigénios Tn, TF, Lewis e também o antigénio presente em glicolípidos, GM2.....	5
Figura 2: Esquema das principais características estruturais encontradas na MGL humana. Adaptado de Drickamer et al. ²⁹	8
Figura 3: A- Representação do espectro de $^1\text{H},^{15}\text{N}$ -HSQC de uma titulação em regime de equilíbrio lento. Neste tipo de regime, existe a presença de dois sinais correspondentes à forma livre e forma complexada da proteína, cada um representado a população desses dois estados. B- Representação do espectro de $^1\text{H},^{15}\text{N}$ -HSQC de uma titulação em regime de equilíbrio lento. Este tipo de regime é caracterizado pela existência de apenas um sinal correspondente à média ponderada entre as populações da proteína no estado livre e no estado complexado.	13
Figura 4: Escala de tempo representando os tipos de movimentos presentes em proteínas e as respetivas técnicas de RMN normalmente utilizadas para o seu estudo.	14
Figura 5: Estrutura química dos ligandos reconhecidos pela MGL anteriormente estudados e as suas respetivas representações esquemáticas.	16
Figura 6: Espectro de $^1\text{H},^{15}\text{N}$ -HSQC do CRD da MGL na ausência (preto) e na presença de α -MeGalNAc (vermelho). Os sinais indicados com um * representam os resíduos que apareceram na presença de α -MeGalNAc e os que não contêm um * indicam aqueles que foram identificados no estudo de CSP. Adaptado de Diniz et al. ⁴⁴	17
Figura 7: Histograma do CSP do CRD da MGL na presença de α -MeGalNAc. A vermelho encontram-se assinalados os resíduos que foram perturbados pelo ligando neste estudo. As duas linhas cinzentas a tracejado representam os valores de $\Delta\delta_{\text{comb}}$ delineados que consideram quais os resíduos que são perturbados em maior extensão.	

A primeira linha foi marcada a um valor de 0,08 ppm e a segunda a 0,15 ppm. Adaptado de Diniz et al. ⁴⁵	18
Figura 8: Mapa da interação da MGL-CRD com o ligando α -MeGalNAc. Os resíduos identificados pela análise do CSP encontram-se assinalados a vermelho, assim como as regiões que os contêm. Adaptado de Diniz et al. ⁴⁵	18
Figura 9: Histograma do CSP do MGL-CRD na presença dos vários ligandos estudados usando como referência o espetro de $^1\text{H},^{15}\text{N}$ -HSQC de α -MeGalNAc. A vermelho encontram-se assinalados os resíduos que foram perturbados pelo ligando neste estudo. Adaptado de Diniz et al. ⁴⁵	20
Figura 10: Experiência de STD do glicopéptido Tn-T8 MUC1. Os valores de % de STD encontram-se representados na estrutura por círculos com a cor correspondente à sua % de STD recebida. Adaptado de Marcelo et al. ⁴⁴	21
Figura 11: Cromatograma da cromatografia de afinidade. O pico correspondente à eluição da MGL-CRD encontra-se assinalado a vermelho. O tampão de eluição é composto por Tris-HCl 10 mM, NaCl 75 mM, CaCl_2 20 mM e 150 mM de lactose a pH 7,5.	35
Figura 12: Gel de eletroforese (SDS-PAGE) a 10% da fração de MGL-CRD. No gel observa-se uma banda correspondente ao peso molecular do domínio MGL-CRD e também uma banda correspondente ao peso molecular da formação de dímeros do domínio MGL-CRD. Os marcadores de pesos moleculares NZYColour Protein Marker II encontram-se no poço identificado com a letra M.	36
Figura 13: Estruturas químicas do ligando Tn-Thr e dos glicopéptidos da MUC1 contendo o açúcar GalNAc estudados.	39
Figura 14: A- Sobreposição dos espetros de $^1\text{H},^{15}\text{N}$ -HSQC do CRD MGL na forma livre (azul) e na presença do ligando Tn-Thr (vermelho) adquiridos a 600MHz e 293K. Os espetros foram adquiridos com uma concentração de MGL-CRD de 90 μM e de 180 μM de antigénio numa proporção de 1:2. A negrito encontram-se assinalados os resíduos que surgiram na presença do ligando e também os resíduos com uma perturbação de $\Delta\delta_{\text{comb}}$ significativo. B- Histograma com a análise de CSP do ligado Tn-Thr. A linha a laranja	

e a linha a cinzento representam os valores de $\Delta\delta_{\text{comb}}$ considerados para os resíduos fortemente e moderadamente perturbados, respetivamente..... 40

Figura 15: Histograma da análise de CSP do ligando Tn-Thr utilizando como referência o ligando α -MeGalNAc. Os resíduos identificados encontram-se assinalados na figura.. 41

Figura 16: Sobreposição dos espectros de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD na presença do ligando α -MeGalNAc (azul) e do ligando Tn-Thr (vermelho) dos resíduos identificados na análise de CSP..... 42

Figura 17: Modelo 3D de homologia do complexo MGL-CRD/Tn-Thr gerado por dinâmica molecular. Os resíduos a azul (H274, D294 e C295) foram identificados pela análise de CSP entre os complexos MGL-CRD/Tn-Thr e MGL-CRD/ α -MeGalNAc. A verde encontra-se representada a estrutura do ligando Tn- Thr..... 43

Figura 18: Representação esquemática dos dois monoglicopéptidos Tn-T3 e Tn-T15. A amarelo encontra-se representado o açúcar GalNAc. 44

Figura 19: A- Histograma da análise de CSP do ligando Tn-T3 utilizando como referência o ligando Tn-Thr. **B-** Histograma da análise de CSP do ligando Tn-T15 utilizando como referência o ligando Tn-Thr. 45

Figura 20: Ampliação de alguns resíduos que apresentaram diferenças no desvio químico da MGL-CRD na presença do ligando Tn-T3 em relação ao outros dois ligandos. 46

Figura 21: Ampliação de alguns resíduos que apresentaram diferenças no desvio químico da MGL-CRD na presença do ligando Tn-T15 em relação ao outros dois ligandos. 47

Figura 22: Ampliação de alguns resíduos que apresentaram diferenças no desvio químico da MGL-CRD tanto na presença do ligando Tn-T3 bem como na presença do ligando Tn-T15 em relação ao ligando Tn-Thr..... 48

Figura 23: A- Modelo 3D do complexo MGL-CRD/ Tn-T3 MUC1 gerado por dinâmica molecular. **B-** Modelo 3D do complexo MGL-CRD/Tn-T15 gerado por dinâmica molecular. As zonas assinaladas a azul mostram os resíduos identificados que mostram diferenças na MGL-CRD na presença do ligando correspondente em relação à MGL-CRD

na presença do ligando Tn-Thr. As zonas a vermelho identificam os resíduos duplamente perturbadas em ambos os glicopeptidos em relação ao ligando Tn-Thr..... 49

Figura 24: A e B- Modo de ligação previsto pela MGL-CRD em que uma molécula de MGL-CRD se liga preferencialmente a um dos dois sítios glicosilados com o açúcar GalNAc no glicopéptido Tn-T3, T15 MUC1. **C-** Modo de ligação previsto em que duas moléculas de MGL-CRD interagem com os dois sítios de glicosilação presente em uma molécula de glicopéptido Tn-T3, T15 MUC1..... 50

Figura 25: Ampliação de alguns resíduos que apresentam dois sinais correspondentes a duas conformações diferentes no espectro de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD na presença do ligando Tn-T3, T15 MUC1. Os sinais correspondentes a cada uma das conformações encontram-se assinalados na figura..... 51

Figura 26: Histograma das intensidades dos resíduos que apresentaram dois sinais bem definidos para cada uma das populações existentes no espectro de $^1\text{H},^{15}\text{N}$ -HSQC do glicopéptido Tn-T3, T15 MUC1..... 52

Figura 27: A- Modelo 3D de homologia do domínio MGL-CRD. A vermelho encontram-se assinalados os aminoácidos os quais viram os seus sinais aparecer no espectro de $^1\text{H},^{15}\text{N}$ -HSQC na presença dos vários ligandos estudados. **B-** Histograma da diferença das intensidades dos resíduos entre o complexo MGL-CRD/ α -MeGalNAc e o domínio MGL-CRD na forma livre. A vermelho estão assinalados as intensidades correspondentes aos resíduos que surgiram no espectro de $^1\text{H},^{15}\text{N}$ -HSQC na forma complexada. **C-** Análise das flutuações do $\text{C}\alpha$ da cadeia principal por dinâmica molecular do domínio MGL-CRD na forma livre e na presença do complexo MGL-CRD/ α -MeGalNAc. Adaptado de Diniz et al.⁴⁵..... 54

Figura 28: Estruturas químicas dos ligandos e das suas respetivas representações esquemáticas que foram utilizados no estudos da dinâmica do domínio MGL-CRD. ... 55

Figura 29: A e B- Valores de HetNOE e de R_2 , respetivamente, para cada um dos resíduos do domínio MGL-CRD (preto) e do complexo MGL-CRD/1 (vermelho) Encontra-se representado também o histograma da análise do CSP do complexo MGL-CRD/1 em relação ao domínio MGL-CRD. **C-** Histograma da diferença das intensidades dos resíduos

entre o complexo MGL-CRD/ α -MeGalNAc e o domínio MGL-CRD na forma livre. A vermelho estão assinalados as intensidades correspondentes aos resíduos que surgiram no espectro de $^1\text{H},^{15}\text{N}$ -HSQC na forma complexada. 60

Figura 30: Histogramas da perturbação do desvio químico de diferentes ligandos utilizando como referência o complexo MGL-CRD/1: **A**-MGL-CRD/2 **B**-MGL-CRD/3 **C**-MGL-CRD/4 **D**-MGL-CRD/5. Encontram-se representados no eixo secundário os valores de HetNOE para cada resíduo para cada um dos respetivos complexos a vermelho e os valores de HetNOE para cada resíduo do complexo MGL-CRD/1 a preto. 62

Figura 31: Histogramas da perturbação do desvio químico de diferentes ligandos utilizando como referência o complexo MGL-CRD/1: **A**-MGL-CRD/2 **B**-MGL-CRD/3 **C**-MGL-CRD/4 **D**-MGL-CRD/5. Encontram-se representados no eixo secundário os valores de R_2 para cada resíduo para cada um dos respetivos complexos a vermelho e os valores de R_2 para cada resíduo do complexo MGL-CRD/1 a preto. 63

Figura 32: Análise das flutuações do $C\alpha$ da cadeia principal por dinâmica molecular do complexo MGL-CRD/1 e na presença dos complexos MGL-CRD/2-5. Adaptado de Diniz et al.⁴⁵ 64

Figura 33: Vetor de Expressão pET21a contendo o gene do domínio MGL-CDR, contem 5443 pares de bases e resistência a ampicilina. 73

Figura 34: **A**- Sobreposição dos espectros de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD (azul) e da MGL-CRD na presença do ligando Tn-T3. **B**- Sobreposição dos espectros de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD (azul) e da MGL-CRD na presença do ligando Tn-T15. Os resíduos assinalados em ambos os espectros foram identificados a partir da análise de CSP. Os espectros foram adquiridos a 293K a pH 7,4. 78

Figura 35: **A**- Histograma com a análise de CSP da MGL-CRD na presença do ligando Tn-T3 MUC1. **B**- Histograma com a análise de CSP da MGL-CRD na presença do ligando Tn-T15 MUC1. A linha a laranja e a linha a cinzento representam os valores de $\Delta\delta_{\text{comb}}$ considerados para os resíduos fortemente e moderadamente perturbados, respetivamente. Os resíduos identificados encontram-se assinalados na figura. 79

Figura 36: Espectro de $^1\text{H},^{15}\text{N}$ -HSQC do domínio MGL-CRD na presença do ligando Tn-T3, T15. Os resíduos assinalados correspondem aos resíduos identificados que apresentam duas ressonâncias correspondentes a cada um dos modos de união do domínio MGL-CRD existente em solução. O espectro foi adquirido a 293K a pH 7,4. 80

Figura 37: **A-** Histograma dos valores de HetNOE para cada um dos resíduos para o domínio MGL-CRD. **B-** Histograma dos valores de R_1 para cada um dos resíduos para o domínio MGL-CRD. **C-** Histograma dos valores de R_2 para cada um dos resíduos para o domínio MGL-CRD. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD. 103

Figura 38: **A-** Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/1. **B-** Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/1. **C-** Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/1. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD. 104

Figura 39: **A-** Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/2. **B-** Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/2. **C-** Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/2. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD. 105

Figura 40: **A-** Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/3. **B-** Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/3. **C-** Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/3. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD. 106

Figura 41: **A-** Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/4. **B-** Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/4. **C-** Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/4. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD. 107

Figura 42: **A-** Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/5. **B-** Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/5. **C-** Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/5. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD..... 108

Índice de tabelas

Tabela 1: Média dos valores de R_1 , R_2 e HetNOE e valor de τ_c estimado do domínio de MGL-CRD obtido nas diferentes condições. Entre parêntesis estão o número total de resíduos usados para o cálculo de cada um dos parâmetros de relaxação e τ_c	57
Tabela 2- Composição do meio LB.....	74
Tabela 3- Composição do meio LB Agar.	74
Tabela 4- Composição do meio LB Agar. Os valores indicados são para a preparação de 1L.....	75
Tabela 5- Preparação de 5ml da solução de gel de empacotamento para SDS-PAGE.	75
Tabela 6- Preparação de 15ml da solução de gel de separação para SDS-PAGE.	76
Tabela 7- Preparação de 50mL de solução de sample buffer para SDS-PAGE.....	76
Tabela 8- Preparação de 1L da solução tampão 10x Ânodo 2M Tris	76
Tabela 9- Preparação de 1L da solução tampão Tris-Tricina.....	76
Tabela 10- Valores obtidos da experiência de relaxação HetNOE e respectivos erros do domínio MGL-CRD na forma livre e na presença dos complexos MGL-CRD/1-5.....	81
Tabela 11- Valores obtidos da experiência de relaxação R_1 e respectivos erros do domínio MGL-CRD na forma livre e na presença dos complexos MGL-CRD/1-5.....	87
Tabela 12- Valores obtidos da experiência de relaxação R_2 e respectivos erros do domínio MGL-CRD na forma livre e na presença dos complexos MGL-CRD/1-5.....	93
Tabela 13- Valores calculados do rácio R_2/R_1 para cada resíduo do domínio MGL-CRD na forma livre e na presença dos complexos MGL-CRD/1-5.....	99

Abreviaturas

ϵ - Coeficiente de extinção molar

$\Delta\delta_{comb}$ – Desvio químico combinado

δ – Desvio químico

σ - Desvio padrão

τ_c – Tempo de correlação

Ab_{280nm} – Absorvância medida a 280 nm

Antigénio TF – Antigénio Thomsen-Friedenreich

APS – Persulfato de Amónia (do inglês *Ammonium Persulfate*)

Células NK - do inglês *Natural Killer cells*

CRD – Domínio de reconhecimento de hidratos de carbono (do inglês *Carbohydrate recognition domain*)

CSP – Perturbação do desvio químico (do inglês *Chemical shift perturbation*)

CTL - Lectina do tipo C (do inglês *C type lectin*)

Da - Dalton

DC-SIGN - proteína de adesão intercelular não-integrina das células dendríticas (do inglês *Dendritic cell-specific intercelular adhesion molecule-3-grabbing non-integrin*)

DEC205 - recetor da célula dendríticas para endocitose (do inglês *Dendritic cell receptor for endocytosis*)

DNA – Ácido Desoxirribonucleico (do inglês *Deoxyribonucleic acid*)

DOSY – do inglês *diffusion ordered spectroscopy*

DTT – Ditioneitol

E. coli – *Escherichia Coli*

EDTA – Ácido etilendiamino tetra-acético (do inglês *Ethylenediamine tetraacetic acid*)

Fuc – Fucose

Gal - Galactose

GalNAc – *N*- Acetilgalactosamina

Glc – Glucose

GlcNAc – *N*-Acetilglucosamina

HSQC – do inglês *Heteronuclear Single Quantum Coherence*

IPTG – Isopropil- β -D-1-tiogalactopiranosídeo

J – Constante de acoplamento escalar

LPS – Lipopolissacáridos

ManNAc – *N*-Acetilmanosamina

Meio LB – Meio Luria-Bertani

MGL – Lectina galactose presente em macrófagos (do inglês *Macrophage galactose type lectin*)

MHC - Complexo principal de histocompatibilidade (do inglês *Major histocompatibility complex*)

MPB - Proteína de união à manose (do inglês *Mannose binding protein*)

MUC1 – Mucina 1

Neu5Ac – Ácido *N*-Acetilneuramínico

NOE – Efeito Nuclear de Overhauser (do inglês *nuclear Overhauser effect*)

OD_{600nm} – Densidade ótica a 600 nm

PAMPs - Padrões moleculares associados a antígenos (do inglês *Pathogen-associated molecular pattern*)

PDB – do inglês *Protein Data Bank*

ppm – partes por milhão

PRRs - Padrões de recetores de reconhecimento (do inglês *Pattern pathogen- associated molecular pattern*)

R₁ – Taxa de relaxação longitudinal

R₂ – Taxa de relaxação transversal

RMN – Ressonância Magnética Nuclear

rms – Raiz do valor quadrático Médio

rpm – Rotações por minuto

SDS-PAGE – eletroforese em gel de poliacrilamida na presença de dodecil sulfato de sódio (do inglês *sodium dodecyl sulfate – polyacrilamide gel electrophoresis*)

Ser- Serina

Siglecs – lectinas do tipo imunoglobulina que ligam a ácidos siálicos (do inglês *Sialic acid-binding immunoglobulin-type lectins*)

STD – do inglês *Saturation transfer difference*

Thr – Treonina (do inglês *Threonine*)

Tris – Tris (Hidroximetil) aminometano

TrNOE – do inglês *Transferred nuclear Overhauser effect*

TSP – Ácido Trimetilsilil propiónico

UV-VIS- Ultravioleta visível

1-Introdução

1.1-Sistema Imunitário

O sistema imunitário, sistema imunológico ou sistema imune de um organismo desempenha como principal papel a defesa do mesmo contra uma grande variedade de agressores a que está sujeito no seu dia-a-dia, desde agentes patogénicos, como vírus e bactérias, a alergénicos, ou até na proteção contra agressores endógenos, como, por exemplo, células não “saudáveis” ou células infetadas. A capacidade do sistema imunitário em desempenhar o seu papel depende da sua eficiência em conseguir distinguir entre células intrínsecas e saudáveis do próprio corpo de células externas (invasoras, agressoras) ou próprias, mas não saudáveis. Para isso, conta com a existência de células específicas, denominadas imunitárias, e de um complexo sistema de processos biológicos capazes de reconhecer e discriminar estas duas identidades.¹ O sistema imunitário contempla dois tipos de mecanismos de resposta: inata e adquirida.

1.2 – Sistema Imunitário Inato

Na imunidade inata, estão compreendidas as barreiras físicas e químicas do indivíduo que constituem a primeira linha de defesa contra os agentes agressores, como a pele, as barreiras mucosas ou então o suco gástrico do estômago. Se as células invasoras conseguirem escapar esta primeira barreira, entram em atuação as células do sistema imunitário, como os neutrófilos e os macrófagos, que circulam livremente na corrente sanguínea prontos para intervirem na resposta imediata do organismo, conhecida como resposta humoral².

Este tipo de imunidade é caracterizado por ser um tipo de resposta não específica perante a forma do agressor que é apresentado ao organismo, dependendo, para isso, de um inúmero tipo de padrões moleculares associados a agentes patogénicos (do inglês *Pathogen-associated molecular pattern*, PAMPs) conservados em microrganismos que são reconhecidos por recetores presentes nas células imunitárias (do inglês *Pattern*

recognition receptor, PRRs). Alguns exemplos destes padrões moleculares incluem os lipopolissacáridos (LPS), presentes nas membranas de bactérias gram-negativas, ou a estrutura do peptidoglicano, presente na parede celular em quase todas as bactérias³.

1.3 – Sistema Imunitário Adquirido

Ao contrário do sistema imunitário inato, o sistema imunitário adquirido comporta-se de uma maneira específica/seletiva, sendo por isso uma resposta mais lenta e mais avançada evolutivamente, estando presente apenas em alguns animais vertebrados. Os principais intervenientes neste tipo de resposta são os linfócitos B e T e as células NK (do inglês *Natural killer cells*).⁴

A sua especificidade deve-se ao grande número de recetores de antígenos existentes nos linfócitos B e T, que são adquiridos durante o seu desenvolvimento na medula óssea e timo, respetivamente. Esta grande abundância de recetores é alcançada a partir de diferentes rearranjos entre uma linhagem de genes que, nos humanos, permitem obter uma diversidade de cerca de 10^8 de recetores diferentes, assegurando assim uma proteção a uma vasta gama de microrganismos e de antígenos existentes em todas as formas de vida conhecidas⁴.

Em primeiro lugar, o antígeno é reconhecido pelo recetor específico num linfócito B ou T levando à ativação desse linfócito, induzindo a resposta imunitária e a diferenciação e proliferação, no caso dos linfócitos T, ou a indução da libertação de anticorpos específicos pelos linfócitos B. Após desenvolver uma resposta ao encontrar um antígeno pela primeira vez, o sistema imunitário adquirido guarda essa informação e perante uma nova exposição a esse mesmo antígeno, provoca uma resposta mais forte e mais rápida^{2,5}.

Apesar de serem caracterizados individualmente, estes dois tipos de imunidade trabalham em sinergia na defesa do organismo, sendo as suas funções potenciadas aquando do normal funcionamento destes, por exemplo, na apresentação de antígenos ao sistema adquirido pelas células da imunidade inata.

1.4 – Reconhecimento de células pelo sistema imunitário

O sistema imunitário dispõe de diversas estratégias para controlar e identificar as suas próprias células de outras pertencentes a diferentes organismos e também a capacidade de detetar as células do seu corpo que sofreram alterações. A estratégia utilizada na imunidade inata é a capacidade de expressar uma grande quantidade de recetores não específicos nas superfícies das suas células imunitárias que irão reconhecer diversos motivos evolutivamente conservados, dependendo da classe do organismo invasor, conhecidos como PAMPs⁶. Acredita-se que em vertebrados complexos, como os seres humanos, existam várias centenas de PRRs. Este tipo de discriminação é pouco específica, permitindo assim, na primeira fase de defesa, um grande espectro de atuação perante um diverso número de organismos de diferentes origens. Na classe de PRRs encontram-se diversas famílias de proteínas, entre elas as lectinas, uma família que tem como principal característica o reconhecimento de resíduos de açúcares⁷.

Em contraste com o sistema inato, o sistema adquirido dispõe de uma outra estratégia, que consiste na presença de um grupo de proteínas membranares presentes na superfície das células do próprio organismo, conhecido como o complexo principal de histocompatibilidade ou MHC (do inglês *Major histocompatibility complex*). Quando existe o contato entre uma célula própria saudável e uma célula do sistema imunitário adquirido, a célula imunitária irá reconhecer e ligar-se ao MHC, sinalizando à mesma que a célula pertence ao próprio organismo. Na presença de antigénios invasores, este complexo é alterado e apresenta os epítomos correspondentes à célula invasora, fazendo com que a interação entre este complexo e as células do sistema imunitário adquirido seja perturbada, levando à ativação das mesmas^{4,8}.

1.5- A função das glicoproteínas no sistema imunitário

Os glicanos (açúcares, hidratos de carbono) são as biomoléculas mais abundantes da natureza e apresentam uma enorme diversidade estrutural⁹. Existem associados a proteínas e lípidos por ligações covalentes na superfície das membranas celulares. Devido a estas características, desempenham diversas funções, regulando variados mecanismos celulares, como o reconhecimento, a sinalização e a comunicação celular,

a ativação de receptores celulares e de transdução de sinal, e participando em processos de endocitose pelas células^{9,10,11}.

A glicosilação é o mecanismo pós-traducional mais importante nas células eucarióticas e inicia-se no retículo endoplasmático, ocorrendo maioritariamente no complexo de Golgi¹². Os padrões de glicosilação de uma célula são modulados por duas famílias de enzimas, as glicosiltransferases e as glicosidases, dependendo do tipo de célula em que se encontram.

Os mamíferos desenvolveram, ao longo de milhares de anos, um grande número de estruturas de glicanos, evolutivamente conservados, diferenciando-se das estruturas de glicanos presentes em organismos menos evoluídos. Desta forma, os glicanos são uma das principais estruturas utilizadas pelo sistema imunitário para discriminar células próprias e saudáveis de células invasoras ou próprias, mas não saudáveis.

Diversas proteínas, existentes à superfície das células imunitárias, são responsáveis pelo reconhecimento de glicanos. Como tal, alterações na estrutura dos glicanos serão detetadas induzindo, ou diminuindo, a resposta imunitária das células imunitárias⁷.

1.6- Glicosilação, cancro e sistema imunitário

O padrão de glicosilação existente à superfície das células (glicoproteínas, glicolípidos) é um processo extraordinariamente regulado, sendo controlado pela ação concertada de um conjunto de enzimas (glicosiltransferases e glicosidases)⁹.

Durante o desenvolvimento de uma célula tumoral, existem diversas mutações genéticas na mesma, sendo que, em particular, os níveis de glicosidases e de glicosiltransferases são extremamente alterados, levando a que padrões de glicosilação aberrantes passem a ser expressos nas células. Efetivamente, a glicosilação aberrante é uma das características das células tumorais¹³. De entre as estruturas de glicanos, reconhecidos como antígenos mais expressos nas células cancerígenas, encontram-se diversos tipos de *O*-glicanos, tais como, o antígeno Tn (resíduo de *N*-acetilgalactosamina (GalNAc) ligado a uma serina ou treonina), o antígeno Thomsen-Friedenreich (TF, dissacárido Gal β 1-3GalNAc) e os antígenos de Lewis (Gal β 1-3GalNAc α 1-4Fuc), e as suas

respetivas estruturas sialiladas, os quais se encontram ligados a proteínas, como as mucinas^{14,15}. A sobre expressão de determinados glicolípido também é uma característica do cancro¹⁶ (Figura 1).

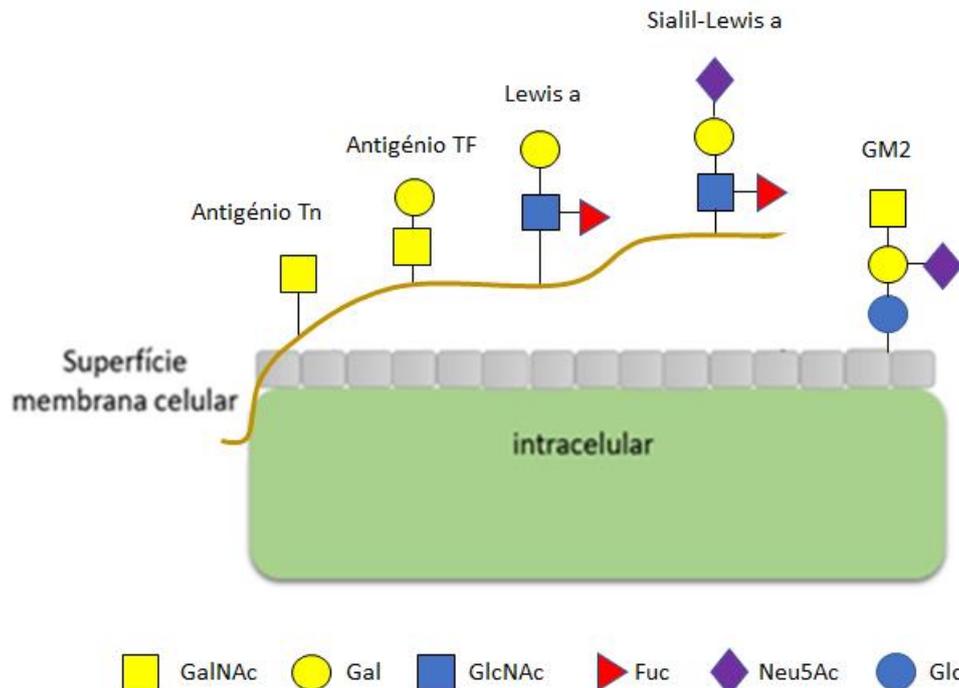


Figura 1: Estruturas de hidratos de carbono reconhecidas como antígenos presentes em células tumorais. Entre os antígenos apresentados estão os antígenos: Tn, TF, Lewis a e correspondente versão sialilada, bem como o antígeno presente em glicolípido, GM2.

O sistema imunitário defende o organismo do aparecimento de células geneticamente alteradas, conseguindo, numa fase prematura, eliminar as células tumorais, nomeadamente através das células NK, que, ao encontrarem uma célula tumoral, detetam uma diminuição na expressão de proteínas do MHC, característica da diferenciação de uma célula saudável para uma célula tumoral⁸.

As alterações no padrão de glicosilação das células tumorais conferem a estas células certas vantagens do ponto de vista evolutivo, tais como, uma grande capacidade de proliferação e de evasão do sistema imunitário^{17,18}. Neste sentido, uma das principais estratégias utilizadas é a expressão abundante de estruturas de glicanos muito semelhantes aos naturais, que, ao serem detetados por lectinas presentes na superfície das células do sistema imunitário, conduz, em muitos casos, à supressão da resposta imunitária^{13,19}.

1.7- As Lectinas e o seu papel no Sistema Imunitário

As lectinas são uma família de proteínas que são reconhecidas pela sua habilidade em identificar e interagir com resíduos de açúcares. As lectinas estão divididas em diferentes classes consoante o tipo de açúcar que interagem²⁰.

Algumas lectinas são expressas em grande quantidade nas células do sistema imunitário participando, maioritariamente, em processos de reconhecimento de antígenos, sendo um dos tipos de PRRs mais abundantes na resposta inata de um organismo³. As suas funções no sistema imunitário incluem a participação em processos de reconhecimento hospedeiro-agente patogénico e na interface entre o sistema inato e adquirido, em mecanismos de apresentação de antígenos ao sistema imunitário adquirido²¹.

Através da interação com os glicanos, as lectinas expressas em células do sistema imunitário, podem estar envolvidas no aumento ou na supressão da resposta imunitária, ajudando a manter os níveis basais de ativação do sistema imunitário perante as células do seu próprio organismo e aumentando a resposta aquando da interação com estruturas de glicanos anormais ao próprio corpo.

1.8- Lectina-galactose presente em Macrófagos (MGL)

A maior e mais diversificada família de lectinas presentes em mamíferos e expressas em células do sistema imune, são as lectinas do tipo C, (CTL, do inglês *C-type lectins*). Esta classe inclui, entre outras, a lectina-galactose presente em macrófagos (MGL, do inglês *Macrophage galactose type lectin*), a langerina, o recetor da célula dendrítica para endocitose (DEC205, do inglês *Dendritic cell receptor for endocytosis*), a proteína de união à manose (MPB, do inglês *Mannose binding protein*) e a proteína de adesão intercelular não-integrina das células dendríticas (DC-SIGN, do inglês *Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin*)^{20,22}. São lectinas caracterizadas por conterem um domínio de reconhecimento de hidratos de carbono (CRD, do inglês *carbohydrate recognition domain*) que interage com o seu ligando de uma maneira dependente do ião Ca^{2+} . Devido às suas funções na adesão celular, reconhecimento de antígenos e sinalização celular, as CTLs conseguem regular e

modelar o efeito da resposta imunitária, conseguindo diferenciar entre glicanos endógenos e glicanos de células invasoras ou células tumorais.

Em particular, a MGL humana é expressa exclusivamente em macrófagos e em células dendríticas não maduras, diminuindo o seu nível de expressão após maturação da célula dendrítica^{23,24}. Entre o tipo das lectinas do tipo C, é a única que demonstra afinidade para resíduos terminais de GalNAc, como, por exemplo, para o antígeno Tn (GalNAc-Ser/Thr), presente em células tumorais, e em glicoproteínas das células T (CD43 e CD45), ou para o epítipo LacdiNAc (GalNAc β 1-4GlcNAc), o qual é comum em agentes patogénicos²³.

O reconhecimento de determinados ligandos associados a tumores por parte da MGL parece desencadear a supressão do sistema imunitário. A interação da MGL com o antígeno Tn, presente nas células tumorais e na proteína CD45 das células T ativadas, vai reduzir a proliferação das células T e a produção de citocinas inflamatórias, o que conduz a uma diminuição da resposta imunitária²³. Além disso, existem diversos ligandos que contêm o resíduo de GalNAc exposto, tais como gangliosídios GM2/GD2, antígeno Forssman e antígeno do grupo sanguíneo do tipo A, os quais interagem com a MGL e parecem estar relacionados com a indução da supressão da resposta imune pela mesma.^{25,26}

Foi também documentado que a interação da MGL com células contendo estruturas com resíduos de GalNAc poderá estar relacionada com a diminuição da migração das células dendríticas até aos nódulos linfáticos, onde a resposta imunitária adquirida é despoletada, atrasando e abrandando assim a resposta imune²⁷.

1.8.1- Principais características estruturais da MGL humana

A MGL é uma proteína transmembranar do tipo II que pertence a família das lectinas do tipo C²². A sua estrutura é formada por um domínio citoplasmático, uma região transmembranar, um *neck domain* e um CRD^{24,28} (Figura 2).

O CRD contém uma zona de ligação ao ião Ca²⁺ que consiste numa sequência de 3 aminoácidos, evolutivamente conservados, os quais são fundamentais e indispensáveis

para a interação com o ligando, visto que esta interação acontece através do átomo de Ca^{2+} , o qual se coordena depois ao resíduo de açúcar.

As CTLs dividem-se em dois subgrupos com base no tipo de ligação que existe entre o átomo de Ca^{2+} e o resíduo de açúcar. As lectinas que contêm no seu CRD a sequência conservada EPN têm uma preferência para resíduos de manose, *N*-acetilglucosamina (GlcNAc) e glucose, que contêm os grupos hidroxilo OH-3 e hidroxilo OH-4, ambos em posição equatorial. No caso da MGL, o CRD contém a sequência QPD, a qual lhe confere especificidade para ligandos com os grupos hidroxilo OH-3 em posição equatorial e OH-4 em posição axial, ou seja, para os resíduos de GalNAc e galactose (Gal)^{29,30}.

O domínio intracelular da MGL é caracterizado por conter o motivo de tirosina YENF, uma sequência tipicamente encontrada em recetores endocíticos, que é requerido para a interação com vesículas revestidas de clatrina³¹. Já a região do *neck domain*, consiste num motivo formado por repetições de sete aminoácidos que contêm resíduos hidrofóbicos em sítios específicos da sequência³². Resultados da literatura parecem demonstrar que o *neck domain* está envolvido na capacidade do domínio extracelular da MGL em oligomerizar e formar trímeros entre si. A oligomerização no caso das CTLs é um fenómeno conhecido, o qual se encontra bem descrito e documentado no caso das lectinas MBP e a DC-SIGN²⁴.

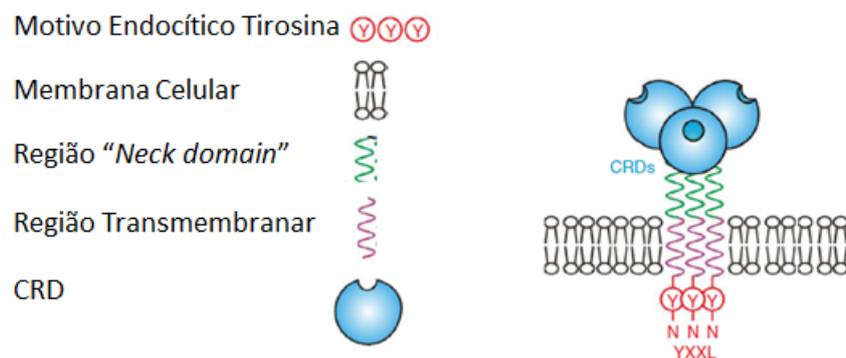


Figura 2: Esquema dos componentes estruturais presentes na MGL humana. Adaptado de Drickamer et al.²⁹

1.9- Interações lectina-glicano

O estudo a nível atómico dos complexos proteína-glicano é crucial para se compreender a especificidade dos eventos de reconhecimento molecular envolvendo açúcares, perceber a sua repercussão a nível biológico, bem como desenvolver novas estratégias onde seja possível modular as interações proteína-glicano (fármacos, vacinas, entre outros). Do ponto de vista estrutural, o reconhecimento molecular entre um glicano e uma lectina é composto por diferentes tipos de interações intermoleculares.

A estrutura molecular dos glicanos é composta por um elevado número de grupos hidroxilo, o que leva a que o tipo de interação mais observada num complexo proteína-glicano seja a interação por pontes de hidrogénio. Os grupos hidroxilos dos açúcares podem funcionar tanto como um grupo dador como um grupo aceitador de pontes de hidrogénio. Para além dos grupos hidroxilos presentes em todos os tipos de unidades de açúcar, também podem existir interações do tipo ponte de hidrogénio com grupos carboxilo ou grupos amina em alguns tipos de açúcares, tais como os resíduos de GalNAc, GlcNAc ou *N*-acetilmanosamina (ManNAc). Do lado da lectina a ponte de hidrogénio normalmente envolve os grupos polares das cadeias laterais de aminoácidos, como, por exemplo, argininas, glutaminas, ácidos glutâmicos, asparaginas, bem como com a cadeia principal de aminoácidos através dos grupos NH e C=O, da ligação peptídica³³.

Outro tipo de interação documentada entre estes tipos de complexos são as interações CH- π existentes entre as ligações C-H do açúcar e os sistemas aromáticos de certos aminoácidos, tais como, tirosina, fenilalanina e triptofano, normalmente muito frequentes nos locais de ligação das lectinas. Este tipo de ligação não covalente ocorre pois o sistema π do anel aromático é rico em eletrões e, como tal, pode interagir electrostaticamente com zonas hidrofóbicas criadas pela existência de várias ligações C-H com a conformação espacial correta e existentes em todos os açúcares³³.

Mais pontualmente, existem também interações por pontes salinas. Glicanos que possuam resíduos de ácidos siálicos são reconhecidos por uma vasta família de lectinas, tais como as lectinas do tipo imunoglobulina que ligam a ácidos siálicos (Siglecs, do inglês *Sialic acid-binding immunoglobulin-type lectins*), através da formação de uma

ponte salina entre o carboxilo do ácido siálico e uma arginina, presente no sítio de ligação e conservada em todas Siglecs³⁴.

Finalmente, é também bem conhecida a participação de íões divalentes no processo de reconhecimento de glicanos por diversas lectinas, principalmente na família das CTLs, em que o íão divalente é o Ca^{2+} ²⁰. Uma vez mais, a interação com o íão é realizada através dos grupos hidroxilo do açúcar. A orientação dos grupos hidroxilo no açúcar é essencial para a correta coordenação com o íão²⁹. Por outro lado, a coordenação do íão à lectina envolve a participação de aminoácidos carregados negativamente, tais como, aspartatos e glutamatos.

1.10- O RMN no estudo de complexos proteína-ligando

Em conjunto com a técnica de Cristalografia por Raios-X, a técnica de Ressonância Magnética Nuclear (RMN) é uma das técnicas mais utilizadas no estudo de complexos proteína-ligando. A técnica de RMN permite obter informação a nível atômico acerca da estrutura, da dinâmica e das interações estabelecidas entre uma proteína e um ligando, em solução (condições mais próximas das existentes fisiologicamente)³⁵. Em particular, no caso dos complexos proteína-glicano, a flexibilidade dos glicanos, sobretudo dos glicanos mais complexos, faz com que seja difícil de obter informação estrutural por Cristalografia de Raios-X acerca do glicano (devido à baixa resolução). Desta forma, a técnica de RMN torna-se mais versátil e robusta no que respeita à elucidação de complexos proteína-glicano³⁶.

A técnica de RMN permite a determinação de diversos parâmetros, tais como, o desvio químico (δ), a taxa de relaxação (R_1 e R_2), difusão, efeito nuclear de Overhauser (NOE), constantes de acoplamento (J), os quais permitem retirar informações acerca da estrutura de uma (bio)molécula. Além disso, qualquer um destes parâmetros é extremamente sensível aquando a formação de um complexo proteína-ligando, adquirindo valores muito diferentes do estado livre para o estado associado³⁷.

Existem diversas técnicas de RMN utilizadas para estudar interações proteína-ligando, sendo possível dividi-las entre as técnicas que detetam alterações estruturais na

proteína e as técnicas que observam alterações na estrutura do ligando. No estudo de interações proteína-ligando em que se observa o reconhecimento molecular do ponto de vista da proteína são utilizadas experiências que têm normalmente como base a realização de espectros 2D $^1\text{H},^{15}\text{N}$ -HSQC (do inglês *Heteronuclear Single Quantum Coherence*). Para a realização deste tipo experiências é absolutamente vital a marcação isotópica da proteína. Através da experiência de 2D $^1\text{H},^{15}\text{N}$ -HSQC é possível: 1) calcular a perturbação de desvio químico (CSP, do inglês *Chemical shift perturbation*) aquando a adição de um ligando, 2) realizar estudos de permuta hidrogénio/deutério, 3) estudar a dinâmica molecular de uma proteína (determinação de taxas de relaxação R_1 e R_2 e do efeito heteronuclear de Overhauser (HetNOE)). Do ponto de vista do ligando, existem também diversas técnicas que podem ser utilizadas, tendo como principal vantagem não ser necessária a marcação isotópica. Em caso de união à proteína a mobilidade do ligando altera-se o que conduz a uma diminuição de T_2 , a um aumento do tempo de correlação e a uma diminuição do coeficiente de difusão do ligando. Por exemplo experiências de relaxação T_1 e T_2 ou de difusão do ligando, na ausência e na presença de proteína, permitem inferir se existe ou não interação do ligando com a proteína. Também se podem realizar experiências com base no efeito nuclear de Overhauser (NOE), tais como o TrNOE (do inglês *Transferred-nuclear Overhauser effect*) e o STD (do inglês *Saturation transfer difference*), através das quais é possível elucidar a conformação bioativa do ligando e identificar os prótons do ligando que se encontram mais envolvidos no reconhecimento molecular com a proteína, respetivamente.

1.10.1- Estudo de um complexo proteína-ligando do ponto de vista da proteína: $^1\text{H},^{15}\text{N}$ -HSQC

A experiência de $^1\text{H},^{15}\text{N}$ -HSQC é uma experiência 2D que utiliza a capacidade de marcar isotopicamente o átomo de azoto, e que permite observar o sinal de NH da ligação peptídica e das cadeias laterais dos aminoácidos que constituem uma proteína³⁸. Normalmente, o espectro de $^1\text{H},^{15}\text{N}$ -HSQC de uma proteína é considerado a “impressão digital”, como tal cada proteína possui o seu espectro de $^1\text{H},^{15}\text{N}$ -HSQC característico que é único. Através da observação do espectro de $^1\text{H},^{15}\text{N}$ -HSQC é também possível inferir se

a proteína se encontra estruturada ou, se por outro lado, se encontra desnaturada ou estruturalmente não funcional³⁹. Um espectro de $^1\text{H},^{15}\text{N}$ -HSQC de uma proteína num estado desnaturado tipicamente apresenta uma dispersão limitada dos seus sinais, uma vez que, o ambiente químico sentido pela maioria dos aminoácidos é bastante idêntico. Contrariamente, uma proteína bem estruturada apresenta um espectro de $^1\text{H},^{15}\text{N}$ -HSQC com uma grande dispersão dos sinais, onde os aminoácidos possuem desvios químicos significativamente distintos.

No âmbito do estudo de interações proteína-ligando a experiência de $^1\text{H},^{15}\text{N}$ -HSQC é amplamente utilizada. Isto deve-se ao fato de os desvios químicos dos NHs dos aminoácidos da proteína serem fortemente alterados conforme o ambiente químico sentido pelo núcleo em questão. Desta forma, durante um evento de reconhecimento molecular proteína-ligando, existe uma alteração no ambiente químico sentido pelos resíduos da proteína devido à presença do ligando, o que leva a uma perturbação do desvio químico dos NHs. Experimentalmente, realiza-se a aquisição de espectros de $^1\text{H},^{15}\text{N}$ -HSQC, na ausência e na presença de diferentes concentrações de ligando. Com isto, um determinado resíduo da proteína pode experienciar dois ambientes químicos diferentes, na ausência e na presença do ligando, os quais irão ser detetados no espectro de $^1\text{H},^{15}\text{N}$ -HSQC, dependendo da frequência do espectrómetro e da diferença entre as frequências da forma livre e complexada da proteína. Este fenómeno é conhecido como permuta química. O regime de permuta química pode ser caracterizado em três diferentes tipos, com base na constante de permuta existente entre as duas formas livre e complexada, conhecida por k_{ex} , e pela diferença da frequência das duas formas $\Delta\omega = \omega_{\text{complexado}} - \omega_{\text{livre}}$. Na situação em que $k_{\text{ex}} \ll \Delta\omega$, o regime é denominado de equilíbrio lento, e os δ s de ambos os sinais correspondentes às duas diferentes formas surgem resolvidos no espectro (Figura 3-A). Quando $k_{\text{ex}} \gg \Delta\omega$, o regime é denominado de equilíbrio rápido, e o espectro é marcado pelo aparecimento de apenas um sinal em que o desvio químico é ponderado pela população existente em cada um dos estados da proteína livre ou complexada, calculado por $\delta_{\text{obs}} = p_{\text{livre}}\delta_{\text{livre}} + p_{\text{complexado}}\delta_{\text{complexado}}$ (Figura 3-B). Entre estes dois regimes, existe um regime intermédio em que $k_{\text{ex}} \approx \Delta\omega$ e os dois sinais coalescem, resultando num alargamento do sinal no espectro^{37,39}.

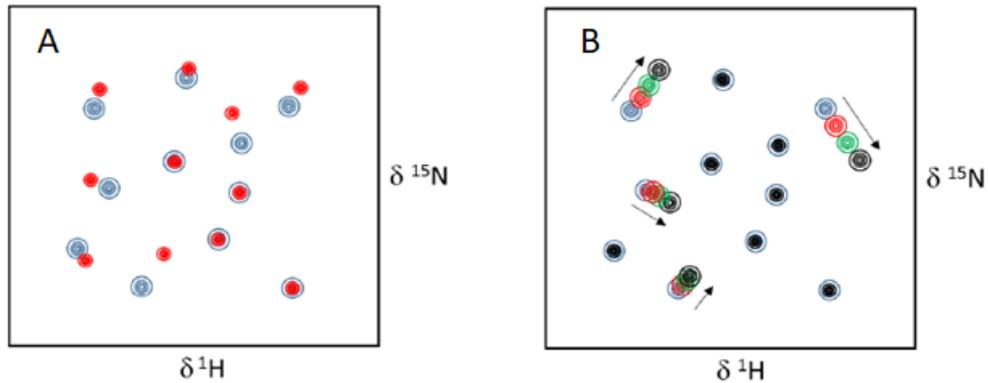


Figura 3: Representação esquemática de $^1\text{H},^{15}\text{N}$ -HSQCs ilustrando permuta lenta e rápida entre uma proteína em estado livre e em estado complexado. **A-** Regime de equilíbrio lento entre a forma livre e a forma complexada. Neste tipo de regime, existe a presença de dois sinais correspondentes à forma livre e forma complexada da proteína, cada um representando a população desses dois estados. **B-** Regime de equilíbrio rápido. Este tipo de regime é caracterizado pela existência de apenas um sinal correspondente à média ponderada entre as populações da proteína no estado livre e no estado complexado.

1.11- Estudo da dinâmica no contexto do reconhecimento molecular

A dinâmica de uma proteína é uma propriedade que influencia a capacidade de uma proteína desempenhar a sua função, em particular de se unir ou não a determinado ligando. Em solução, à temperatura ambiente, as proteínas assumem diversas conformações e experienciam um grande número de movimentos em diferentes escalas de tempo. Todas estas alterações conformacionais são dirigidas pela termodinâmica desses estados, nomeadamente pelas barreiras de energia que os separam. A possibilidade de uma proteína se unir a um determinado ligando, e com isso desencadear uma resposta biológica, está dependente de ela assumir um ou mais estado(s) conformacionais ativos durante um tempo de vida adequado, bem como da população existente nesse estado^{40,41}.

Com vista ao estudo da dinâmica de uma proteína, o RMN demonstra ter um enorme potencial para ser utilizado neste tipo de análise. A evolução da técnica e do tipo de experiências possíveis de realizar tornou possível observar vários tipos de movimentos, que tipicamente ocorrem em biomoléculas, e em diversas escalas de tempo, desde os pico segundos (ps), em que são detetados movimentos relacionados com vibrações de

ligações e movimentos locais, aos segundos (s), em que existem movimentos de enrolamento (do inglês *folding*) de proteínas⁴² (Figura 4).

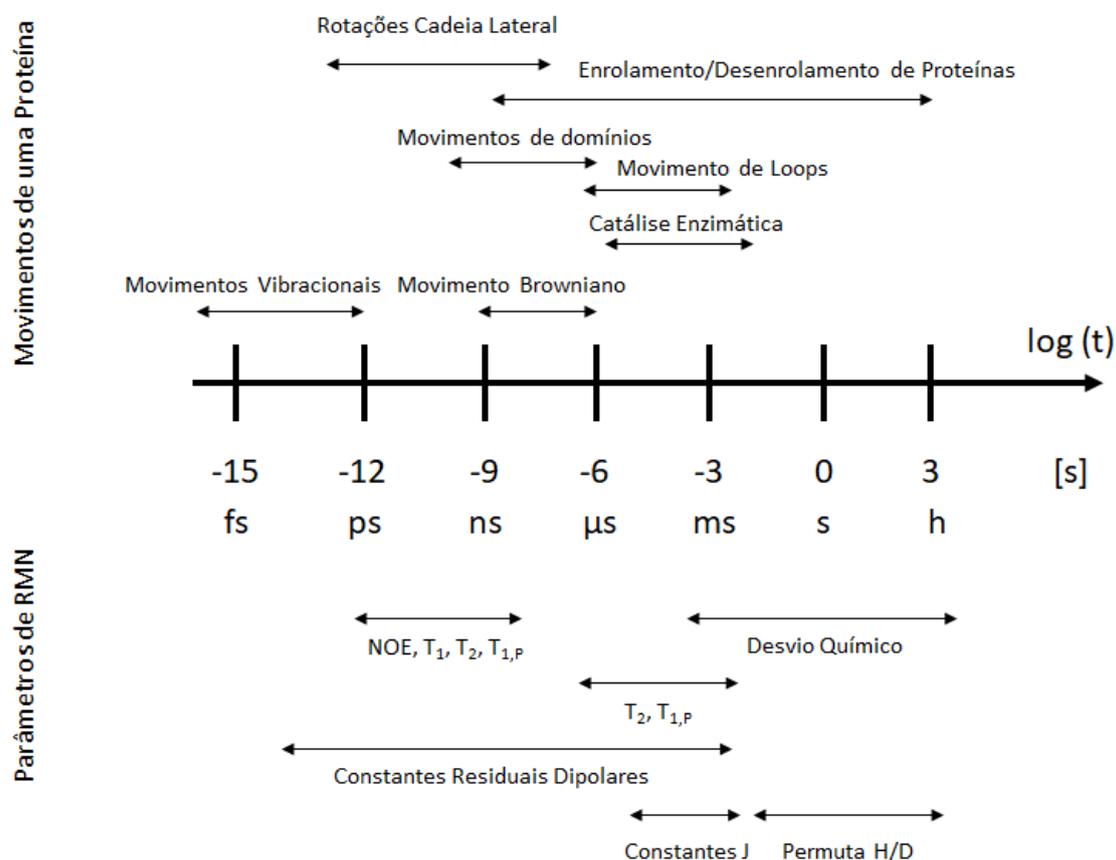


Figura 4: Escala de tempo representando os tipos de movimentos presentes em proteínas e as respectivas técnicas de RMN normalmente utilizadas para o seu estudo. Adaptado de Chao et al.⁴¹

A análise da dinâmica de uma proteína no estado livre e no estado associado permite obter então informação acerca do efeito do ligando na mobilidade da cadeia peptídica ou das cadeias laterais da região onde o ligando se liga, mas também permite avaliar alterações, caso ocorram, na dinâmica da proteína em regiões distantes do sítio de interação com o ligando.

Regra geral, nestes estudos, utiliza-se o modelo desenvolvido por Lipari e Szabo, denominado de *Model-Free*⁴⁰, que permite retirar informações acerca da dinâmica de uma proteína a partir de dados de relaxação adquiridos por RMN. Este modelo é particularmente sensível ao movimento molecular do vetor da ligação NH da cadeia principal, o qual é perturbado no contexto da interação com o ligando, tratando o

movimento de cada vetor de forma independente do movimento rotacional global da molécula.

Os parâmetros medidos para este estudo são normalmente a taxa de relaxação longitudinal ($R_1=1/T_1$) e a taxa de relaxação transversal ($R_2=1/T_2$), bem como o efeito HetNOE, e os quais podem ser extraídos a partir da análise de espectros de ^1H - ^{15}N -HSQC^{40,42}. Estes parâmetros, quando medidos a dois campos magnéticos diferentes, permitem obter parâmetros com uma maior precisão e uma maior informação acerca de movimentos locais numa escala de tempo contida entre os ps e os nanossegundos (ns). A análise de estes dados usando o modelo desenvolvido por Lipari e Szabo permite avaliar a existência de alterações na dinâmica local da proteína aquando a união de um ligando, extrair informação acerca do tamanho e da forma aparente da proteína e da existência ou não de estrutura quaternária de uma proteína. Com os dados de dinâmica é possível, através da média da razão R_2/R_1 dos resíduos NH da proteína, retirar informações sobre o movimento global de uma proteína, ou seja, o tempo de correlação, τ_c , (tempo médio que uma molécula demora a girar um radiano sobre um eixo) da proteína e se ele se altera ou não na presença do ligando^{40,43}.

Os fenómenos de relaxação de uma proteína em RMN são controlados pelos movimentos existentes das moléculas em solução. As taxas de relaxação longitudinal, R_1 , e transversal, R_2 , são sensíveis aos movimentos internos na ordem dos ps e do ns, onde predominam as interações dipolo-dipolo resultantes das flutuações do vetor de ligação NH em relação ao campo magnético permanente (a relaxação é dominada pelo protão devido à sua distância do átomo de azoto) e também da anisotropia do desvio químico devido à presença do grupo amida que é restringido no seu movimento. Para além disso, a taxa de relaxação R_2 é também sensível a movimentos mais lentos na ordem dos milissegundos (ms) a microssegundos (μs), conhecidos por movimentos de permuta química, que alteram a conformação de resíduos da proteína numa escala de tempo que resulta numa alteração do valor de R_2 medido.

Por fim, o valor de HetNOE permite inferir acerca do nível das flutuações existentes na ordem dos ps a ms do vetor da ligação NH da cadeia peptídica e é caracterizado por um

valor entre 0 e 1, sendo um valor mais alto relacionado com uma maior restrição dos movimentos e um valor mais baixo com uma maior liberdade de movimentos.

A partir dos dados de R_1 , R_2 e HetNOE torna-se possível, a partir do *Model-Free*, determinar movimentos que ocorrem numa escala de tempo rápida e os quais podem ser descrito por três parâmetros: o tempo de correlação global da molécula, o tempo de correlação local e o parâmetro S^2 , que compara a ordem entre estes dois tempos de correlação, tomando valores entre 0 e 1, e onde os valores mais baixos significam uma maior amplitude de movimentos locais.

1.12- Estudos prévios realizados sobre a MGL

O reconhecimento molecular de glicanos contendo a unidade GalNAc por parte da lectina MGL tem sido objeto de estudo do grupo de investigação *(Bio)Molecular Structure and Interactions by NMR* ^{28,44,45}.

Recentemente, estudou-se o reconhecimento molecular do ponto de vista da proteína de diferentes glicanos contendo a unidade GalNAc, incluindo o glicósido de metilo da *N*-acetilgalactosamina (α -MeGalNAc), e outros ligandos, tais como, o antígeno do grupo sanguíneo do tipo A, o antígeno *Forssman* e as estruturas derivadas de glicoesfingolípidos, tais como, o GM2 e asialo GM2 (Figura 5)⁴⁵.

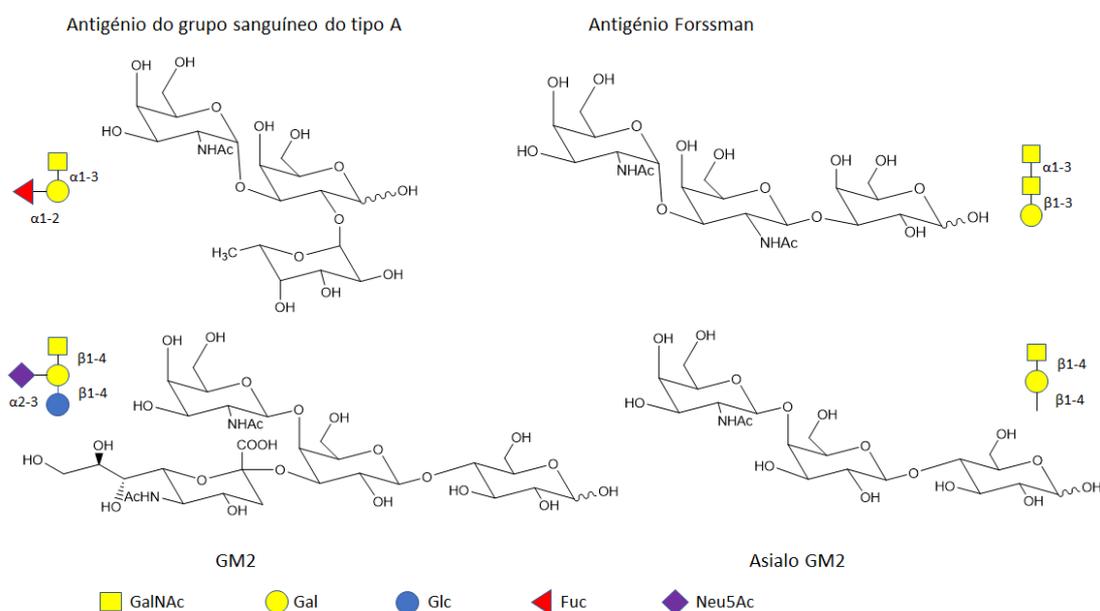


Figura 5: Estrutura química dos ligandos reconhecidos pela MGL anteriormente estudados e as suas respetivas representações esquemáticas.

Numa primeira fase realizou-se a atribuição dos sinais do espectro de $^1\text{H},^{15}\text{N}$ -HSQC da cadeia principal domínio CRD da MGL (MGL-CRD) na forma livre e complexada com o ligando α -MeGalNAc. Este procedimento foi essencial por diferentes motivos: 1) devido ao aparecimento de novas ressonâncias no espectro, na presença de α -MeGalNAc, e 2) devido à interação com o ligando ocorrer em regime de permuta lenta. Após a atribuição dos sinais dos resíduos da proteína em ambos os estados, foi possível determinar a região da proteína mais afetada pela presença do ligando α -MeGalNAc.

Examinando os espectros de $^1\text{H},^{15}\text{N}$ -HSQC do domínio MGL-CRD, na forma livre e na presença do ligando α -MeGalNAc, é possível observar o aparecimento de 11 resíduos pertencentes ao CRD da MGL que não existiam no espectro do MGL-CRD sua forma livre. Estes resíduos correspondem aos aminoácidos: D270, G272, H274, G275, L276, D288, N292, D293, D294, V295 e a R298 (Figura 6, aminoácidos marcados com *). Foi também detetado, do espectro de $^1\text{H},^{15}\text{N}$ -HSQC do domínio MGL-CRD na forma livre para a forma complexada, um aumento da intensidade de várias ressonâncias. Este resultado indica uma alteração na dinâmica da proteína aquando da adição do ligando.

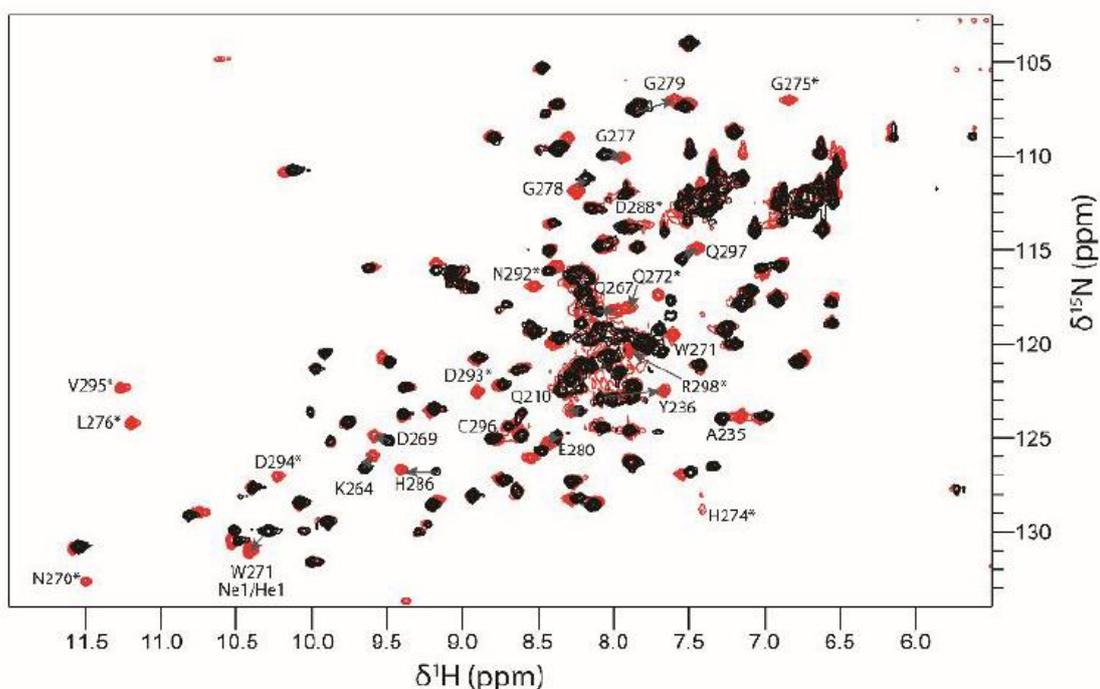


Figura 6: Espectro de $^1\text{H},^{15}\text{N}$ -HSQC do domínio MGL-CRD na ausência (preto) e na presença de α -MeGalNAc (vermelho). Os sinais indicados com um * representam os resíduos que apareceram na presença de α -MeGalNAc e os que não contêm um * indicam aqueles que foram identificados no estudo de CSP. Adaptado de Diniz et al.⁴⁵

Para além disso, a partir da análise de CSP demonstrada na figura 7, foram detetados também diversos resíduos que sofreram alterações do seu desvio químico na presença do ligando α -MeGalNAc, entre os quais os resíduos A235, Y236, K264, Q267, D269, W271, G277, G278, G279, E280, H286, C296, Q297 e H301.

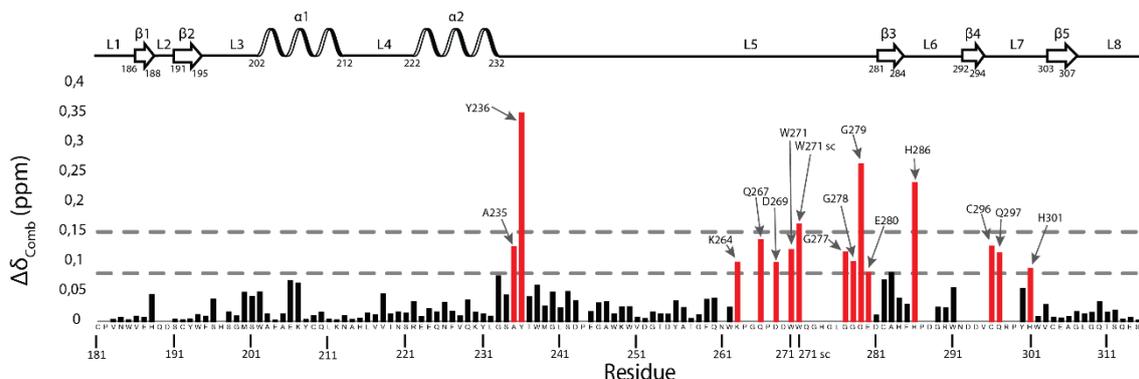


Figura 7: Histograma do CSP do domínio MGL-CRD na presença de α -MeGalNAc. As barras a vermelho identificam os resíduos considerados como significativos na perturbação da estrutura do MGL-CRD na presença do ligando. As duas linhas a cinzento e a tracejado representam os valores de $\Delta\delta_{comb}$ que discriminam os resíduos que induzem uma forte ou moderada perturbação no desvio químico. A primeira linha foi marcada a um valor de 0,08 ppm (perturbação moderada) e a segunda a 0,15 ppm (perturbação forte). Adaptado de Diniz et al.⁴⁵

Todos os resíduos afetados pela presença α -MeGalNAc foram então mapeados no modelo 3D da MGL, previamente gerado pelo grupo (Figura 8). O modelo 3D usado é um modelo de homologia gerado a partir do CRD da lectina asialoglicoproteína, a qual partilha cerca de 73% de homologia em termos de sequência primária com a MGL-CRD, e do qual existe estrutura de cristalografia (PDB: 1DV8).⁴⁶

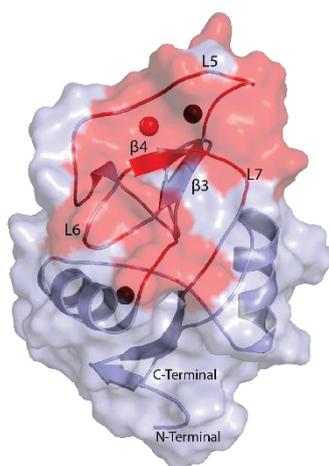


Figura 8: Mapa da interação da MGL-CRD com o ligando α -MeGalNAc. Este mapa inclui os resíduos que aparecem e os que foram identificados através da análise de CSP. Adaptado de Diniz et al.⁴⁵

A Figura 8 mostra que uma grande região da proteína é perturbada pela presença do ligando. A região perturbada é muito mais extensa e vai para além do motivo de reconhecimento definido pelos três aminoácidos QPD (Q267 e D269 na sequência). Este resultado indica que o MGL-CRD na forma livre é uma estrutura altamente dinâmica, adotando diversas conformações (não detetáveis por RMN), e que o ligando induz fortes alterações na estrutura e dinâmica do domínio MGL-CRD. De salientar que a grande maioria dos sinais perturbados pertencem à região do *loop* 5, seguido de alguns aminoácidos pertencentes aos *loops* 6 e 7. Em particular, o aparecimento de cinco resíduos pertencentes ao *loop* 5, aquando da presença do ligando, permite considerá-lo como uma região muito flexível e de grande importância no fenómeno de reconhecimento pelo domínio MGL-CRD.

Foi também estudado anteriormente, do ponto de vista do ligando, a interação entre a MGL (MGL-FL, do inglês MGL-full length,) e o ligando α -MeGalNAc utilizando a técnica de STD e também de dinâmica molecular (MD, do inglês *Molecular dynamics*), utilizando o modelo de homologia criado para o CRD da MGL, tendo sido propostos dois modos de ligação para o ligando α -MeGalNAc (modo A e modo B), onde o anel do ligando α -MeGalNAc pode assumir duas orientações possíveis no sítio de união da proteína ⁴⁴. Apenas assumindo a presença de dois modos de ligação foi possível interpretar os resultados de STD. Contudo, os dados recentes de ¹H,¹⁵N-HSQC não são suficientes para identificar a existência de ambos os modos de união nem permitem inferir qual deles poderá existir maioritariamente em solução.

Como foi dito anteriormente, mais recentemente estudou-se o reconhecimento do domínio MGL-CRD, utilizando experiências de RMN que permitem analisar, do ponto de vista da proteína, a interação entre domínio MGL-CRD e alguns ligandos mais complexos⁴⁵, tais como, o antigénio do grupo sanguíneo do tipo A, o antigénio *Forssman* e estruturas derivadas de glicosíngolípido, tais como, o GM2 e asialo GM2. Analisando os espetros de ¹H,¹⁵N-HSQC do domínio MGL-CRD na presença dos novos ligandos e comparando-os com o espetro de ¹H,¹⁵N-HSQC na presença da referência α -MeGalNAc foi possível calcular CSPs e observar diferenças no estado unido do domínio MGL-CRD consoante o ligando em estudo (Figura 9).

O antígeno do grupo sanguíneo do tipo A mostrou-se ser o ligando mais parecido com o α -MeGalNAc, não tendo sido detetadas diferenças significativas de desvio químico entre os $^1\text{H},^{15}\text{N}$ -HSQCs do domínio MGL-CRD unido a estes dois ligandos. Por outro lado, o antígeno *Forssman* mostrou na análise de CSP, com respeito ao ligando α -MeGalNAc, uma perturbação nos resíduos H274, L276 e G279 (todos pertencentes ao *loop* 5). Contudo, os valores de CSP mais significativos foram apresentados para os ligandos GM2 e asialo GM2. Em particular com estes ligandos foram observadas diferenças significativas nos resíduos K264, Q267, D269, H274, W291, V295 e Q297, na presença do ligando GM2, e nos resíduos K264, Q267, H274, H286 e W291, aquando da presença do ligando asialo GM2.

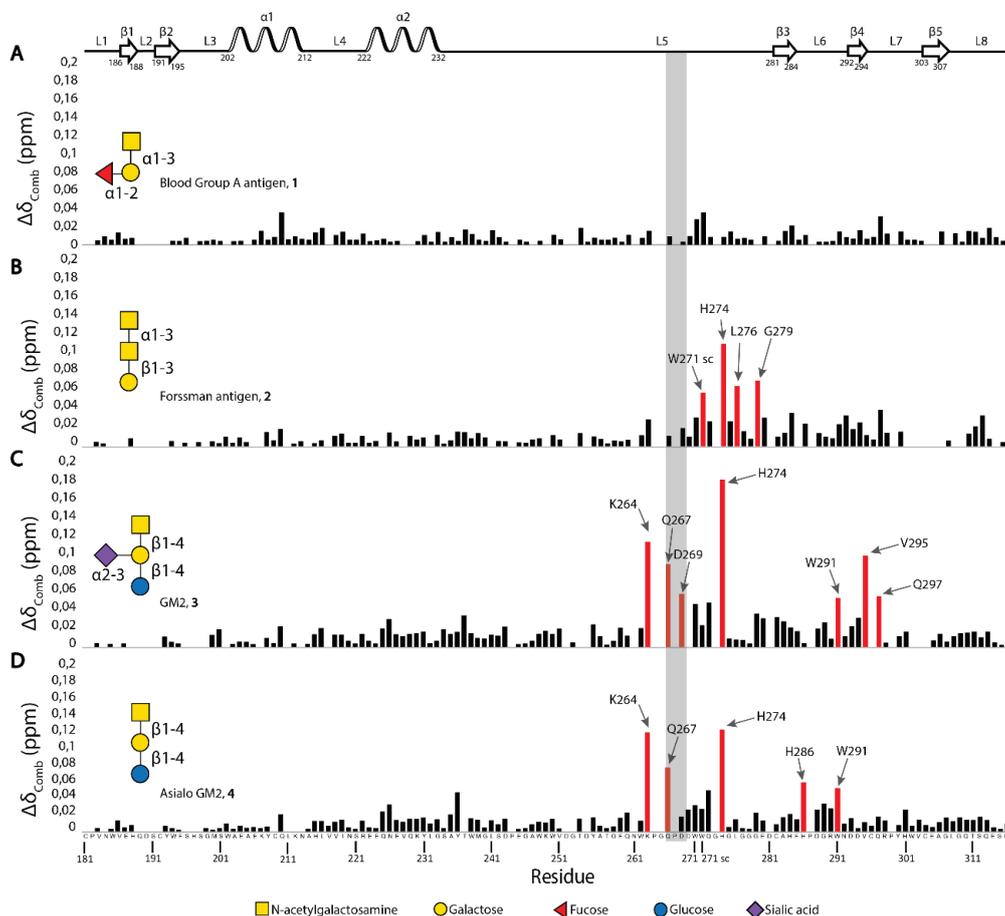


Figura 9: Histograma do CSP obtido entre os espectros de $^1\text{H},^{15}\text{N}$ -HSQC dos complexos de MGL-CRD na presença dos vários ligandos estudados e o espectro de $^1\text{H},^{15}\text{N}$ -HSQC do complexo MGL-CRD/ α -MeGalNAc. A vermelho encontram-se assinalados os resíduos que apresentavam CSP mais elevado (> 0.05 ppm). Adaptado de Diniz et al.⁴⁵

O reconhecimento molecular dos mesmos ligandos pelo CRD da MGL foi também investigado do ponto de vista do ligando, com recurso à técnica de STD. O ligando do grupo sanguíneo do tipo A apresentou apenas resultado de STD para a unidade GalNAc, o que está em concordância com os resultados de CSP. Os outros ligandos, em particular os ligandos GM2 e asialo GM2, mostraram que os outros resíduos, incluindo a unidade redutora, estão também envolvidos no reconhecimento molecular por parte do CRD da MGL.

O modelo 3D dos diferentes complexos foi também gerado por MD, utilizando o modelo de homologia anteriormente usado para o ligando α -MeGalNAc. Ao contrário dos dois modos de ligação detetados para o ligando α -MeGalNAc, apenas um dos modos de ligação se mostrou estável o suficiente para a criação do complexo MGL-CRD com os novos ligandos, devido à existência de diversos impedimentos estereoquímicos entre a proteína e os novos ligandos. As estruturas 3D geradas são estáveis e compatíveis com os dados experimentais de CSP e de STD. Em suma, estes resultados demonstram que a estrutura do CRD permite acomodar diferentes ligandos contendo a unidade de GalNAc, e que esta é bastante sensível à estrutura do ligando em questão. Esta conclusão relaciona-se também com resultados obtidos anteriormente pelo grupo⁴⁴ e onde tinha sido identificado que os aminoácidos em torno do antígeno Tn (GalNAc-Ser/Thr), usando um glicopéptido derivado da proteína mucina-1 (MUC1), recebem resposta de STD (Figura 10). Esta evidência parece demonstrar que pelo menos os aminoácidos localizados em torno do antígeno Tn estabelecem interações com o domínio MGL-CRD.

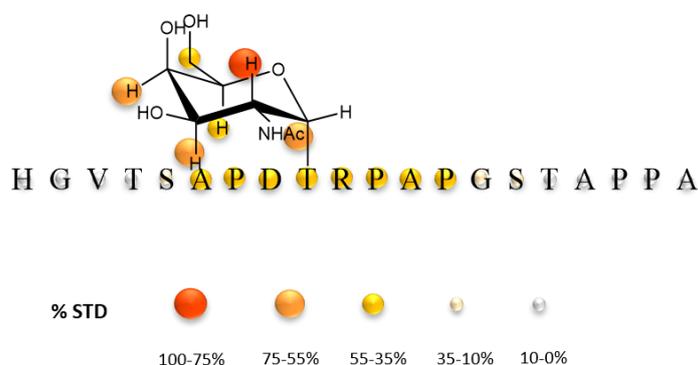


Figura 10: Experiência de STD do glicopéptido Tn-T8 MUC1. Os valores de % de STD encontram-se representados na estrutura por círculos com a cor correspondente à sua % de STD recebida. Adaptado de Marcelo et al.⁴⁴

2- Objetivos

A glicosilação aberrante é uma das principais características das células cancerígenas que contribuem para a sua capacidade de evadir e diminuir a resposta do sistema imunitário no cancro. Através da interação entre determinados antigénios presentes nas superfícies das células tumorais e lectinas presentes em células do sistema imunitário, as células cancerígenas conseguem modular o tipo de resposta que é despoletada pelas células do sistema imunitário. Neste contexto, a lectina MGL é uma lectina presente nos macrófagos e nas células dendríticas imaturas, a qual reconhece especificamente resíduos de GalNAc, normalmente presentes em células tumorais, em glicoproteínas das células T e em agentes patogénicos. Diversos estudos realizados indicam que a MGL tem a capacidade de modelar a resposta imunitária das células do sistema imune, num ambiente de reconhecimento de agentes invasores ou num ambiente tumoral, através da interação com ligandos contendo a unidade de GalNAc. Desta forma a elucidação do mecanismo de reconhecimento molecular a nível atómico da MGL frente a distintos ligandos contendo a unidade GalNAc é crucial para se compreender a resposta imunológica desencadeada pela MGL bem como ajudar no desenvolvimento de miméticos que possam interferir na interação MGL/GalNAc e assim modular a resposta imune.

O presente trabalho é inspirado nos resultados recentes do grupo com a proteína MGL e teve como objetivo principal investigar o reconhecimento molecular de distintos ligandos, contendo a unidade GalNAc com distintas apresentações, por parte do domínio MGL-CRD, com recurso à técnica de RMN. O trabalho envolveu dois estudos distintos e como tal a presente tese está dividida em dois capítulos.

Na primeira parte investigou-se o mecanismo de reconhecimento molecular do MGL-CRD frente a diferentes glicopéptidos derivados da proteína MUC1 (Figura 13), contendo o antigénio Tn (GalNAc-Thr) em distintos locais de glicosilação do péptido com o objetivo de determinar o papel do péptido no reconhecimento molecular. Em detalhe, através da análise detalhada dos espetros de $^1\text{H},^{15}\text{N}$ -HSQC, pretendeu-se identificar a contribuição da sequência peptídica na apresentação do antigénio Tn ao domínio MGL-CRD e conseqüentemente o nível de interações que existe entre o péptido e a proteína.

A segunda parte do trabalho teve como base os estudos recentemente publicados pelo grupo, os quais claramente indicaram que o MGL-CRD é altamente flexível e dinâmico. Com isto, neste trabalho, iniciou-se o estudo da dinâmica molecular da cadeia principal do domínio de MGL-CRD por RMN, na ausência e presença de α -MeGalNAc, bem como na presença dos ligandos previamente estudados e ilustrados na figura 5. Técnicas de relaxação baseadas em ^{15}N foram aplicadas com o intuito de detetar eventuais alterações na dinâmica dos resíduos do CRD da MGL dependendo do contexto do ligando.

3-Métodos e Procedimento Experimental

3.1-Expressão e Purificação do domínio MGL-CRD uniformemente marcada em ¹⁵N

O procedimento experimental foi realizado com base no protocolo publicado anteriormente³², com algumas diferenças apresentadas neste capítulo.

O gene contendo a região da MGL-CRD correspondente aos resíduos Q181-H316 foi subclonado no vetor de expressão pET21a pela NZYTech (Anexo 7.1).

3.1.1-Produção de células de *E. coli* competentes

Cresceram-se células de *Escherichia coli* (*E. Coli*) BL21 (DE3) em placas de Petri com meio LB Agar (Anexo 7.2) durante aproximadamente 16 h a 37°C (incubadora Orbital Shaker ES-20). De seguida, inoculou-se uma colónia isolada em meio LB (Anexo 7.2) e deixou-se a crescer durante, aproximadamente, 16 h a 37°C a 190 rpm (incubadora Orbital Shaker ES-20). A cultura foi diluída numa razão 1:100 em meio LB e as células cresceram durante 2 h a 37°C com 190 rpm até atingirem uma OD_{600nm} de 0,6 a 0,8. Recuperaram-se as células por meio de uma centrifugação a 4000 rpm durante 5 min a 4°C (Centrifuga 5804R eppendorf, rotor F-34-6-38). As células foram ressuspensas em 15 mL de MgCl₂ 0,1 M frio. Procedeu-se à sua centrifugação a 4000 rpm durante 5 min a 4°C, descartou-se o sobrenadante e ressuspendeu-se em 15 mL de CaCl₂ 0,1 M, em gelo. Por fim, as células foram novamente recuperadas e ressuspensas em 5 mL de CaCl₂ 0,1 M e glicerol 30%. Foram feitas alíquotas e armazenadas a -80°C.

3.1.2-Transformação de células de *E. coli* competentes

Pipetou-se 1 µL do DNA contendo o gene da proteína MGL-CRD em 50 µL de células *E. coli* BL21 (DE3) competentes e incubou-se durante 15 min em gelo. De seguida as células foram submetidas a um choque térmico a 42°C durante 40 s (AccuBlock™ Digital Dry Bath, Labnet International, Inc.) e incubadas novamente em gelo durante 15 min. Adicionaram-se 500 µL de meio LB à cultura de células e incubou-se a 37°C durante 1 h. Seguidamente, as células foram centrifugadas a 14000 rpm durante 30 s e descartou-se o sobrenadante. Inocularam-se, aproximadamente, 50 µL de células de *E. coli*

recombinantes numa placa de Petri contendo um meio LB agar com 100 µg/mL de ampicilina e incubou-se a 37°C durante aproximadamente 16 h.

3.1.3-Expressão e purificação da MGL-CRD isotopicamente marcada com ¹⁵N

A MGL-CRD foi sobreexpressa em células *E. Coli* BL21 (DE3) que cresceram em meio mínimo M9 (Anexo 7.2) suplementado com ¹⁵NH₄Cl e induzido com IPTG 1 mM a 25°C durante, aproximadamente, 16 h a 190 rpm. Posteriormente, o meio de cultura foi centrifugado a 6000 rpm durante 12 min a 4°C (Centrifuga Avanti J-26 XPI Beckman Coulter, rotor JA-10) e foi recolhido o *pellet* com o conteúdo celular.

O *pellet* contendo as células foi ressuscitado em tampão Tris-HCl 50 mM, NaCl 150 mM, a pH 8 e a lise das células foi realizada com recurso à técnica de sonicação. Efetuou-se a sonicação com 10 ciclos de 1 min on/off com uma amplitude de 80% em gelo (sonicadora Hielscher Ultrasound Technology UP100H). De seguida o lisado foi centrifugado a 11000 rpm durante 15 min a 8°C, o sobrenadante foi descartado e foi recolhido o *pellet*, devido à expressão da proteína ocorrer em corpos de inclusão.

Os corpos de inclusão foram ressuscitados em tampão Tris-HCl 50 mM, NaCl 150 mM, EDTA 10 mM e Triton X-100 0,5% a pH 8, e recuperados por meio de uma centrifugação a 11000 rpm durante 15 min a 8°C e este procedimento foi repetido duas vezes, descartando-se o sobrenadantes ambas as vezes. Posteriormente, os corpos de inclusão foram ressuscitados em H₂O desionizada e recuperados por meio de uma centrifugação a 11000 rpm durante 15 min a 8°C. Por fim, os corpos de inclusão foram solubilizados em NH₄OH 2 M em agitação durante 3,5 h. Após este período, a solução dos corpos de inclusão solubilizados foi dialisada contra uma solução tampão Tris-HCl 10 mM, NaCl 75 mM e CaCl₂ 20 mM a pH 7,5 durante, aproximadamente, 16 h (fita de diálise SnakeSkin® Dialysis tubing 7K MWCO). As soluções foram recolhidas da diálise e centrifugadas a 19000 rpm durante 30 min a 8°C.

Finalmente, procedeu-se à purificação da proteína MGL-CRD por cromatografia de afinidade (ÄKTA start), sendo utilizada uma coluna de afinidade de lactose (α-Lactose-Agarose Sigma Aldrich L7634). O sobrenadante resultante da centrifugação foi então

injetado na coluna. A coluna foi lavada com tampão de lavagem Tris-HCl 10 mM, NaCl 75 mM e CaCl₂ 20 mM a pH 7,5 durante 50 mL. Para eluir a proteína foi usado o tampão de eluição Tris-HCl 10 mM, NaCl 75 mM, CaCl₂ 20 mM e 150 mM de lactose a pH 7,5. As frações eluídas contendo a proteína foram sujeitas a uma cromatografia em coluna de *desalting* (HiTrap™ Desalting, GE Healthcare), seguida de um passo de diálise contra a solução de Tris-HCl 10 mM, NaCl 75 mM, CaCl₂ 20 mM a pH 7,5.

Por último, as frações dialisadas contendo o domínio MGL-CRD sofreram um passo de concentração em que foram utilizados concentradores (Vivaspin 4 Turbo 10k MWCO).

A monitorização da expressão e purificação do domínio de MGL-CRD foi realizada através da técnica de eletroforese em gel de poliacrilamida em condições desnaturantes, em particular na presença de dodecil sulfato de sódio (SDS-PAGE). A corrida foi realizada com um gel a 10% de acrilamida durante 45 min a uma voltagem constante de 180 V com a amperagem a variar (Bio-Rad). Em anexo estão descritas as soluções usadas na preparação do gel de concentração e do gel de separação e também dos tampões de corrida (Anexo 7.3).

3.2- Espectroscopia de Ressonância Magnética Nuclear

As experiências de Ressonância Magnética Nuclear (RMN) foram realizadas no espectrómetro Bruker Avance III de 600 MHz equipado com uma crio-sonda de tripla ressonância.

3.2.1- Interação da MGL-CRD com os diferentes glicopéptidos

As interações entre a MGL-CRD com os diferentes ligandos foram analisadas através da aquisição de espectros ¹H,¹⁵N-HSQC da proteína MGL-CRD, isotopicamente marcada com ¹⁵N, no estado livre e no estado complexado com os diferentes ligandos. Os espectros ¹H,¹⁵N-HSQC foram adquiridos com uma resolução de 2048 × 128 pontos, com 32 scans e uma janela espectral de 9615,4 Hz (centrado a 600,13 MHz) em ¹H por 1946,2 Hz (centrado a 60,82 MHz) em ¹⁵N. Os espectros adquiridos foram processados no programa

Bruker TopSpin 4.0.6 (Bruker) e analisados no programa Cara 1.9.1.5⁴⁷ e CcpNMR Analysis 2.4.⁴⁸.

Todas as experiências foram realizadas em tampão Tris-HCl 10 mM, NaCl 75 mM, CaCl₂ 20 mM a pH 7 a uma temperatura de 293K. Todos os estados complexados foram estudados com uma razão de 1:2 proteína: ligando, com uma concentração de MGL-CRD de 100 μM, 90 μM e 200 μM no caso dos complexos MGL-CRD/α-MeGalNAc, MGL-CRD/Tn-Thr e MGL-CRD/Tn-T3 MUC1, Tn-T15 MUC1 e Tn-T3/T15 MUC1, respetivamente. O espectro de ¹H,¹⁵N-HSQC da MGL-CRD no estado livre foi adquirido antes de cada adição de ligando.

3.2.2- Experiências de dinâmica da MGL-CRD: Estudos de Heteronuclear NOE, R₁ e R₂

As experiências de dinâmica da MGL-CRD HetNOE, R₁ e R₂ foram realizadas com a proteína MGL-CRD a 200 μM no estado livre e na presença de diferentes ligandos, nomeadamente α-MeGalNAc, antigénio do grupo sanguíneo do grupo A, antigénio *Forssman*, GM2 e asialo GM2, com uma razão de 1:10 proteína: ligando, em tampão Tris-HCl 10 mM, NaCl 75 mM, CaCl₂ 20 mM a pH 7,4 a uma temperatura de 293K.

Os espectros de HetNOE foram adquiridos com um tempo de relaxação entre scans de 5 s, numa matriz de 2048 × 256 pontos, com 32 scans e uma janela espectral de 9615,4 Hz (centrado a 600,13 MHz) por 1946,2 Hz (centrado a 60,82 MHz).

As experiências de R₁ e de R₂ foram adquiridas como uma experiência pseudo-3D, onde foram adquiridos um conjunto de espectros 2D ¹H,¹⁵N- HSQC onde se variou o tempo de relaxação para cada um dos espectros 2D ¹H,¹⁵N- HSQC.

Foram adquiridos, para a determinação da taxa de relaxação R₁, um conjunto de quinze espectros com os tempos de relaxação de 0,002 s, 0,050 s, 0,100 s, 0,200 s, 0,300 s, 0,450 s, 0,600 s, 0,800 s, 1 s, 1,2 s, 1,5 s e 1,8 s. Os espectros foram adquiridos numa matriz de 2048 × 128 pontos com 32 scans, com um tempo de relaxação entre scans de 1,5 s e uma janela espectral de 9615,385 Hz (centrado a 600,13 MHz) por 1946,168 Hz (centrado a 60,82 MHz).

Para a determinação da taxa de relaxação R_2 , foram adquiridos um conjunto de quinze espectros com os tempos de relaxação de 0,0169 s, 0,0338 s, 0,0507 s, 0,0676 s, 0,0845 s, 0,1014 s, 0,1183 s, 0,1352 s, 0,1521 s, 0,1690 s, 0,1859 s, 0,2028 s, 0,2197 s, 0,2366 s e 0,2535 s, adquiridos por uma ordem alternada para evitar o aquecimento da amostra. Os espectros foram adquiridos com uma resolução de 2048×128 pontos, 32 scans, com um tempo de relaxação entre scans de 1,5 s e uma janela espectral de 9615,4 Hz (centrado a 600,13 MHz) por 1946,2 Hz (centrado a 60,82 MHz).

Os espectros adquiridos foram processados no programa Bruker TopSpin 4.0.6 (Bruker) com o mesmo fator de intensidade (nc_proc) e analisados no programa Cara 1.9.1.5, de onde foram retirados os valores de intensidade para o cálculo dos valores de NOE, de R_1 e de R_2 . O programa Relax foi utilizado para realizar o ajuste das curvas de intensidades necessário para estimar os valores de R_1 , R_2 e para o cálculo do valor de HetNOE.⁴⁹

3.2.3- Análise $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD unido a diferentes ligandos: Determinação da perturbação do desvio químico (CSP)

O cálculo do CSP é uma das principais metodologias utilizadas para estudar interações proteína:ligando. Para determinar quais os resíduos de NH que têm uma maior alteração no seu desvio químico devido à presença de ligando, procedeu-se ao cálculo do valor de desvio do desvio químico combinado ($\Delta\delta_{comb}$), um valor que melhor representa a diferença do desvio químico de cada um dos resíduos, na presença do ligando, em relação ao desvio químico obtido na ausência do mesmo. O valor de $\Delta\delta_{comb}$ é calculado com base na equação 1³⁹:

$$\Delta\delta_{comb} = \sqrt{(\Delta\delta_H)^2} + \sqrt{(0.1\Delta\delta_N)^2} \quad (\text{equação 1})$$

onde $\Delta\delta_H$ e $\Delta\delta_N$ representam a diferença do sinal do desvio químico do próton e azoto, respetivamente, obtido no espectro da forma complexada com o desvio químico do sinal correspondente na ausência de ligando. Para classificar os resíduos quanto ao seu grau de interação com os ligandos em estudo, delimitou-se dois valores de $\Delta\delta_{comb}$, o valor de 0,15 ppm, para os resíduos que mais são alterados, e um valor entre 0,08 e 0,15 ppm, para aqueles que sofrem uma perturbação significativa, mas não tão elevada.

3.2.4-Análise da dinâmica molecular do MGL-CRD unido a diferentes ligandos: HetNOE, R₁ e R₂

Na técnica de HetNOE adquirem-se dois espectros 2D ¹H,¹⁵N-HSQC: na ausência de saturação (espectro referência), e na presença de saturação dos núcleos de ¹H (espectro saturado). De seguida, extraem-se os valores de intensidade dos sinais no espectro de ¹H,¹⁵N-HSQC e o valor de NOE é calculado com base na equação 2⁵⁰:

$$NOE = \frac{I_{sat}}{I_{ref}} \text{ (equação 2)}$$

Com base nesta equação, o valor de 1 refere-se aos resíduos que menos movimentos locais sofrem e o valor de 0 corresponde a resíduos extremamente móveis e flexíveis. Os valores de incertezas das intensidades para cada espectro foram definidos como o valor da raiz do valor quadrático médio (RMS) de três medições do valor do ruído em zonas diferentes do espectro. As intensidades e os valores das incertezas foram extraídos usando o programa Cara 1.9.1.5. Por último foi utilizado o programa Relax, onde se calculou os valores de HetNOE para cada resíduo e o seu erro associado com recurso a simulações de Monte-Carlo executadas pelo mesmo.

A determinação dos valores de R₁ e de R₂ requiere a aquisição de um conjunto de espectros em que é variado o tempo de relaxação, ajustando depois os valores das intensidades de cada resíduo às equações exponenciais, equações 3 e 4, de onde são retirados os valores de R₁ e R₂⁵⁰:

$$I = I_0(1 - e^{\frac{-\tau}{R_1}}) \text{ (equação 3)}$$

$$I = I_0 e^{\frac{-\tau}{R_2}} \text{ (equação 4)}$$

onde I representa a intensidade de um resíduo no tempo de relaxação τ , e I₀ representa o valor de intensidade do mesmo resíduo em equilíbrio.

As equações 3 e 4 foram utilizadas e ajustadas por intermédio do programa Relax ⁴⁹, onde foram inseridos os valores das intensidades de cada um dos resíduos em cada tempo de relaxação retirados a partir do programa Cara 1.9.1.5, assim como os valores

das incertezas dos mesmos. De forma semelhante ao cálculo dos valores de NOE, o programa Relax permitiu também calcular o erro associado às medições com o uso do modelo de Monte-Carlo.

3.2.5- Determinação do tempo de correlação a partir da razão R_2/R_1

O tempo de correlação da MGL-CRD na forma livre e na presença dos seus ligandos foi estimado a partir da razão entre os valores médios estimados de R_2 e de R_1 ⁴⁰.

Para efetuar o cálculo foi necessário analisar resíduo a resíduo pois nem todos os resíduos podem entrar para o cálculo da média de R_1 e R_2 . Primeiro, os resíduos que apresentavam um valor de NOE < 0,65 foram excluídos porque demonstram uma grande flexibilidade a nível local. De seguida, foram também retirados todos os resíduos que demonstravam na sua dinâmica uma ocorrência substancial de permuta química na escala de tempo de μ s a ms, designado por R_{ex} , com base na condição da equação 5⁵¹:

$$\frac{\langle T_2 \rangle - T_{2,n}}{\langle T_2 \rangle} - \frac{\langle T_1 \rangle - T_{1,n}}{\langle T_1 \rangle} > 1.5 \times \sigma \text{ (equação 5)}$$

em que $\langle T_1 \rangle$ e $\langle T_2 \rangle$ é a média dos valores de T_1 ($1/R_1$) e T_2 ($1/R_2$) para cada ligando, $T_{2,n}$ é o valor do resíduo n em estudo e sigma (σ) é o desvio-padrão calculado da equação em cima.

Após a exclusão dos resíduos que não satisfazem estas condições, é feita a média dos valores da razão R_2/R_1 e obtido o valor estimado de τ_c .

4- Análise e Discussão de Resultados

4.1.-Expressão e Purificação da MGL-CRD isotopicamente marcada com ¹⁵N

A expressão do CRD da MGL humana marcado com ¹⁵N realizou-se em células de *E. Coli*. Contudo, uma vez que o domínio da MGL-CRD é expresso em corpos de inclusão, foi necessário proceder à sua solubilização (ver secção 3.1.3 – Expressão e purificação da MGL-CRD isotopicamente marcada). Após solubilização com 2 M de NH₄OH, a solução foi dialisada em tampão contendo CaCl₂ durante 16h (ver secção 3.1.3 – Expressão e purificação da MGL-CRD isotopicamente marcada). De seguida, efetuou-se a purificação do MGL-CRD solúvel por cromatografia de afinidade, em que foi utilizada uma coluna de agarose α -lactose. O domínio de MGL-CRD contém afinidade para a molécula de lactose imobilizada na coluna, o que permite então a sua separação de outras proteínas, uma vez que fica adsorvida na coluna. Após um passo de lavagem da coluna com o tampão de eluição, o passo de injeção da solução contendo o domínio de MGL-CRD e a passagem do tampão de lavagem, a eluição do domínio de MGL-CRD foi realizada pela passagem com o tampão de eluição, que contém 150 mM de lactose (Figura 11). Deste modo, a lactose ao competir com os resíduos de lactose imobilizados na coluna, faz com que o domínio de MGL-CRD seja eluído e removido da coluna⁵². O passo de cromatografia de afinidade de lactose permite diferenciar entre MGL-CRD funcional, a qual apresenta afinidade para a coluna, e não-funcional, que não mostra afinidade para a coluna. A proteína não-funcional é eluída mais rapidamente que a proteína funcional (com afinidade para lactose).

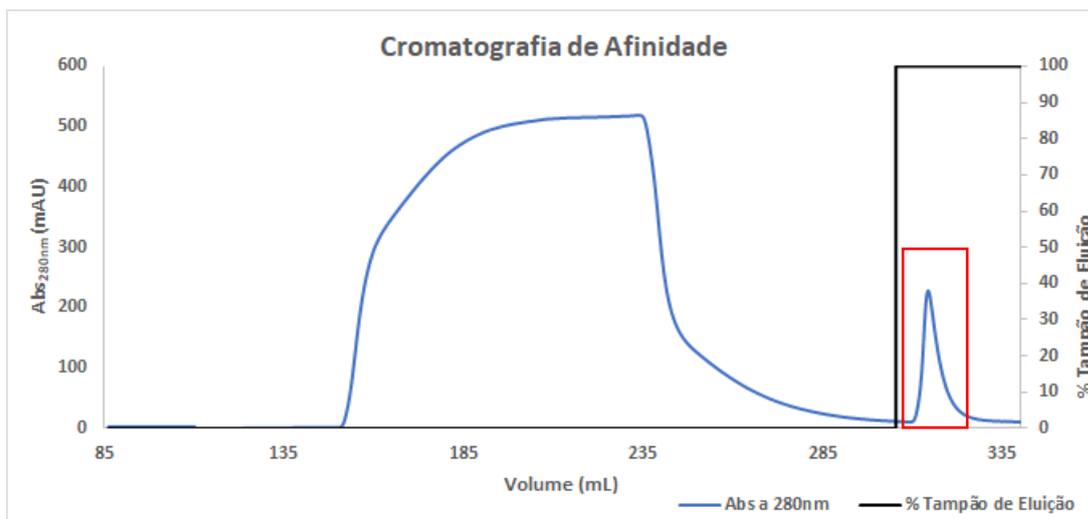


Figura 11: Cromatograma da cromatografia de afinidade. O pico correspondente à eluição do domínio de MGL-CRD encontra-se assinalado a vermelho. O tampão de eluição é composto por Tris-HCl 10 mM, NaCl 75 mM, CaCl₂ 20 mM e 150 mM de lactose a pH 7,5.

De seguida, as frações de eluição separadas na cromatografia de afinidade que continham a MGL-CRD foram sujeitas a uma coluna de *desalting* e a um processo de diálise com o intuito de remover toda a lactose presente (ver secção 3.1.3 - Expressão e purificação da MGL-CRD isotopicamente marcada com ¹⁵N).

Após estes passos, a pureza da proteína foi analisada por eletroforese em gel de poliacrilamida em condições desnaturantes com SDS, SDS-PAGE (do inglês, *sodium dodecyl sulfate – polyacrilamide gel electrophoresis*) onde foi possível observar a banda correspondente á MGL-CRD por volta dos 16 kDa (peso molecular estimado de 15,72 kDa para o MGL-CRD através da ferramenta ExPASy⁵³) (Figura 12). No gel podemos observar uma banda por volta dos 35 kDa, a qual corresponde à possível formação de dímeros.

O rendimento da expressão foi calculado por espectroscopia UV/VIS , tendo por base a lei de Lambert-Beer, usando o valor de coeficiente de extinção molar (ϵ) a 280nm de 63940 M⁻¹cm⁻¹ (retirado da ferramenta ExPASy)⁵³. O rendimento final obtido foi de 2,8 mg por litro de cultura.

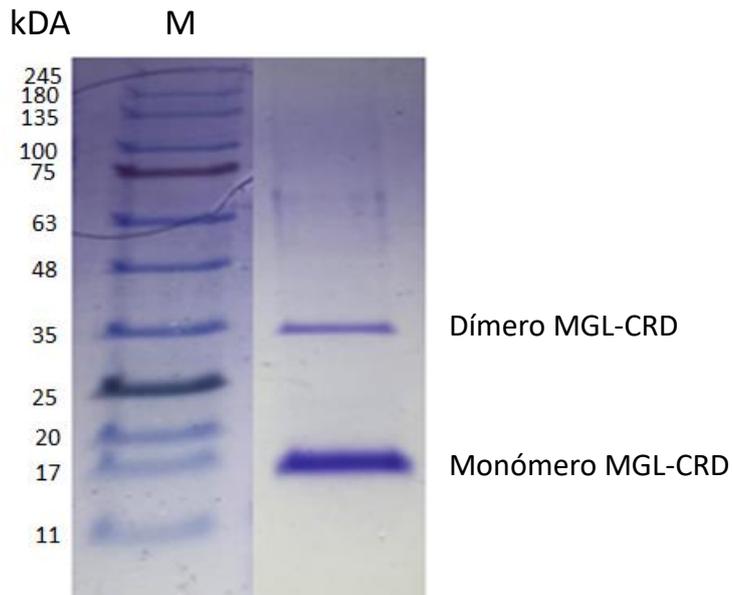


Figura 12: Gel de eletroforese (SDS-PAGE) da fração final de MGL-CRD após purificação. No gel observa-se uma banda correspondente ao peso molecular do domínio MGL-CRD e também uma banda correspondente ao peso molecular equivalente a uma possível formação de dímeros do domínio MGL-CRD. Os marcadores de pesos moleculares NZYColour Protein Marker II encontram-se no poço identificado com a letra M.

4.2-Estudos de interação da MGL-CRD com glicopéptidos cancerígenos por RMN

A participação da lectina MGL no reconhecimento molecular de estruturas cancerígenas *O*-glicosiladas, nomeadamente a sua interação com o antígeno Tn (GalNAc-Thr/Ser), parece conduzir a uma supressão da resposta imunitária no cancro²⁵.

As mucinas, proteínas transmembranares altamente glicosiladas, sofrem um aumento da sua expressão em células tumorais o qual é acompanhado por uma alteração no padrão da glicosilação e de onde surgem diversas estruturas de *O*-glicosilação ramificadas como, por exemplo, o antígeno Tn reconhecido pela MGL. A mucina-1 (MUC1) é uma glicoproteína em que o domínio extracelular contém um domínio de 20 aminoácidos repetido múltiplas vezes (domínio TR, **GVTSAPDTRPAPGSTAPPAH**). Cada domínio TR contém três treoninas e duas serinas perfazendo um total de cinco locais de *O*-glicosilação possíveis (identificados a negrito na sequência). A presença do antígeno Tn em células tumorais indica uma maior probabilidade de metástases e parece estar relacionado com um pior prognóstico⁵⁴. Devido a isso, é imperativo estudar as interações moleculares que estão na base do reconhecimento molecular dos glicopéptidos derivados MUC1 que contêm o antígeno Tn (Tn-glicopeptidos) e a lectina MGL.

Neste trabalho, estudaram-se quatro ligandos (Figura 13), um contendo apenas o aminoácido treonina glicosilado com o resíduo de GalNAc (Tn-Thr), dois péptidos Tn-monoglicosilados (Tn-T3 e Tn-T15) e um péptido Tn-diglicosilado (Tn-T3, T15). Com esta seleção de ligandos pretendeu-se estudar o papel da sequência peptídica e da densidade do antígeno Tn no processo de reconhecimento molecular de derivados da MUC1, aberrantemente expressos no cancro, por parte da MGL.

De forma alcançar este objetivo, foram adquiridos espectros de ¹H,¹⁵N-HSQC do domínio de MGL-CRD na ausência e presença dos ligandos e monitorizadas as alterações através da análise da perturbação do desvio químico (CSP) (ver secção 3.2.3 - Análise ¹H,¹⁵N-HSQC do MGL-CRD unido a diferentes ligandos: Determinação da perturbação do desvio químico (CSP)).

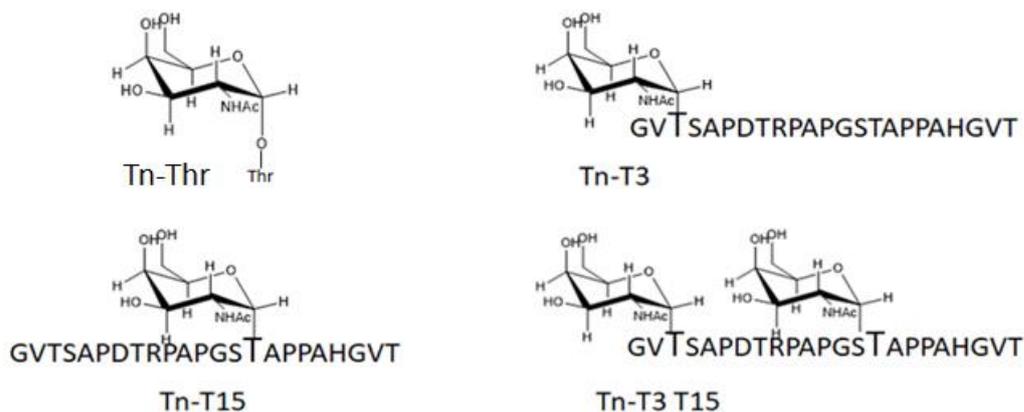


Figura 13: Estruturas químicas do ligando Tn-Thr e dos glicopéptidos da MUC1 contendo o antígeno Tn.

4.2.1- MGL-CRD complexado com Tn-Thr

Primeiramente, efetuou-se o estudo do ligando Tn-Thr (Figura 13). Através do cálculo do CSP entre os espectros de $^1\text{H},^{15}\text{N}$ -HSQCs do domínio MGL-CRD na forma livre e associada a Tn-Thr, foi possível determinar quais os resíduos da proteína que sofriam uma maior alteração na presença do ligando.

A interação entre o Tn-Thr e o domínio MGL-CRD ocorre num regime de permuta lenta, a qual resulta na presença de duas ressonâncias correspondentes às formas livre e complexada da MGL-CRD, e onde as intensidades correspondem às suas respectivas populações em solução. Este comportamento está em concordância com os resultados previamente publicados pelo grupo⁴⁵. Comparando os espectros do domínio MGL-CRD na ausência e presença do ligando Tn-Thr (razão proteína: ligando 1:2), verificamos a existência do aparecimento dos 11 resíduos anteriormente identificados, na presença de α -MeGalNAc, nomeadamente os aminoácidos D270, Q272, H274, G275, L276, D288, N292, D293, D294, V295 e R298 (Figura 14-A, assinalados a negrito). A análise do CSP é apresentada sob a forma de histograma (Figura 14-B), onde os valores de $\Delta\delta_{comb}$ são representados em função da sequência peptídica (ver secção 3.2.3 - Análise $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD unido a diferentes ligandos: Determinação da perturbação do desvio químico (CSP)). Dois valores de $\Delta\delta_{comb}$ (0,15 ppm e 0,08 ppm) foram considerados de forma a discriminar entre os aminoácidos que são fortemente ou moderadamente perturbados na presença do ligando. Desta forma os resíduos Y236, G279, H286 e C296, bem como a cadeia lateral do resíduo W271 demonstraram um valor superior de $\Delta\delta_{comb}$

de 0,15 ppm's, ou seja, mostraram uma maior perturbação sentida pela presença do ligando Tn-Thr. No intervalo de $\Delta\delta_{comb}$ superior a 0,08 ppm's, mas inferior a 0,15 ppm's (moderadamente perturbados) foram identificados os resíduos: A235, K264, Q267, D269, W271, G277, G278, E280, C282, Q297 e a H301.

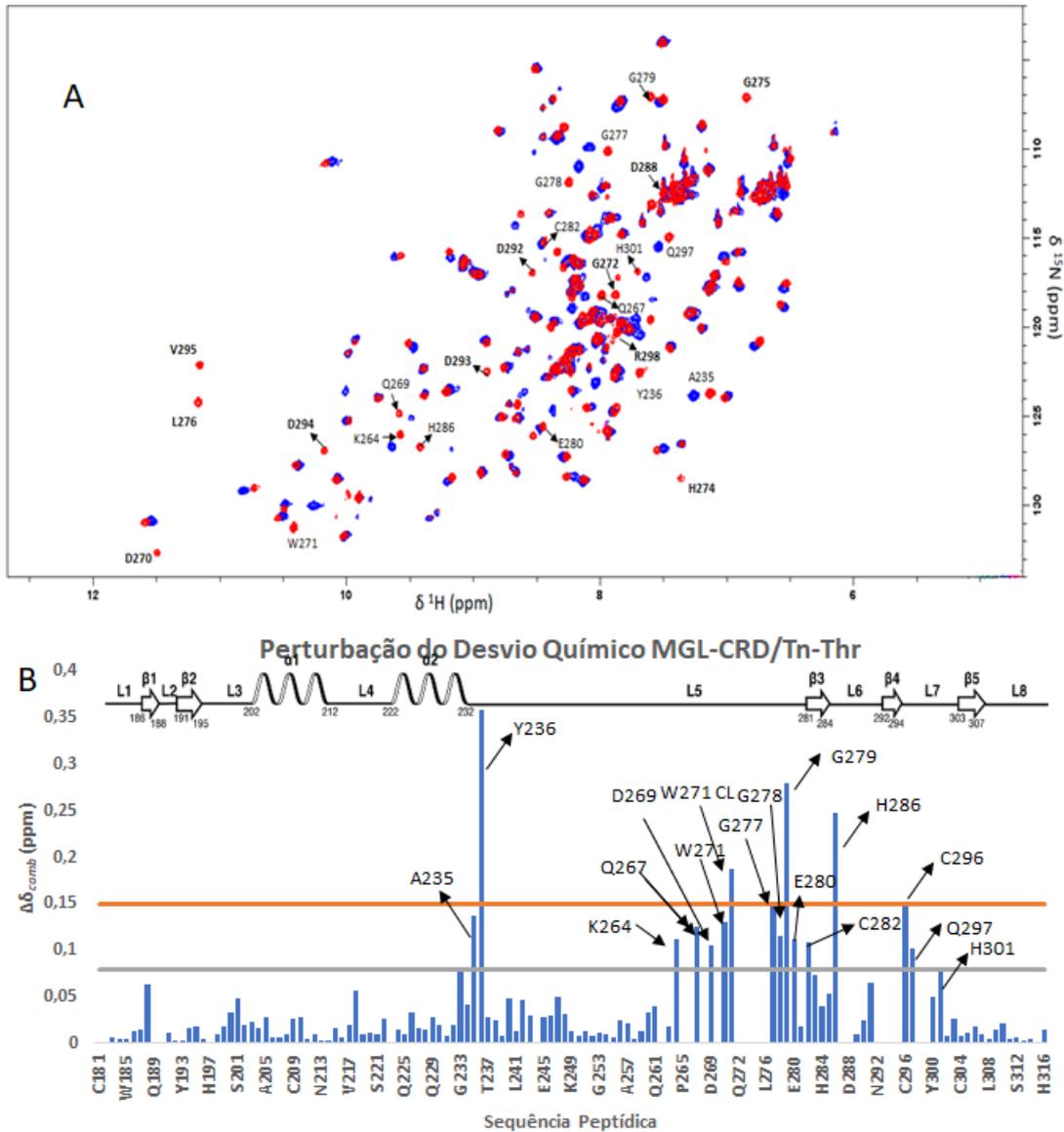


Figura 14: Reconhecimento molecular do antígeno Tn-Thr pela MGL-CRD estudado por $^1\text{H},^{15}\text{N}$ -HSQC. **A-** Sobreposição dos espectros de $^1\text{H},^{15}\text{N}$ -HSQC do CRD MGL na forma livre (azul) e na presença do ligando Tn-Thr (vermelho) adquiridos a 600 MHz e 293 K. Os espectros foram adquiridos com uma concentração de MGL-CRD de 90 μM e de 180 μM de antígeno numa proporção de proteína/ligando 1:2. Assinalados, a negrito, encontram-se os resíduos que surgiram na presença do ligando Tn-Thr. Os resíduos que apresentaram uma perturbação de $\Delta\delta_{comb}$ significativo (> 0.08 ppm) também se encontram identificados no espetro. **B-** Histograma com a análise de CSP na presença do ligando Tn-Thr. A linha a laranja e a linha a cinzento representam os valores de $\Delta\delta_{comb}$ para os resíduos fortemente (> 0.15 ppm) e moderadamente (0.08 ppm) perturbados, respetivamente.

De forma a comparar os $^1\text{H},^{15}\text{N}$ -HSQC do domínio MGL-CRD, na presença de Tn-Thr e α -MeGalNAc, calcularam-se os CSPs entre os dois estados (Figura 15).

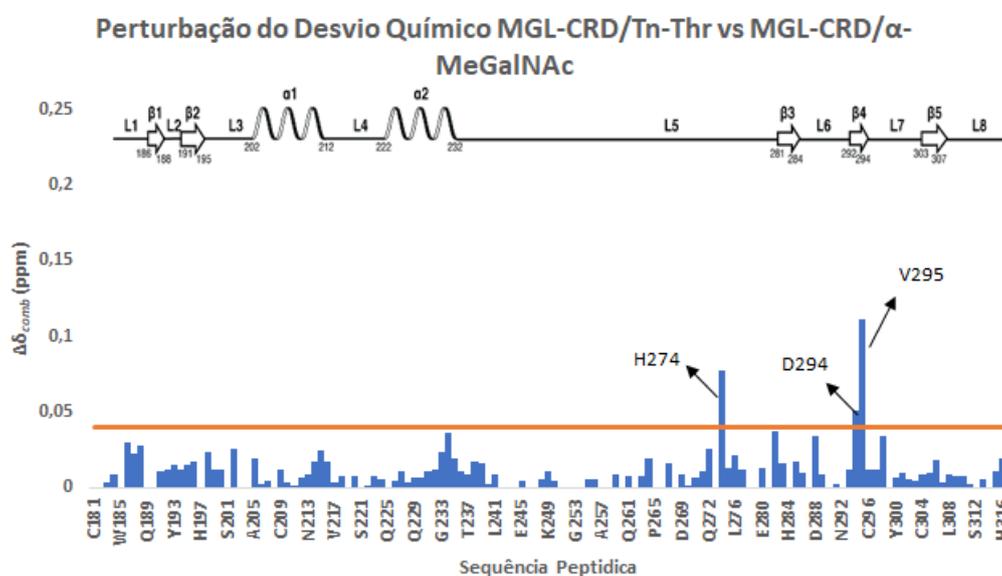


Figura 15: Histograma da análise de CSP do complexo MGL-CRD/Tn-Thr utilizando como referência o complexo MGL-CRD/ α -MeGalNAc. Os resíduos com maior valor de CSP encontram-se assinalados na figura.

Da análise do histograma concluímos que, à exceção dos resíduos H274, D294 e V295, não são detetadas diferenças significativas nos restantes aminoácidos da proteína. Os aminoácidos que se destacam (Figura 16) representam as diferenças no desvio químico detetadas entre as duas formas unidas da lectina. De um modo geral, podemos afirmar que o modo de união do ligando Tn-Thr é semelhante ao α -MeGalNAc, induzindo perturbações semelhantes ao α -MeGalNAc na estrutura do MGL-CRD.

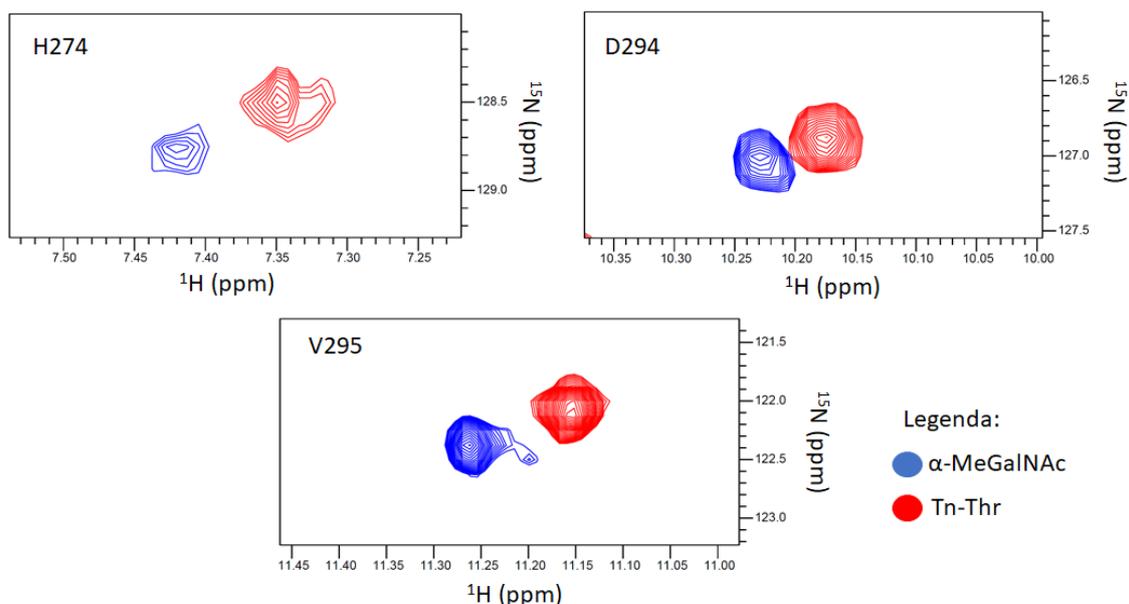


Figura 16: Sobreposição das expansões dos espectros de ^1H , ^{15}N -HSQC da MGL-CRD na presença do ligando α -MeGalNAc (azul) e do ligando Tn-Thr (vermelho) na região dos resíduos H274, D294 e V295.

Para interpretar estes dados experimentais, gerou-se o modelo 3D do complexo MGL-CRD/Tn-Thr por dinâmica molecular, seguindo um protocolo em tudo semelhante ao descrito anteriormente pelo grupo⁴⁵. Este trabalho foi realizado em colaboração com o Prof. Doutor Francisco Corzana da Universidad de La Rioja em Logroño. A figura 17 mostra uma das estruturas obtidas ao longo da dinâmica molecular e onde os resíduos H274, D294 e V295 estão assinalados a cor azul. Efetivamente, de acordo com o modelo, estes resíduos estão espacialmente próximos do aminoácido Thr formando um potencial local para acomodar o aminoácido Thr. A análise dos dados de dinâmica molecular está em curso, contudo podemos especular a formação de uma rede de pontes de hidrogénio formada por esses resíduos e a treonina do ligando.

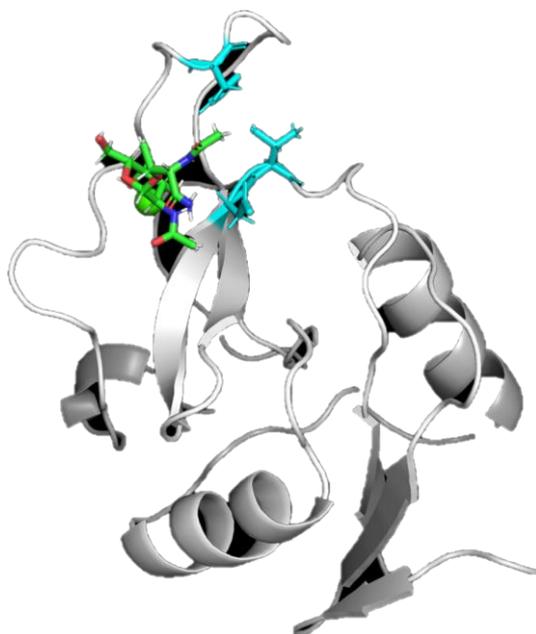


Figura 17: Modelo 3D de homologia do complexo MGL-CRD/Tn-Thr gerado por dinâmica molecular. Os resíduos a azul (H274, D294 e C295) foram identificados pela análise de CSP entre os complexos MGL-CRD/Tn-Thr e MGL-CRD/ α -MeGalNAc. A verde encontra-se representada a estrutura do ligando Tn- Thr.

4.2.2.- MGL-CRD complexado com os monoglicopéptidos MUC1 Tn-T3 e MUC1 Tn-T15

Seguidamente, foram analisados os $^1\text{H},^{15}\text{N}$ -HSQCs do MGL-CRD na presença de dois monoglicopeptidos contendo o antigénio Tn, o monoglicopéptido Tn-T3 e o monoglicopéptido Tn-T15 (Figura 18).



Figura 18: Representação esquemática dos dois monoglicopéptidos Tn-T3 e Tn-T15. O quadrado a amarelo representa o açúcar GalNAc.

Estes dois monoglicopéptidos diferem entre si no local de glicosilação do ligando GalNAc, e como tal, na sequência peptídica à volta do resíduo de GalNAc. Ou seja, são estruturas adequadas para compreender em que medida a sequência peptídica da MUC1 influencia no processo de reconhecimento molecular do antigénio Tn pela lectina MGL. Os espetros de $^1\text{H},^{15}\text{N}$ -HSQC de MGL-CRD, na ausência e na presença de ambos os monoglicopeptidos, encontram-se em anexo (Figura 34-A e 34-B no anexo 7.5). A análise dos CSPs de ambos os glicopéptidos contendo a comparação entre o $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD na forma livre e o $^1\text{H},^{15}\text{N}$ -HSQC do domínio MGL-CRD unido a Tn-T3 ou Tn-T15 é apresentada em anexo na Figura 35-A e 35-B no anexo 7.6.

Posteriormente, foi realizada a comparação dos $^1\text{H},^{15}\text{N}$ -HSQCs do domínio MGL-CRD unido aos glicopéptidos Tn-T3 e Tn-T15 *versus* o $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD unido ao ligando Tn-Thr. Os histogramas obtidos são apresentados na Figura 19 A e B.

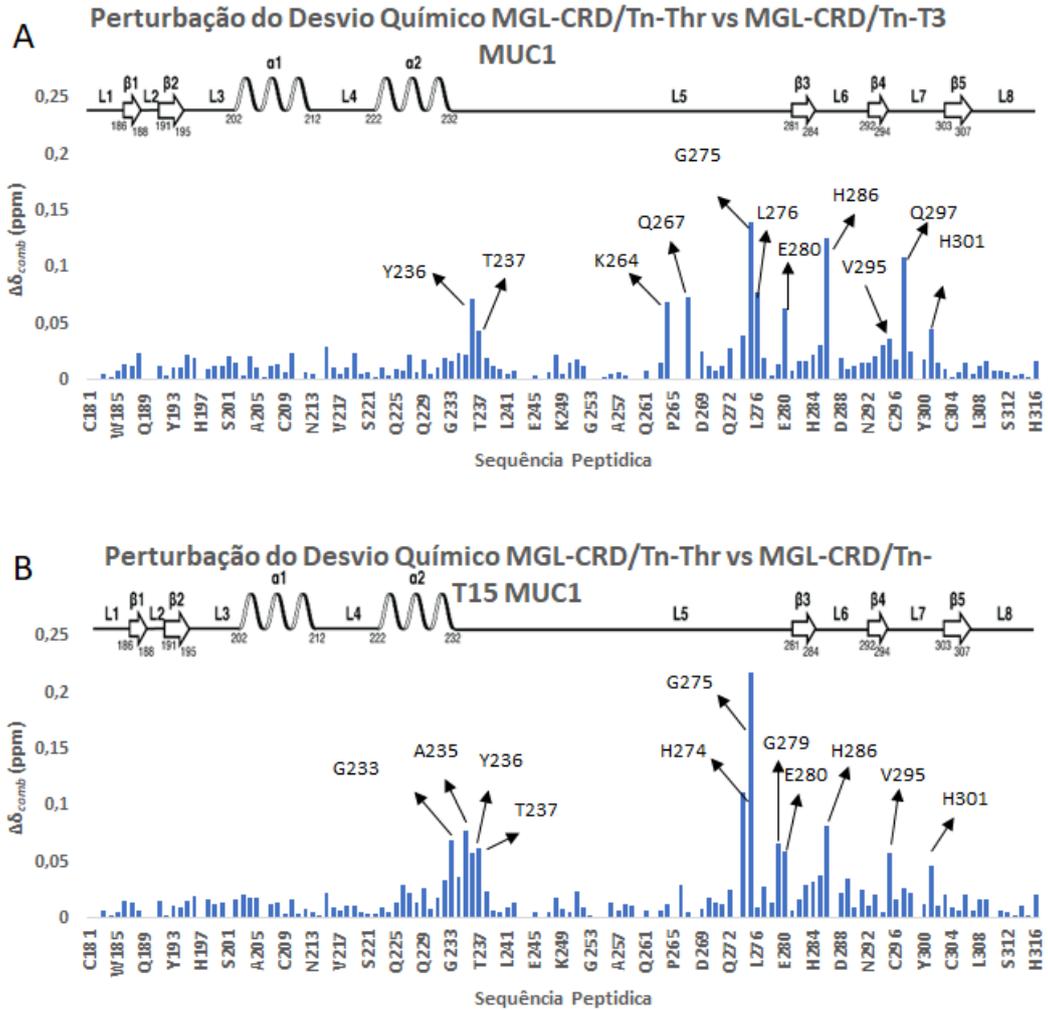


Figura 19: Análise da interação dos monoglicopéptidos Tn-T3 e Tn-T15 com o domínio MGL-CRD. Diferenças com respeito ao complexo MGL-CRD/Tn-Thr. **A-** Histograma da análise de CSP do complexo MGL-CRD/Tn-T3 utilizando como referência o complexo MGL-CRD/Tn-Thr. **B-** Histograma da análise de CSP do complexo MGL-CRD/Tn-T15 utilizando como referência o complexo MGL-CRD/Tn-Thr. O valor de $\Delta\delta_{comb}$ de 0.04 ppm foi considerado para discriminar os resíduos que são perturbados de forma significativa com respeito ao complexo MGL-CRD/Tn-Thr.

Foi considerado um valor de $\Delta\delta_{comb}$ de 0,04 ppm para discriminar os resíduos de MGL-CRD, que devido à presença dos monoglicopéptidos Tn-T3 e Tn-T15, sofriam uma variação significativa no seu desvio químico em relação ao complexo MGL-CRD/Tn-Thr. Os resíduos identificados quando a presença de Tn-T3 foram a Y236, T237, K264, Q267, G275, L276, E280, H286, Q297 e H301. No caso do glicopéptido Tn-T15, foram identificados os seguintes resíduos: G233, A235, Y236, T237, H274, G275, G279, E280, H286, V295 e H301. Desta forma, é evidente que existem diferenças entre os complexos MGL-CRD/Tn-T3 e MGL-CRD/Tn-T15 e complexo MGL-CRD/Tn-Thr em diversos resíduos. Algumas são comuns aos dois glicopéptidos, e outras são exclusivas de cada um dos

glicopéptidos da MUC1. Em comum a ambos os glicopéptidos, existem os resíduos Y236, T237, G275, E280, H286, V295 e H301. Contudo, a amplitude do CSP para alguns destes resíduos é distinta, dependendo do glicopéptido em questão, como é caso particular dos resíduos Y236, E280 e H286.

Além disso, para alguns dos resíduos obtiveram-se valores de CSP distintos, dependendo da localização do GalNAc na sequência peptídica do glicopéptido (Figura 19-A e 19-B). Especificamente, o complexo MGL-CRD/Tn-T3 apresentou diferenças relativas ao complexo MGL-CRD/Tn-Thr nos resíduos K264, Q267, L276 e na Q297 em relação ao ligando Tn-Thr, que não foram detetadas no caso do complexo MGL-CRD/Tn-T15. A Figura 20 mostra a ampliação das ressonâncias alguns destes resíduos, extraídos dos espectros de ^1H , ^{15}N -HSQC, para os três ligandos: Tn-Thr, Tn-T3 e Tn-T15.

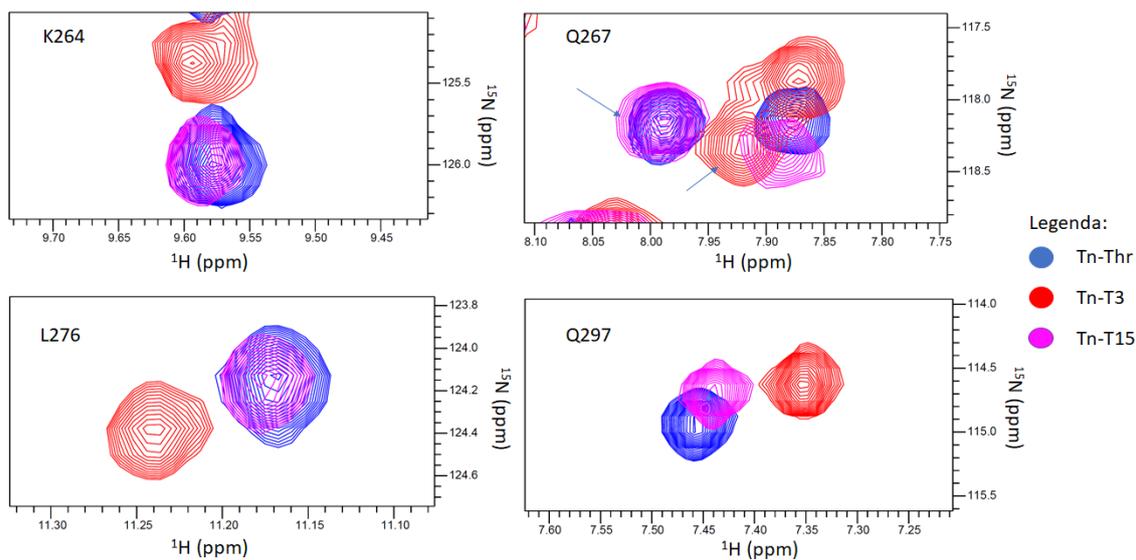


Figura 20: Sobreposição das expansões dos espectros de ^1H , ^{15}N -HSQC do domínio MGL-CRD na presença do Tn-Thr (azul), Tn-T3 (vermelho) e Tn-T15 (roxo) na região dos resíduos K264, Q267, L276, Q297.

O resíduo K264 apresenta no complexo MGL-CRD/Tn-T3 um desvio no desvio químico de ^{15}N para valores inferiores em relação aos complexos MGL-CRD/Tn-Thr e MGL-CRD/Tn-T15. Da mesma forma, no caso dos resíduos Q267 e Q297 existe uma deslocação das respetivas ressonâncias para valores de campo mais altos enquanto para

o resíduo L276 o efeito é oposto, ou seja a ressonância desloca-se para campo alto em comparação com os ligandos Tn-Thr e Tn-T15.

No caso do monoglicopéptido Tn-T15 foram observadas diferenças significativas nos resíduos G233, A235, H274 e G279(Figura 21).

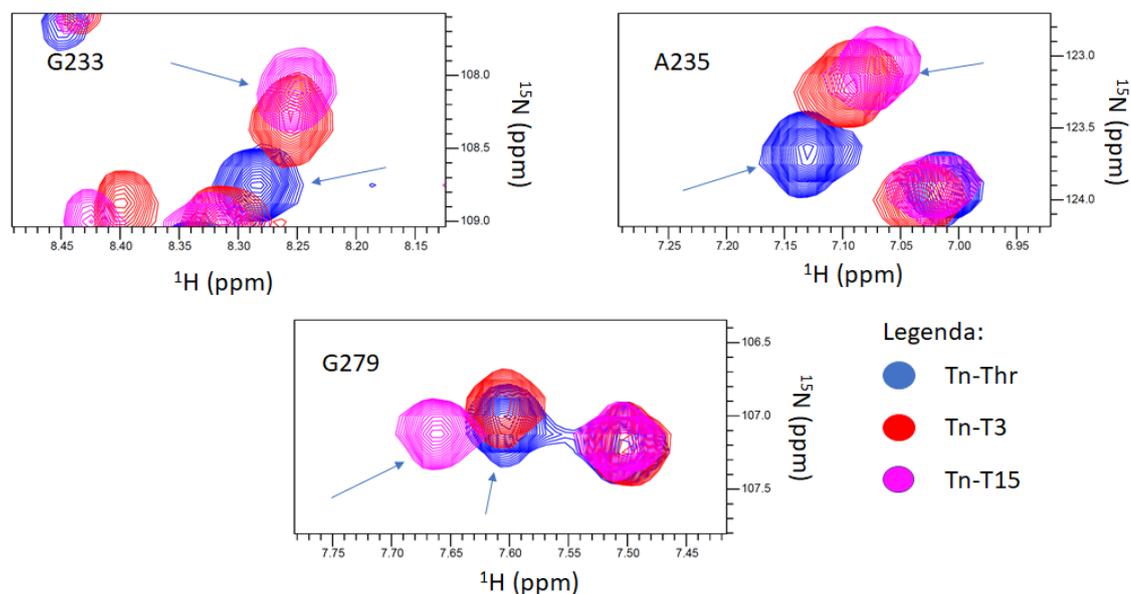


Figura 21: Sobreposição das expansões dos espectros de ^1H , ^{15}N -HSQC do domínio MGL-CRD na presença do Tn-Thr (azul), Tn-T3 (vermelho) e Tn-T15 (roxo) na região dos resíduos G233, A235, G279.

É possível observar nos resíduos G233 e A235 um desvio do sinal para valores de campo mais altos em relação aos complexos MGL-CRD/Tn-Thr e MGL-CRD/Tn-T3 e na direção oposta no caso do resíduo G279.

Por fim, a sobreposição dos dois espectros de ^1H , ^{15}N -HSQC de ambos os monoglicopéptidos Tn-T3 e Tn-T15 (Figura 22), ilustra a sensibilidade de alguns dos aminoácidos à sequência peptídica, como é o caso dos resíduos Y236, G275, E280 e a H286.

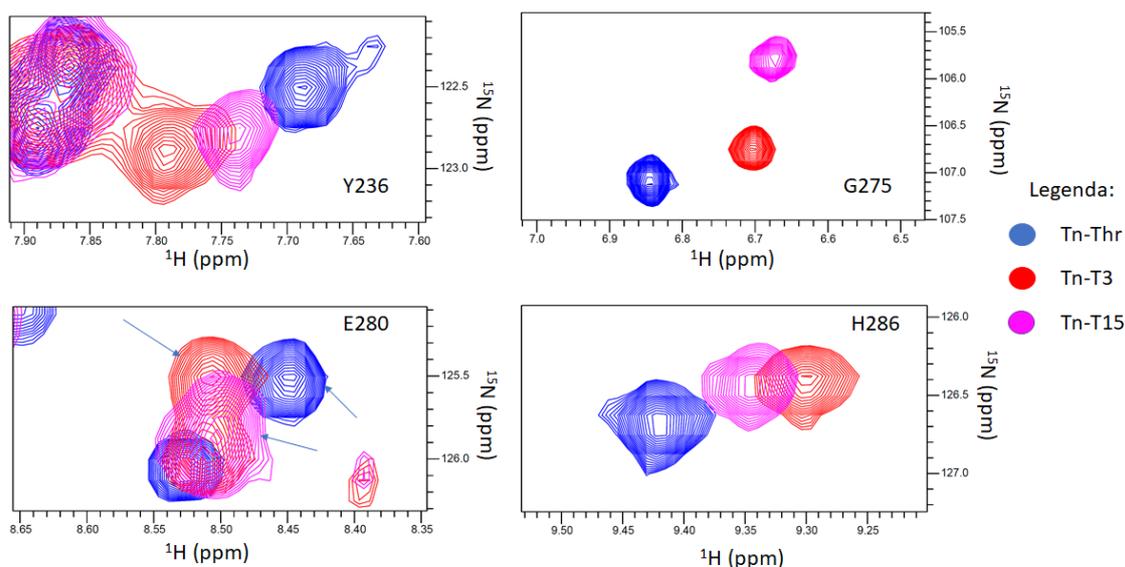


Figura 22: Sobreposição das expansões dos espectros de ^1H , ^{15}N -HSQC do domínio MGL-CRD na presença do Tn-Thr (azul), Tn-T3 (vermelho) e Tn-T15 (roxo) na região dos resíduos Y236, G275, E280 e H286.

Em particular, a relevância do resíduo H286 no reconhecimento molecular de glicopéptidos MUC1 e MUC2 contendo o antígeno Tn foi previamente identificada por microarrays²⁸. Experimentalmente, o mutante H286T mostrou uma menor afinidade para os glicopeptídeos em comparação com a proteína nativa. Simulações de dinâmica molecular revelaram a existência de pontes de hidrogénio entre o resíduo H286 do domínio MGL-CRD e a cadeia peptídica do glicopéptido.

Para ajudar a interpretar estes resultados, geraram-se os modelos 3D dos complexos MGL-CRD/Tn-T3 e MGL-CRD/Tn-T15 seguindo um protocolo em tudo semelhante ao descrito anteriormente pelo grupo⁴⁵. Os modelos dos complexos foram gerados apenas considerando o modo A de união da unidade de GalNAc. O modo B não se revelou estável a partir de 20 ns de dinâmica molecular. Os resíduos afetados na presença de ambos os monoglicopéptidos Tn-T3 e Tn-T15 foram mapeados no modelo 3D do MGL-CRD (Figura 23-A e 23-B). Os aminoácidos afetados, que são comuns a ambos os monoglicopéptidos, estão assinalados a vermelho, enquanto a azul se encontram assinalados os resíduos que são exclusivamente perturbados por cada um dos monoglicopéptidos.

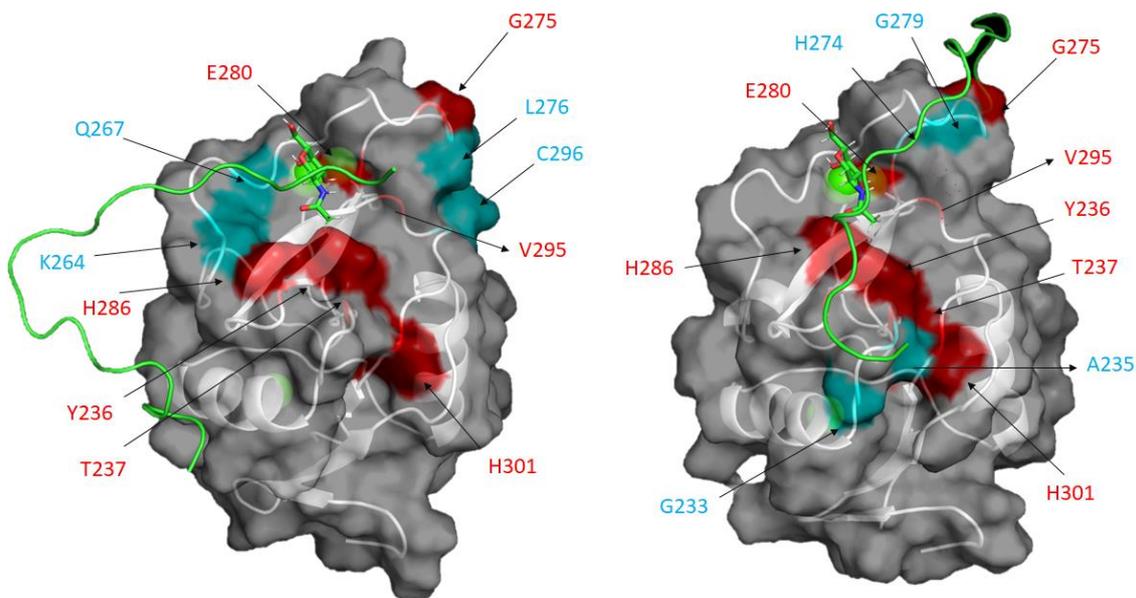


Figura 23: Modelos 3D representativos dos complexos MGL-CRD/MUC1 monoglicopéptidos retirados das dinâmicas moleculares (500 ns) **A-** Modelo do complexo MGL-CRD/Tn-T3. **B-** Modelo do complexo MGL-CRD/Tn-T15. Os resíduos assinalados a vermelho e a azul correspondem aos aminoácidos que apresentam diferenças com respeito ao complexo de referência MGL-CRD/Tn-Thr. Os resíduos assinalados a vermelho correspondem aos aminoácidos que são perturbados em ambos os complexos contendo os monoglicopéptidos Tn-T3 e Tn-T15, enquanto os resíduos assinalados a azul correspondem aos resíduos que são perturbados de forma distinta entre os dois monoglicopéptidos.

Os modelos mostram uma orientação da cadeia peptídica diferente consoante o monoglicopéptido estudado, o que está de acordo com os dados de RMN. Em particular, a orientação distinta da cadeia peptídica explica as diferenças dos resíduos que foram perturbados especificamente dependendo de Tn-T3 (K264, Q267, L276 e Q297) e Tn-T15 (G233, A235, H274 e G279). A análise detalhada da dinâmica molecular destes complexos está em curso e permitirá investigar que tipo de interações (pontes de hidrogénio, CH- π) poderão existir entre a sequência peptídica e o domínio MGL-CRD. De realçar que, as interações estabelecidas entre a proteína e a sequência peptídica dos glicopéptidos certamente contribuíram para a estabilização do complexo lectina-ligando. Por fim, as diferenças nos mapas de interação, consoante o local de glicosilação do antígeno Tn no péptido MUC1, reveladas pelos estudos de RMN e dinâmica molecular, claramente demonstram que a estrutura do domínio MGL-CRD é sensível à sequência peptídica e que a localização do antígeno Tn-Thr na estrutura do péptido MUC1 pode modelar a interação com a lectina.

4.2.3. MGL-CRD complexado com o péptido MUC1 duplamente glicosilado Tn-T3, T15

Por último, estudou-se a interação do domínio MGL-CRD com o péptido da MUC1 duplamente glicosilado na posição Thr3 e Thr15. Com este ligando, pretendíamos averiguar duas hipóteses: A) a preferência por parte da MGL em função da sequência peptídica e B) a eventual possibilidade de dois domínios de MGL-CRD poderem unir-se em simultâneo ao diglicopéptido Tn-T3, T15 (Figura 24-A, B e C). Esta última hipótese é extramente relevante uma vez que: 1) a MGL apresenta-se nas células do sistema imunitário como um trímero (mas onde cada domínio CRD funciona de forma independente), 2) a MUC1 é uma proteína que naturalmente (mesmo em situações de cancro) se encontra glicosilada em diversos locais, 3) a multivalência tanto da MGL como do ligando MUC1 é crucial no contexto da relação estrutura-atividade.

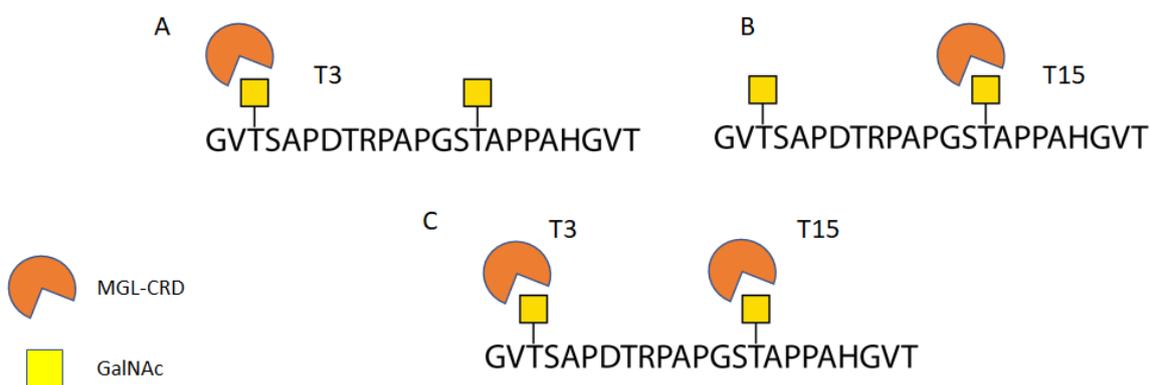


Figura 24: Possíveis modos de ligação do domínio MGL-CRD frente ao diglicopéptido Tn-T3, T15. **A.** Modo de ligação do domínio de MGL-CRD ao GalNAc localizado na Thr 3 **B.** Modo de ligação do domínio de MGL-CRD ao GalNAc localizado na Thr 15. **C.** Modo de ligação onde dois domínios de MGL-CRD se unem em simultâneo ao diglicopéptido Tn-T3, T15. Um dos domínios se une ao GalNAc localizado em Thr3 e outro em Thr15.

Para averiguar a hipótese A) analisamos os espetros de $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD na presença de excesso do péptido diglicosilado Tn-T3, T15 (razão molecular 1:2 proteína: ligando, ou seja, 4 locais de interação por domínio de MGL-CRD). Com este excesso de ligando podemos garantir que em solução vamos ter um modo de união 1:1 de proteína: ligando.

Observando o espectro de $^1\text{H},^{15}\text{N}$ -HSQC do complexo MGL-CRD/Tn-T3, T15 1:2, é possível distinguir os resíduos anteriormente identificados no estudo dos dois monoglicopéptidos Tn-T3 e Tn-T15 (Figura 36 em anexo 7.7). A Figura 25 mostra as ressonâncias de alguns dos aminoácidos onde é claramente observada a existência de duas populações correspondentes aos complexos MGL-CRD/Tn-T3 e MGL-CRD/Tn-T15, os quais parecem existir alternadamente em solução na presença do diglicopéptido MGL-CRD/Tn-T3, T15. Na Figura 25 apenas foram consideradas as ressonâncias que não se encontravam sobrepostas. Em particular os resíduos Y236, T237 e E280 não mostraram dois sinais bem definidos para as duas populações existentes.

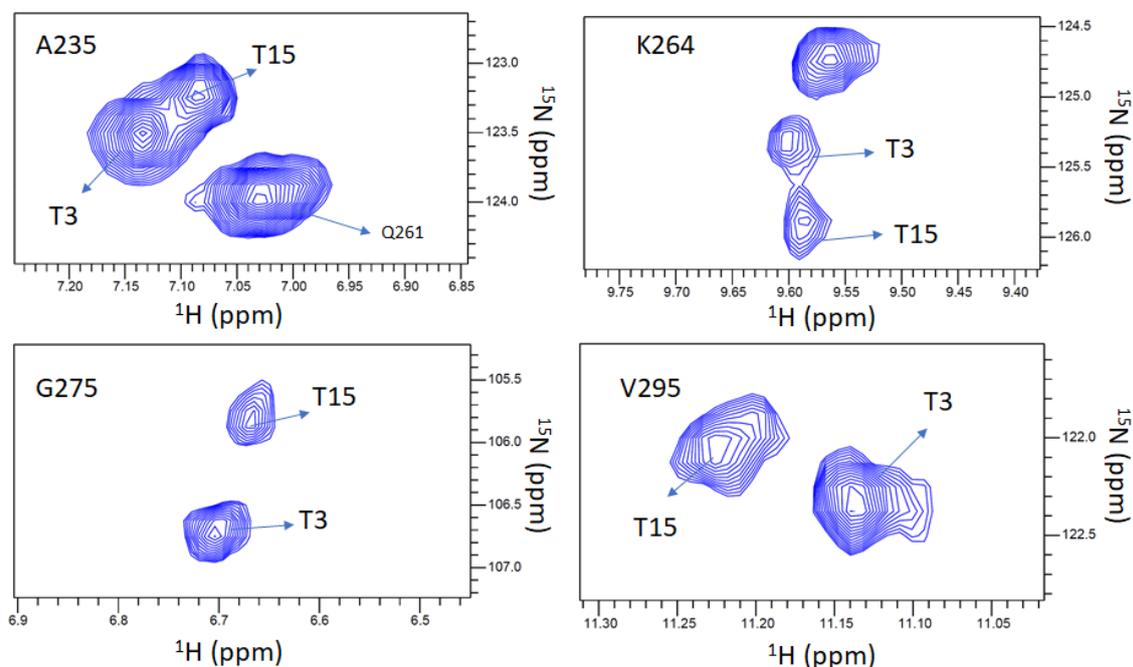


Figura 25: Expansão do espectro de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD na presença do diglicopéptido Tn-T3,T15 na região dos resíduos que apresentam duas ressonâncias correspondentes aos dois modos de união presentes em solução: MGL-CRD unido ao GalNAc localizado na Thr 3 (T3 na figura) e MGL-CRD unido ao GalNAc localizado na Thr15 (T15 na figura).

Com o intuito de perceber se o CRD da MGL tem preferência por alguma das sequências peptídicas, foi realizada a análise das intensidades absolutas das ressonâncias, onde foi possível discriminar a população referente a cada complexo existente em solução (Figura 26). Com base nos valores de intensidade, é possível concluir que a população existente nos dois estados é semelhante para a grande maioria dos resíduos

identificados, o que parece indicar que não existe uma preferência clara da MGL entre as seqüências peptídicas em estudo.

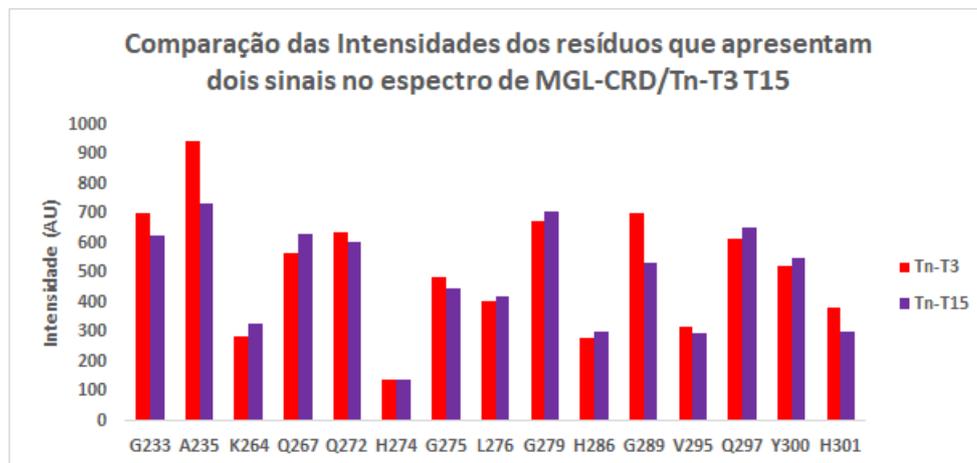


Figura 26: Histograma das intensidades das ressonâncias, detetadas no espectro de $^1\text{H},^{15}\text{N}$ -HSQC de MGL-CRD na presença do diglicopéptido Tn-T3,T15, que possuem dois desvios químicos, os quais correspondem à presença de dois complexos em solução, um correspondente ao complexo MGL-CRD/Tn-T3 (vermelho) e outro correspondente ao complexo MGL-CRD/Tn-T15 (roxo). As ressonâncias no $^1\text{H},^{15}\text{N}$ -HSQC que apresentavam sobreposição de resíduos não foram consideradas na análise.

Relativamente à hipótese B), infelizmente não foi possível validar a coexistência de dois domínios de MGL-CRD unidos ao diglicopéptido Tn-T3, T15. Os dados de $^1\text{H},^{15}\text{N}$ -HSQCs do complexo MGL-CRD/Tn-T3, T15 não nos permitiram extrair informação a esse respeito. Experiências de difusão, em particular DOSY (do inglês diffusion ordered spectroscopy) poderão permitir avaliar a eventual existência deste tipo complexo (2 MGL-CRD: 1 Tn-T3, T15).

4.3.- Estudos de dinâmica molecular do domínio MGL-CRD

A análise prévia dos $^1\text{H},^{15}\text{N}$ -HSQCs do MGL-CRD na forma livre e na forma complexada, com o ligando α -MeGalNAc e distintos ligandos contendo GalNAc em diferentes apresentações, indicaram que a estrutura do domínio de MGL-CRD na forma livre é bastante dinâmica⁴⁵. Na presença de todos os ligandos onze novas ressonâncias (D270, Q272, H274, G275, L276, D288, N292, D293, D294, V295, R298) foram atribuídas no $^1\text{H},^{15}\text{N}$ -HSQC, as quais não foram detetadas na forma livre (Figura 6 na secção 1.12.– Estudos prévios realizados sobre a MGL, os resíduos que surgiram aquando a adição do ligando encontram-se assinalados com *). O não aparecimento destas ressonâncias no espetro de $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD na sua forma livre indica que determinadas regiões da proteína, em particular as regiões localizadas nos *loops* 5, 6 e 7 bem como nas folhas β 3 e β 4 do domínio MGL-CRD assumem várias/distintas conformações na ausência de ligando. A permuta entre estas várias conformações, ocorre num regime intermédio na escala de tempo do desvio químico, razão pela qual estas ressonâncias não aparecem no $^1\text{H},^{15}\text{N}$ -HSQC do domínio MGL-CRD na forma livre (Figura 27-A). Além das ressonâncias que aparecem no espetro $^1\text{H},^{15}\text{N}$ -HSQC, diversas ressonâncias sofreram um aumento significativo da sua intensidade das ressonâncias no $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD na presença dos vários ligandos. A Figura 27-B mostra o histograma da diferença de intensidade do MGL-CRD na presença e na ausência de α -MeGalNAc. A intensidade de um grande número de resíduos do domínio de MGL-CRD, em particular dos resíduos pertencentes aos *loops* 5 e 6 e folhas β 3 e β 4, aumenta significativamente. De notar também que, através da análise das flutuações do $\text{C}\alpha$, durante a dinâmica molecular do modelo 3D da MGL-CRD na forma livre e complexada com o ligando α -MeGalNAc, foi verificado que a união do ligando torna algumas regiões da proteína menos dinâmicas, particularmente uma grande parte do *loop* 5 (Figura 27-C)⁴⁵.

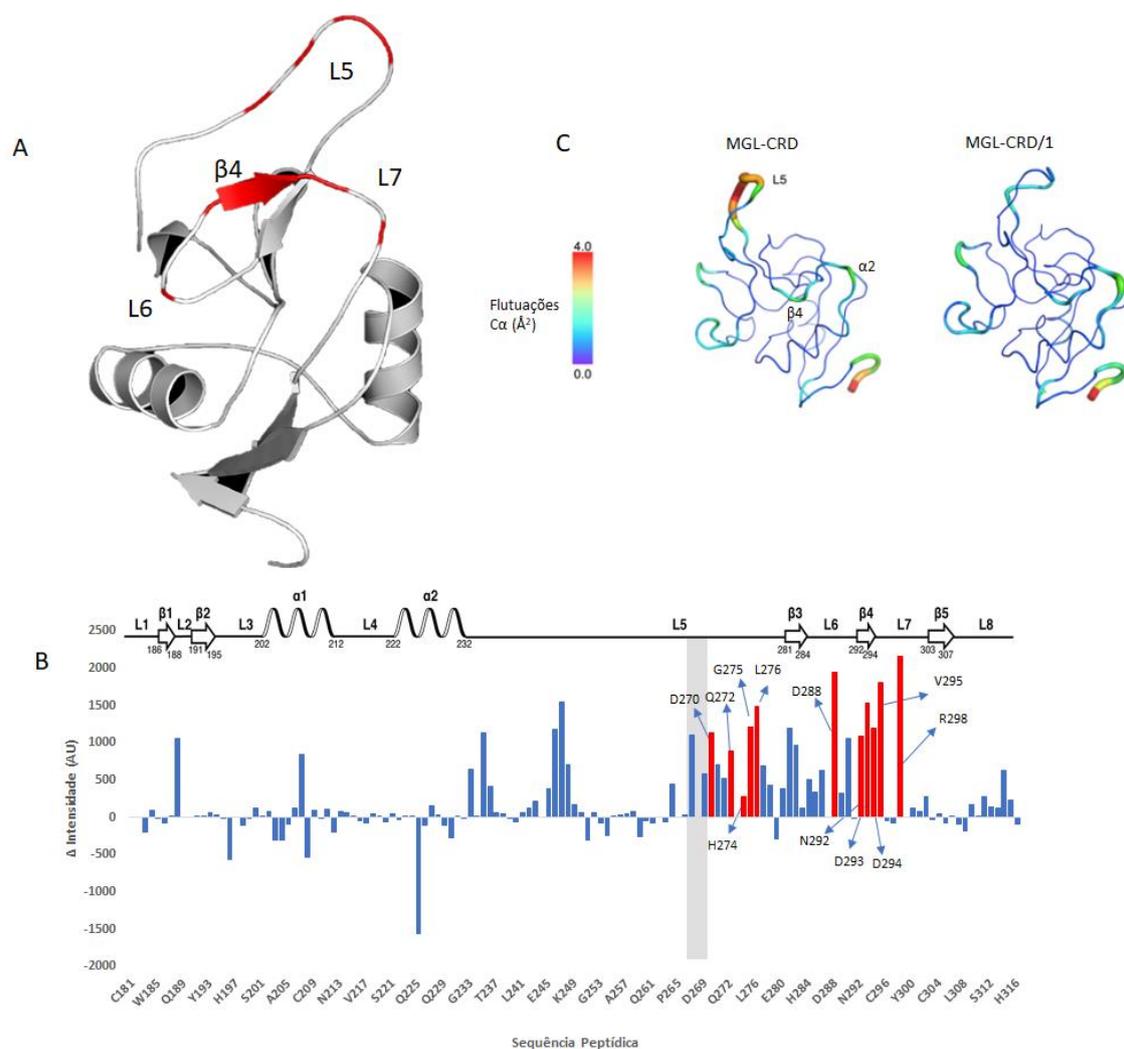


Figura 27: Diferenças de dinâmica entre o domínio de MGL-CRD, na forma livre e na forma complexada com o ligando α -MeGalNAC, deduzidas pela análise do $^1\text{H}, ^{15}\text{N}$ -HSQC e dinâmica molecular. **A-** Modelo 3D de homologia do domínio MGL-CRD. A vermelho encontram-se assinalados os aminoácidos cujas ressonâncias apareceram no espectro de $^1\text{H}, ^{15}\text{N}$ -HSQC na presença dos vários ligandos estudados. **B-** Histograma da diferença das intensidades dos resíduos entre o complexo MGL-CRD/ α -MeGalNAC e o domínio MGL-CRD na forma livre. A vermelho estão assinalados as intensidades correspondentes aos resíduos que surgiram no espectro de $^1\text{H}, ^{15}\text{N}$ -HSQC na forma complexada. **C-** Análise das flutuações do Ca da cadeia principal deduzida por dinâmica molecular do domínio MGL-CRD na forma livre e na presença do complexo MGL-CRD/ α -MeGalNAC. Adaptado de Diniz et al.⁴⁵

Por fim, também se verificou através da análise de CSP, que o domínio do MGL-CRD é sensível à estrutura do ligando que contém a unidade GalNAC e que diferentes interações ocorrem entre o ligando e a proteína, consoante o tipo de ligando que contém a unidade GalNAC⁴⁵ (Figura 9 na secção 1.12.– Estudos prévios realizados sobre a MGL).

Deste modo, na segunda parte do trabalho iniciámos o estudo de dinâmica molecular da cadeia principal do domínio de MGL-CRD por RMN, na ausência e presença de α -MeGalNAc, bem como na presença dos ligandos previamente estudados e ilustrados na Figura 28.

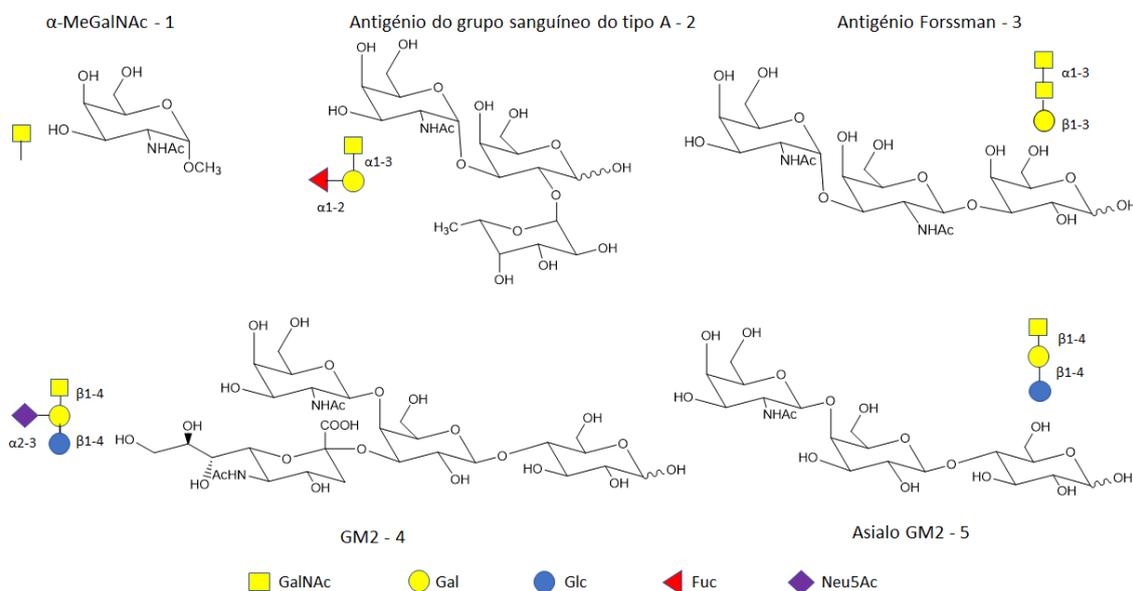


Figura 28: Estruturas químicas representações esquemáticas dos ligandos estudados.

4.3.1 Dados de relaxação globais e cálculo do tempo de correlação

Os dados de relaxação, nomeadamente R_1 , R_2 e HetNOE do MGL-CRD na forma livre e complexado com os ligandos ilustrados na figura 28, foram obtidos através de experiências de $^1\text{H}, ^{15}\text{N}$ -HSQC variando os correspondentes tempos de relaxação e adquiridas a 600MHz a uma temperatura de 293 K. Os valores R_1 , R_2 , HetNOE e razão R_2/R_1 para cada resíduo encontram-se ilustrados nas tabelas 10-13 (anexos 7.8 a 7.11.) e nas Figuras 37-43 (anexo 7.12). A tabela 1 apresenta a média dos valores R_1 , R_2 e HetNOE obtidos e os seus respetivos erros nas diferentes condições, bem como o valor de tempo de correlação (τ_c) estimado a partir da razão R_2/R_1 mas excluindo os valores que falham o critério de seleção⁵¹ (ver secção 3.3-5.-Determinação do tempo de correlação a partir da razão R_2/R_1). O número de resíduos utilizados para o cálculo de cada parâmetro de relaxação e τ_c também se encontra referido na tabela 1.

Tabela 1: Média dos valores de R_1 , R_2 e HetNOE e valor de τ_c estimado do domínio de MGL-CRD obtido nas diferentes condições. Entre parêntesis estão o número total de resíduos usados para o cálculo de cada um dos parâmetros de relaxação e τ_c .

	MGL-CRD	MGL-CRD/1	MGL-CRD/2	MGL-CRD/3	MGL-CRD/4	MGL-CRD/5
R_1 (s^{-1})	1,32 ± 0,10 (109)	1,20 ± 0,08 (120)	1,25 ± 0,24 (120)	1,25 ± 0,10 (119)	1,23 ± 0,14 (120)	1,29 ± 0,18 (120)
R_2 (s^{-1})	12,41 ± 0,39 (110)	11,55 ± 0,54 (120)	11,59 ± 0,44 (120)	11,64 ± 0,47 (119)	12,03 ± 0,50 (120)	11,65 ± 0,37 (119)
HetNOE	0,77 ± 0,07(110)	0,80 ± 0,16 (117)	0,75 ± 0,18 (112)	0,76 ± 0,16 (120)	0,78 ± 0,20 (115)	0,77 ± 0,16 (120)
τ_c (ns)	9,46 ± 0,62 (97)	10,06 ± 0,81(109)	9,45 ± 0,83 (104)	9,42 ± 0,78(102)	10,10 ± 0,95 (102)	9,48 ± 0,63 (106)

De ressaltar que, para o domínio na forma livre, o número de resíduos considerados para determinação de R_1 , R_2 e HetNOE é francamente inferior, comparativamente às restantes condições, uma vez que, um elevado número de ressonâncias não são detetadas na forma livre. Os valores médios de R_1 , R_2 e HetNOE na forma livre do domínio MGL-CRD e na forma complexada com os distintos ligandos são bastante semelhantes entre si.

Os valores de τ_c do domínio MGL-CRD são também idênticos nas diferentes condições experimentais. Este resultado indica que a união dos diferentes ligandos não afeta o movimento global do domínio MGL-CRD em solução, o que poderá estar relacionado com o facto de a massa molecular dos ligandos ser bastante reduzida em comparação com a massa molecular do domínio CRD da proteína.

Por fim, utilizando como regra geral um valor de τ_c calculado com base na massa molecular de uma proteína modelo, é possível estimar com base na massa molecular de 15,7 kDa (estimada pela ferramenta ExPasy⁵³) um valor esperado de τ_c de aproximadamente 8 ns a partir da regra 0.5 ns por 1 kDa⁵⁵. Comparando este valor com os valores de τ_c estimados experimentalmente, pela razão R_2/R_1 , é possível concluir que o $^1H, ^{15}N$ -HSQC da amostra MGL-CRD corresponde a um monómero de MGL-CRD.

4.3.2 Dados de relaxação MGL-CRD vs MGL-CRD/1

Seguidamente, comparou-se os dados de relaxação obtidos para a MGL-CRD na forma livre e do complexo MGL-CRD/1. De forma mais detalhada apenas foram comparados os valores de HetNOE e de R_2 . De notar que, apenas os resíduos com ressonâncias definidas em ambos os estados livre vs unido da MGL-CRD foram considerados.

Os valores de HetNOE e R_2 obtidos para a MGL-CRD na forma livre, bem como para o complexo MGL-CRD/1, encontram-se representados na figura 29-A e 29-B. A figura 29-A e 29-B contêm também a análise do CSP, realizada anteriormente entre as duas formas estudadas. A Figura 29-C mostra a diferença de intensidades absolutas das ressonâncias entre $^1\text{H},^{15}\text{N}$ -HSQC MGL-CRD vs $^1\text{H},^{15}\text{N}$ -HSQC MGL-CRD/1. Os valores de HetNOE demonstraram ser bastante semelhantes entre si, sendo que todas as diferenças observadas na figura 29-A não podem ser consideradas significativas pois encontram-se dentro do erro calculado (Tabela 10 em anexo 7.8). Os valores de HetNOE são sensíveis a movimentos na ordem dos ps a ns em torno do vetor da ligação NH da cadeia peptídica. Desta forma, aparentemente, parecem não existir diferenças, em termos de movimentos locais (ps a ns), em torno da ligação NH da cadeia principal, entre o domínio MGL-CRD livre e o complexo MGL-CRD/1.

Por outro lado, em relação aos valores de R_2 , é possível detetar variações significativas entre os valores apresentados pela forma MGL-CRD livre e os apresentados pelo complexo MGL-CRD/1. Em particular, vários resíduos dos *loops* 5, 6 e 7, e folha β_3 , nomeadamente vários resíduos previamente identificados como envolvidos na interação com o ligando, apresentam valores alterados entre os dois estados da proteína. Especificamente, os valores R_2 são regra geral mais elevados para o domínio MGL-CRD na forma livre do que para a forma complexada com α -MeGalNAc.

Os valores mais elevados de R_2 observados, para o MGL-CRD na forma livre, podem ser explicados com base na mistura de conformações que algumas regiões da proteína adotam. O valor de R_2 é sensível a movimentos locais na ordem dos ps-ns, mas também é afetado por movimentos na ordem dos μs -ms (por exemplo movimento de *loops*). A permuta química causada por alterações conformacionais influencia significativamente os valores de R_2 estimados, e normalmente leva a um aumento do valor estimado de R_2 .

Nestes casos, o valor de R_2 engloba uma grande contribuição da taxa de permuta química entre dois ambientes distintos, ou seja, de R_{ex} .

Desta forma, o facto de na forma livre existirem muitas regiões do domínio MGL-CRD (L5, L6, L7, $\beta 3$) altamente flexíveis, e que apenas na presença de ligando parecem adotar uma conformação mais definida, faz com que o valor estimado de R_2 na forma livre seja superior ao R_2 estimado para o complexo MGL-CRD/1. Os resíduos pertencentes aos *loops* 5, 6 e 7 e $\beta 3$, que apresentam valores elevados de R_2 na forma livre, comparativamente ao complexo MGL-CRD/1, são os mesmos que exibem um aumento das intensidades no $^1H, ^{15}N$ -HSQC quando o MGL-CRD passa da forma livre para a forma complexada (Figura 29-C).

A forma de extrair a contribuição de R_{ex} (permuta conformacional) implica a aquisição dos dados de relaxação a outro campo magnético e a aplicação de um modelo que contemple este parâmetro aquando a análise *Model-free*. A análise, mediante o formalismo do *Model-free*, permitiria então descrever todos os movimentos do domínio MGL-CRD e calcular o parâmetro de ordem S^2 , a permuta conformacional, R_{ex} , e o tempo correlação efetivo. Contudo, este tipo de análise já não fez parte do âmbito desta tese.

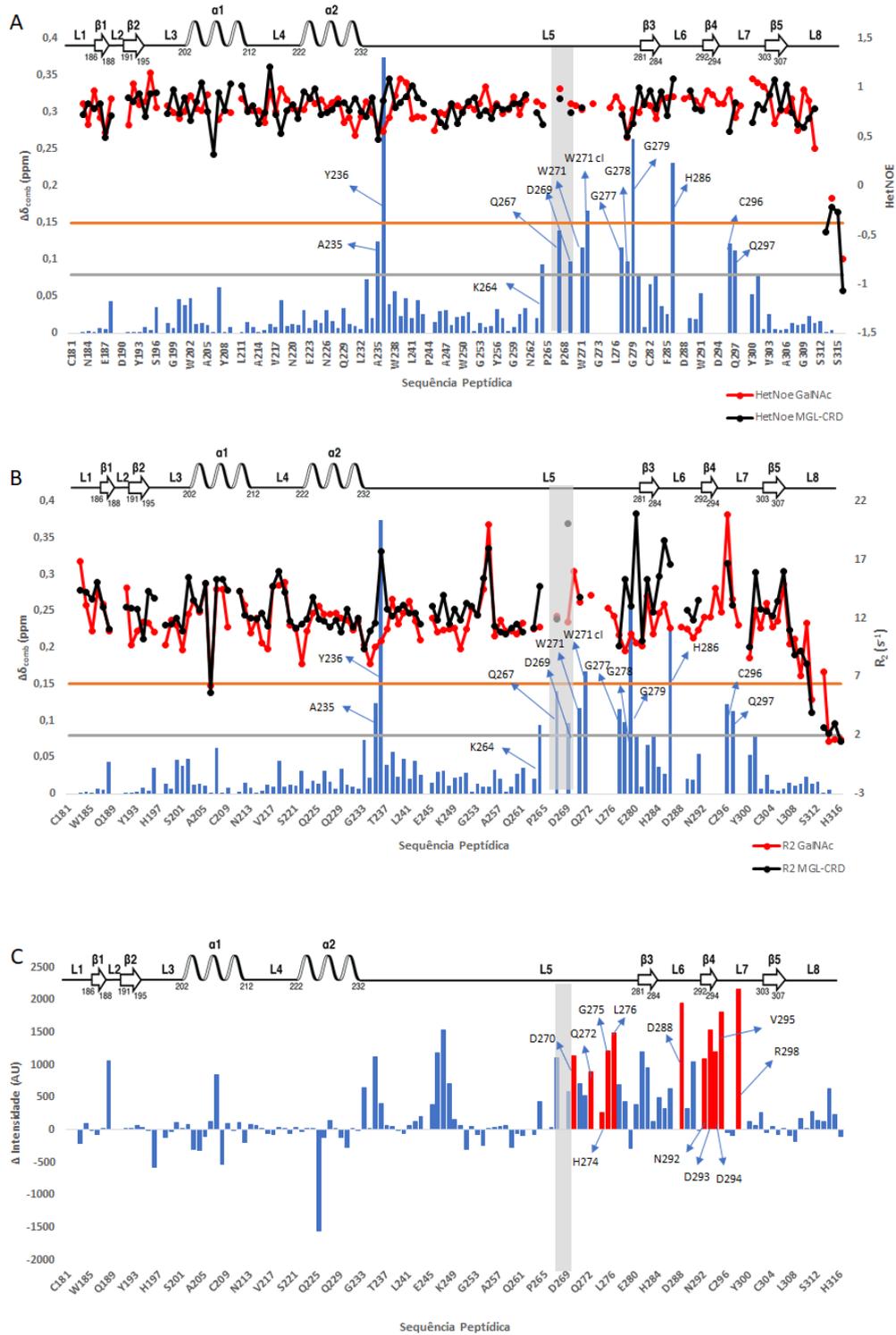


Figura 29: Dados de dinâmica por RMN do domínio MGL-CRD na forma livre e forma complexada com o ligando α -MeGalNAc. **A e B**-Valores de HetNOE e de R_2 , respectivamente, para cada um dos resíduos do domínio MGL-CRD (preto) e do complexo MGL-CRD/1 (vermelho). Encontra-se representado também o histograma da análise do CSP do complexo MGL-CRD/1 em relação ao domínio MGL-CRD. Os resíduos com maiores valores de CSP estão também identificados na figura. **C** - Histograma da diferença das intensidades dos resíduos entre o complexo MGL-CRD/ α -MeGalNAc e o domínio MGL-CRD na forma livre. A vermelho estão assinalados as intensidades correspondentes aos resíduos que surgiram no espectro de $^1\text{H}, ^{15}\text{N}$ -HSQC na forma complexada.

4.3.3 Dados de relaxação do MGL-CRD/1 vs MGL-CRD/2-5

Por fim, comparámos os valores de HetNOE e R_2 do domínio de MGL-CRD na presença dos diferentes ligandos (Figura 30). O complexo MGL-CRD/ α -MeGalNAc foi usado como referência. As Figuras 30 e 31 comparam os valores de HetNOE e R_2 , respetivamente, nas diferentes condições. Cada painel das Figuras 30 e 31 inclui também os valores de CSP calculados entre os $^1\text{H},^{15}\text{N}$ -HSQC dos complexos MGL-CRD/2-5 e $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD/1. Desta forma, procurámos detetar se existia alguma relação entre as alterações observadas no desvio químico de alguns resíduos da proteína, dependendo do ligando, e na dinâmica destes resíduos ou mesmo de outros localizados perto destes. No entanto, analisando os dados contidos na Figura 30, não foram detetadas diferenças significativas entre os diferentes complexos com respeito aos valores de HetNOE.

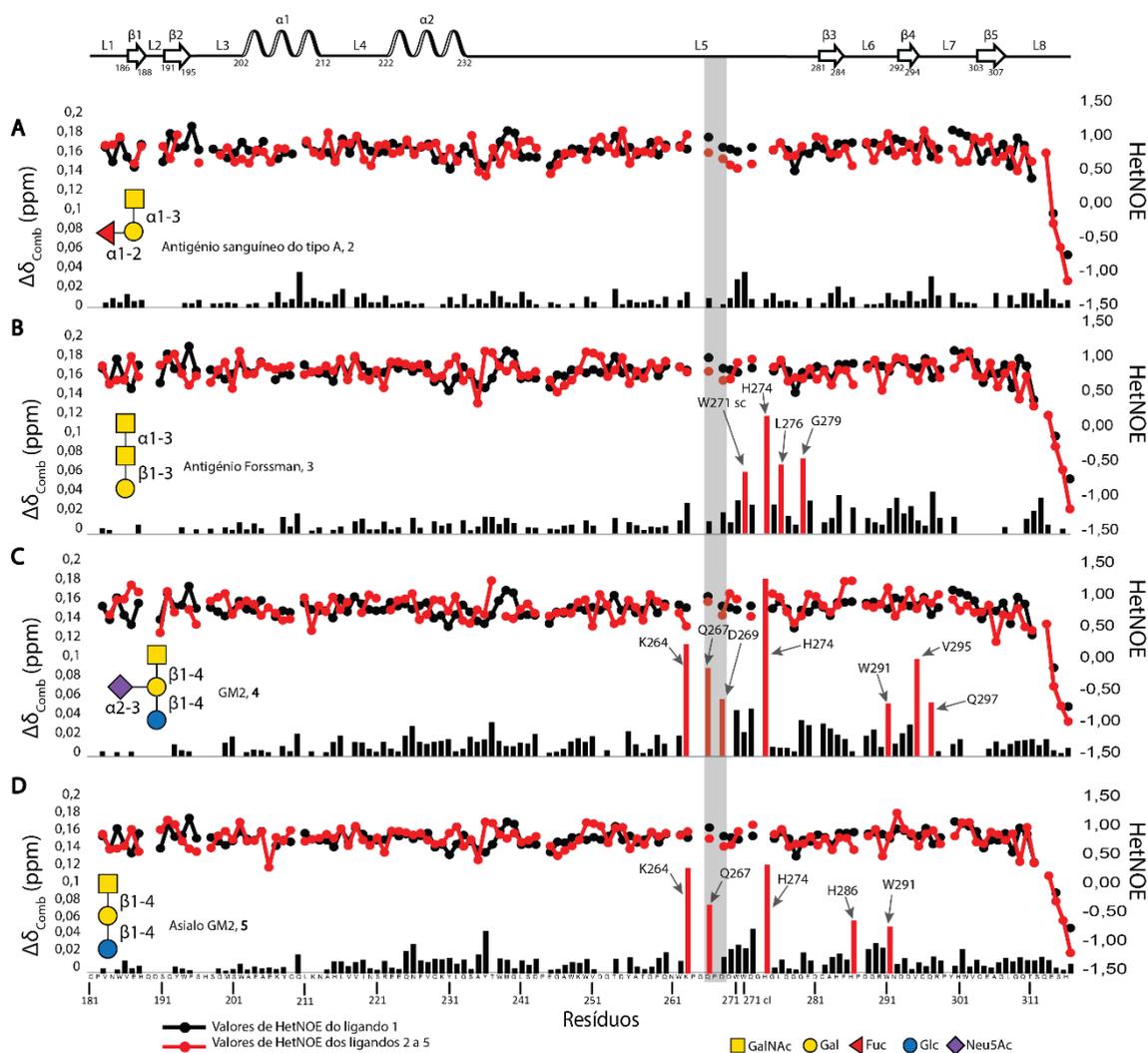


Figura 30: Valores de HetNOE para os diferentes complexos MGL-CRD/2-5 e sua comparação com as alterações no desvio químico de cada complexo com respeito ao complexo de referência MGL-CRD/1. No eixo primário encontram-se representados sob a forma de histograma a perturbação do desvio químico dos diferentes complexos MGL-CRD/2-5 utilizando como referência o complexo MGL-CRD/1. No eixo secundário encontram-se apresentados os valores de HetNOE para cada resíduo e para cada um dos respectivos complexos MGL-CRD/2-5 (vermelho) e os valores de HetNOE para cada resíduo do complexo MGL-CRD/1 (preto) **A**-MGL-CRD/2 **B**-MGL-CRD/3 **C**-MGL-CRD/4 **D**-MGL-CRD/5.

Com respeito aos valores de R_2 , foram detetadas diferenças significativas nos resíduos D270 e C296 (Figura 31). Todos os complexos MGL-CRD/2-5 apresentam, para ambos os resíduos, valores de R_2 inferiores com respeito ao complexo MGL-CRD/1, o que parece indicar um aumento da mobilidade dos resíduos D270 e C296, na presença de ligandos mais complexos que o simples monossacárido α -MeGalNAc. O resíduo D270 localiza-se perto do motivo QPD no *loop* 5. E o resíduo C296 encontra-se localizado no *loop* 3.

Contudo, no geral, não foram encontradas diferenças significativas em R_2 entre os distintos complexos.

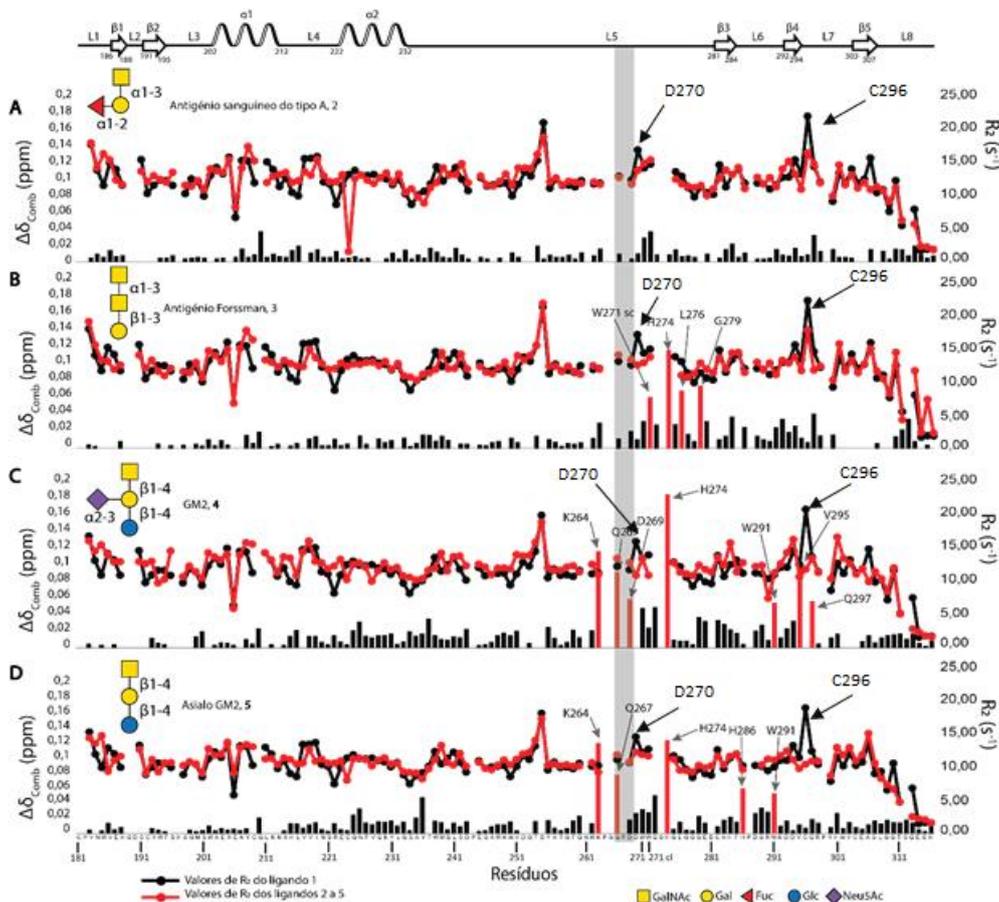


Figura 31: Valores de R_2 para os diferentes complexos MGL-CRD/2-5 e sua comparação com as alterações no desvio químico de cada complexo com respeito ao complexo de referência MGL-CRD/1. No eixo primário encontram-se representados sob a forma de histograma a perturbação do desvio químico dos diferentes complexos MGL-CRD/2-5 utilizando como referência o complexo MGL-CRD/1. No eixo secundário encontram-se apresentados os valores de R_2 para cada resíduo e para cada um dos respectivos complexos MGL-CRD/2-5 (vermelho) e os valores de R_2 para cada resíduo do complexo MGL-CRD/1 (preto) **A-MGL-CRD/2 B-MGL-CRD/3 C-MGL-CRD/4 D-MGL-CRD/5.**

Efetivamente, a análise prévia das flutuações do C α durante a experiência de dinâmica molecular dos complexos MGL-CRD/1-5 (Figura 32) não mostrou diferenças significativas na dinâmica do CRD entre os vários complexos.

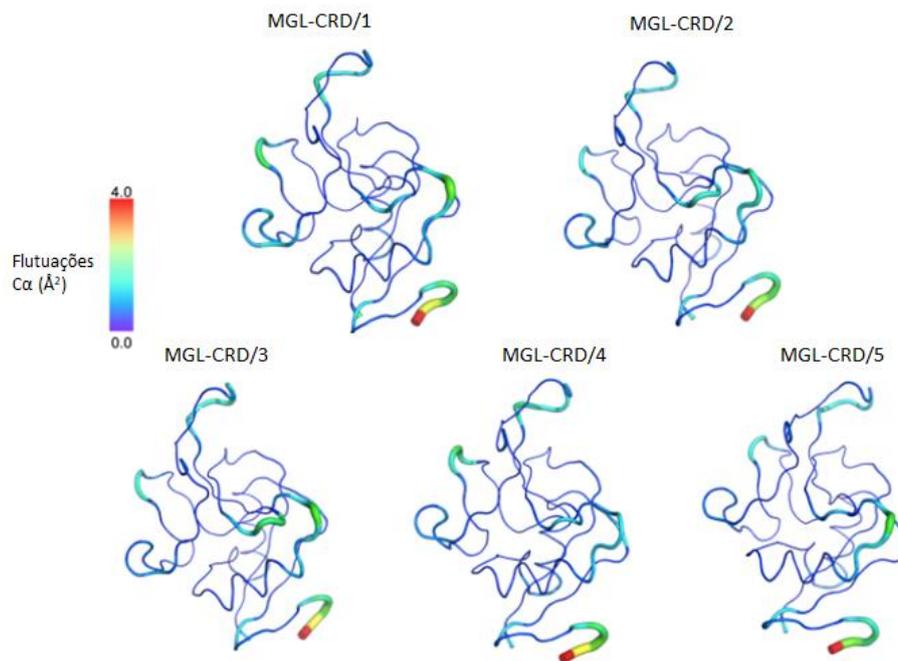


Figura 32: Análise das flutuações do C α da cadeia principal deduzida por dinâmica molecular do complexo MGL-CRD/1 e na presença dos complexos MGL-CRD/2-5. Adaptado de Diniz et al.⁴⁵

5-Conclusões e perspectivas futuras

A interação da lectina MGL com distintos ligandos, contendo a unidade GalNAc, tem sido amplamente estudada devido à sua importância em modelar a resposta imunitária^{23,25,26}. Recentemente, foi demonstrado que o domínio MGL-CRD é bastante dinâmico na sua forma livre e que a sua estrutura altera-se significativamente na presença do ligando⁴⁵. Além disso, foi demonstrado que a estrutura do domínio de MGL-CRD era sensível à estrutura do ligando que continha a unidade GalNAc. Com isto, neste trabalho, levantámos a seguinte hipótese: consoante a estrutura do ligando que contém a unidade GalNAc a estrutura/dinâmica do domínio MGL-CRD é capaz de adotar diferentes conformações, as quais por sua vez definirão que tipo de resposta imunitária será induzida.

Neste contexto, o presente trabalho teve como principal objetivo o estudo do reconhecimento molecular e dinâmica do domínio de MGL-CRD por RMN. Distintos ligandos contendo a unidade principal GalNAc em diferentes modos de apresentação foram estudados.

O estudo do reconhecimento molecular de diferentes monoglicopéptidos, derivados do domínio de repetição da glicoproteína MUC1, contendo o antigénio Tn em distintos locais de glicosilação (Treonina 3 Tn-T3 MUC1 e Treonina 15 Tn-T15 MUC1), mostrou que a sequência peptídica modula o reconhecimento molecular do domínio de MGL-CRD. Em particular, dependendo da glicopéptido em questão, a orientação do péptido face à superfície da lectina é distinta. Com isto, diferentes resíduos do domínio MGL-CRD são perturbados dependendo do tipo de complexo: MGL-CRD/Tn-T3 ou MGL-CRD/Tn-T15. Por outro lado, as experiências de ¹H,¹⁵N-HSQC do MGL-CRD na presença de excesso de diglicopéptido Tn-T3, T15 mostraram que ambos os complexos MGL-CRD/Tn-T3 ou MGL-CRD/Tn-T15 existem de forma alternada na presença de excesso de Tn-T3, T15 em solução. Para uma relação MGL-CRD/Tn-T3, T15 1:2, os complexos MGL-CRD/Tn-T3 e MGL-CRD/Tn-T15 parecem existir numa proporção 1:1 em solução. Este resultado indica que o domínio MGL-CRD não apresenta preferência significativa sobre nenhum dos locais de glicosilação da sequência peptídica da MUC1 estudados. Estudos de RMN e de dinâmica molecular estão em curso para investigar se duas ou mais

unidades de MGL-CRD poderão estar envolvidas em simultâneo no reconhecimento de estruturas de péptidos derivados da MUC1 com distinta densidade de antígeno Tn e com mais do que uma unidade de repetição. Desta forma, pretender-se-á averiguar o reconhecimento da MGL-CRD usando estruturas mais semelhantes às que se encontram nas células cancerígenas.

Por outro lado, os estudos de dinâmica molecular por ^{15}N HetNOE, R_1 e R_2 da cadeia principal do domínio MGL-CRD, na ausência e na presença de distintos ligandos contendo a unidade GalNAc (α -MeGalNAc, antígeno do grupo sanguíneo do tipo A, antígeno Forssman e estruturas derivadas de glicosíngolípido tais como o GM2 e asialo GM2), mostraram diferenças significativas apenas nos valores de R_2 de alguns resíduos e entre a forma livre do domínio MGL-CRD e a forma complexada. Concretamente, os valores de R_2 dos aminoácidos pertencentes aos *loops* 5 e 6 e folha β_3 na ausência de ligando são maiores na forma livre do que na forma complexada. Este resultado é explicado se considerarmos a contribuição devida à permuta conformacional (R_{ex}). A contribuição do parâmetro R_{ex} na dinâmica do domínio MGL-CRD é relevante para o valor de R_2 estimado. Como tal estudos de dinâmica a outro campo magnético terão de se realizar para ser possível determinar a dinâmica local sem contribuição de R_{ex} .

Entre as várias formas complexadas, não foram detetadas diferenças significativas em termos de dinâmica da proteína. Apenas os resíduos D270 e C296 do domínio MGL-CRD complexado com os ligandos: antígeno do grupo sanguíneo do tipo A, antígeno Forssman e estruturas derivadas de glicosíngolípido tais como o GM2 e asialo GM2, sofreram uma diminuição do valor de R_2 com respeito ao complexo MGL-CRD/ α -MeGalNAc. Esta diminuição indica um aumento da mobilidade destes dois resíduos na presença dos ligandos mais complexos.

Por fim, o valor de τ_c da MGL-CRD na ausência e na presença dos diferentes ligandos foi calculado com base na razão R_2/R_1 . Os valores obtidos foram muito idênticos para todos os complexos e também para o domínio MGL-CRD na sua forma livre, indicando que a MGL-CRD é um monómero em solução e que não existem dímeros em solução.

6-Bibliografia

- 1 D. D. Chaplin, *J. Allergy Clin. Immunol.*, 2010, **125**, 1–41.
- 2 I. Mackay and S. R. Rosen, *N. Engl. J. Med.*, 2010, **343**, 1132–1132.
- 3 K. B. Elkon and J. J. Rhiannon, *Innate immunity*, Elsevier Inc., Tenth Edit., 2012.
- 4 J. Parkin and B. Cohen, *Lancet - Immunol.*, 2001, **357**, 1777–1789.
- 5 I. Mackay and S. R. Rosen, *N. Engl. J. Med.*, 2001, **344**, 1140–1144.
- 6 S. Gonzalez, A. P. González-Rodríguez, B. Suárez-Álvarez, A. López-Soto, L. Huergo-Zapico and C. Lopez-Larrea, *Self/NonSelf - Landes Biosci.*, 2011, **2**, 19–25.
- 7 Y. van Kooyk and G. A. Rabinovich, *Nat. Immunol.*, 2008, **9**, 593–601.
- 8 H. Gonzalez, C. Hagerling and Z. Werb, *Genes Dev.*, 2018, **32**, 1267–1284.
- 9 K. Ohtsubo and J. D. Marth, *Cell*, 2006, **126**, 855–867.
- 10 Spiro R.G, *Glycobiology*, 2002, **12**, 43R-56R.
- 11 R. Easton, *Life Sci. - Tech. Bull.*, 2011, **1**, 1–5.
- 12 P. Stanley, *Cold Spring Harb. Perspect. Biol.*, 2011, **3**, 1–13.
- 13 E. Rodríguez, S. T. T. Schetters and Y. Van Kooyk, *Nat. Rev. Immunol.*, 2018, **18**, 204–211.
- 14 J. Chia, G. Goh and F. Bard, *Biochim. Biophys. Acta - Gen. Subj.*, 2016, **1860**, 1623–1639.
- 15 L. A. M. Cornelissen and S. J. Van Vliet, *Biomolecules*, 2016, **6**, 1–19.
- 16 J. L. Daniotti, A. A. Vilcaes, V. T. Demichelis, F. M. Ruggiero and M. Rodriguez-Walker, *Front. Oncol.*, 2013, **3**, 1–12.
- 17 S. M. Candeias and U. S. Gaipl, *Anticancer. Agents Med. Chem.*, 2016, **16**, 101–7.
- 18 S. R. Stowell, T. Ju and R. D. Cummings, *Annu. Rev. Pathol.*, 2016, **25**, 289–313.
- 19 T. Angata, C. M. Nycholat and M. S. Macauley, *Trends Pharmacol. Sci.*, 2015, **36**, 645–660.
- 20 S. Mayer, M. K. Raulf and B. Lepenies, *Histochem. Cell Biol.*, 2017, **147**, 223–237.
- 21 D. R. McDonald and O. Levy, *Principles of Immune Response*, Elsevier Ltd, Fifth Edit., 2018.
- 22 W. I. Weis, M. E. Taylor and K. Drickamer, *Immunol. Rev.*, 1998, **163**, 19–34.
- 23 Y. van Kooyk, J. M. Ilarregui and S. J. van Vliet, *Immunobiology*, 2015, **220**, 185–192.
- 24 N. Higashi, K. Fujioka, K. Denda-Nagai, S. I. Hashimoto, S. Nagai, T. Sato, Y. Fujita, A. Morikawa, M. Tsuiji, M. Miyata-Takeuchi, Y. Sano, N. Suzuki, K. Yamamoto, K. Matsushima and T. Irimura, *J. Biol. Chem.*, 2002, **277**, 20686–20693.
- 25 N. Mortezaei, H. N. Behnken, A. K. Kurze, P. Ludewig, F. Buck, B. Meyer and C.

- Wagener, *Glycobiology*, 2013, **23**, 844–852.
- 26 E. Saeland, S. J. Van Vliet, M. Bäckström, V. C. M. Van Den Berg, T. B. H. Geijtenbeek, G. A. Meijer and Y. Van Kooyk, *Cancer Immunol. Immunother.*, 2007, **56**, 1225–1236.
- 27 S. J. van Vliet, L. C. Paessens, V. C. M. Broks-van den Berg, T. B. H. Geijtenbeek and Y. van Kooyk, *J. Immunol.*, 2008, **181**, 3148–3155.
- 28 F. Marcelo, N. Supekar, F. Corzana, J. C. Van Der Horst, I. M. Vuist, D. Live, G. J. P. H. Boons, D. F. Smith and S. J. Van Vliet, *J. Biol. Chem.*, 2019, **294**, 1300–1311.
- 29 K. Drickamer and M. E. Taylor, *Curr. Opin. Struct. Biol.*, 2015, **34**, 26–34.
- 30 J. Tanaka, A. S. Gleinich, Q. Zhang, R. Whitfield, K. Kempe, D. M. Haddleton, T. P. Davis, S. Perrier, D. A. Mitchell and P. Wilson, *Biomacromolecules*, 2017, **18**, 1624–1633.
- 31 S. J. van Vliet, C. A. Aarnoudse, V. C. M. Broks-van den Nerg, M. Boks, T. B. H. Geijtenbeek and Y. van Kooyk, *Eur. J. Immunol.*, 2007, **37**, 2075–2081.
- 32 N. Suzuki, K. Yamamoto, S. Toyoshima, T. Osawa and T. Irimura, *J. Immunol.*, 1996, **156**, 128–135.
- 33 M. del Carmen Fernandez-Alonso, D. Diaz, M. Alvaro Berbis, F. Marcelo, J. Canada and J. Jimenez-Barbero, *Curr. Protein Pept. Sci.*, 2013, **13**, 816–830.
- 34 J. M. Pröpster, F. Yang, S. Rabbani, B. Ernst, F. H.-T. Allain and M. Schubert, *Proc. Natl. Acad. Sci.*, 2016, **113**, E4170–E4179.
- 35 W. Feng, L. F. Pan and M. J. Zhang, *Sci. China Life Sci.*, 2011, **54**, 101–111.
- 36 P. Valverde, J. I. Quintana, J. I. Santos, A. Ardá and J. Jiménez-Barbero, *ACS Omega*, 2019, **4**, 13618–1630.
- 37 A. L. Carvalho, T. Santos-Silva, M. J. Romao, E. J. Cabrita and F. Marcelo, *Essent. Tech. Med. Life Sci.*, 2018, 30–91.
- 38 V. Breukels, A. Konijnenberg, S. M. Nabuurs, J. F. Doreleijers, N. V. Kovalevskaya and G. W. Vuister, *Curr. Protoc. Protein Sci.*, 2011, **64**, 1–44.
- 39 M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2013, **73**, 1–16.
- 40 V. A. Jarymowycz and M. J. Stone, *Chem. Rev.*, 2006, **106**, 1624–1671.
- 41 F.-A. Chao and R. A. Byrd, *Emerg. Top. Life Sci.*, 2018, **2**, 93–105.
- 42 M. Kovermann, P. Rogne and M. Wolf-Watz, *Q. Rev. Biophys.*, 2016, **49**, 1–43.
- 43 S. C. Sahu, A. K. Bhuyan, A. Majumdar and J. B. Udgaonkar, *Proteins Struct. Funct. Genet.*, 2000, **41**, 460–474.
- 44 F. Marcelo, F. Garcia-Martin, T. Matsushita, J. Sardinha, H. Coelho, A. Oude-Vrielink, C. Koller, S. André, E. J. Cabrita, H. J. Gabius, S. I. Nishimura, J. Jiménez-Barbero and F. J. Cañada, *Chem. - A Eur. J.*, 2014, **20**, 16147–16155.
- 45 A. Diniz, H. Coelho, J. S. Dias, S. J. van Vliet, J. Jiménez-Barbero, F. Corzana, E. J. Cabrita and F. Marcelo, *Chem. - A Eur. J.*, , DOI:10.1002/chem.201902780.

- 46 M. Meier, M. D. Bider, V. N. Malashkevich, M. Spiess and P. Burkhard, *J. Mol. Biol.*, 2000, **300**, 857–865.
- 47 R. Keller, *The computer aided resonance assignment tutorial*, CANTINA Verlag, Zürich, First edit., 2004.
- 48 W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue, *Proteins Struct. Funct. Genet.*, 2005, **59**, 687–696.
- 49 M. Bieri, E. J. D’Auvergne and P. R. Gooley, *J. Biomol. NMR*, 2011, **50**, 147–155.
- 50 T. D. W. Claridge, *High-Resolution NMR Techniques in Organic Chemistry*, Elsevier, Oxford, Second Edi., 2009, vol. 122.
- 51 N. Tjandra, A. Bax, S. E. Feller and R. W. Pastor, *J. Am. Chem. Soc.*, 1995, **117**, 12562–12566.
- 52 J. Pohleven, B. Trukelj and J. Kos, *Affin. Chromatogr.*, , DOI:10.5772/36578.
- 53 E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch, *Proteomics Protoc. Handb. Protein Identif. Anal. Tools ExpASY Serv.*, 2005, 571–607.
- 54 C. Napoletano, A. Rughetti, M. P. Agervig Tarp, J. Coleman, E. P. Bennett, G. Picco, P. Sale, K. Denda-Nagai, T. Irimura, U. Mandel, H. Clausen, L. Frati, J. Taylor-Papadimitriou, J. Burchell and M. Nuti, *Cancer Res.*, 2007, **67**, 8358–8367.
- 55 F. Probert, S. B. M. Whittaker, M. Crispin, D. A. Mitchell and A. M. Dixon, *J. Biol. Chem.*, 2013, **288**, 22745–22757.

7-Anexos

7.1-Sequência Peptídica e vetor de expressão da hMGL-CRD

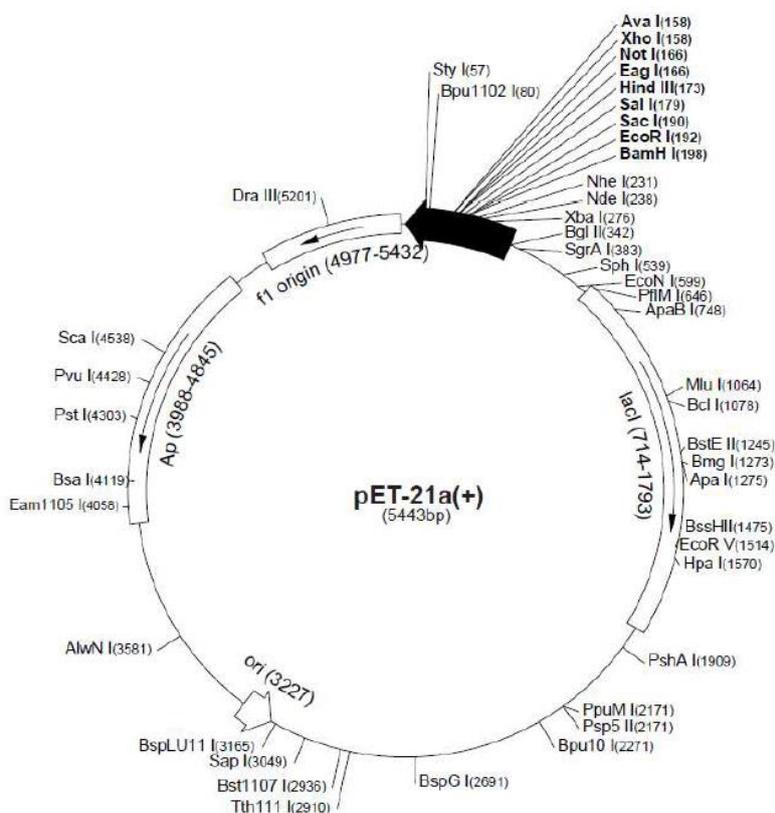


Figura 33: Vetor de Expressão pET21a contendo o gene do domínio MGL-CDR, contém 5443 pares de bases e resistência a ampicilina.

Sequência Peptídica do domínio MGL-CRD:

MCPVNWVEHQDSCYWFSHSGMSWAEAEKYCQLKNAHLVVINSREEQNFVQKYLGSAYTWMG
LSDPEGAWKWVDGTDYATGFQNWKPGQPDDWQGHGLGGGEDCAHFHPDGRWNDDVCQR
PYHWVCEAGLGQTSQESH

7.2-Soluções utilizadas na expressão da MGL-CRD uniformemente marcada com ¹⁵N

Tabela 2- Composição do meio LB.

Solução	Reagentes	Concentração
Meio LB	Extrato Levedura	5g/L
	Triptona	10g/L
	NaCl	10g/L

Tabela 3- Composição do meio LB Agar.

Solução	Reagentes	Concentração
Meio LB Agar	Extrato Levedura	5g/L
	Triptona	10g/L
	Agar	15g/L
	NaCl	10g/L

Tabela 4- Composição do meio LB Agar. Os valores indicados são para a preparação de 1L.

Soluções	Reagentes	Concentração (1 L)
10x sais M9 pH=7.5	NaHPO ₄ .7H ₂ O	60g/L
	KH ₂ PO ₄	30g/L
	NaCl	5g/L
Meio M9	MgSO ₄	2M (1mL)
	CaCl ₂	0,1M (1mL)
	Glucose	4g/L
	Tiamina-HCl	10g/L (1mL)
	FeSO ₄	0,1M (1mL)
	Ampicilina	100mg/mL (1mL)
	¹⁵ NH ₄ ⁺ Cl (¹⁵ N 99% <i>Cambridge Isotope Laboratories</i>)	1g/L

7.3-Soluções utilizadas na técnica de SDS-PAGE

Tabela 5-Preparação de 5ml da solução de gel de empacotamento para SDS-PAGE.

	Reagentes	Volume (mL)
Gel de Empacotamento	H ₂ O	3,9
	3M Tris-HCl SDS pH=8.45	1,55
	Acrilamida 30%	0,8
	APS 30%	0,042
	TEMED	0,014

Tabela 6- Preparação de 15ml da solução de gel de separação para SDS-PAGE.

Gel de Separação	Reagentes	Volume (mL)
	H ₂ O	3,5
	3M Tris-HCl SDS pH=8.45	5
	Acrilamida 30%	5
	Glicerol	1,5
	APS 30%	0,042
	TEMED	0,014

Tabela 7- Preparação de 50mL de solução de *sample buffer* para SDS-PAGE.

<i>Sample Buffer</i>	Reagentes	Quantidade
	1M Tris-HCl pH=6.8	5mL
	Glicerol	12mL
	SDS	4g
	DTT	1,55g
	<i>Blue Coomassie R250</i>	10mg

Tabela 8- Preparação de 1L da solução tampão 10x Ânodo 2M Tris

10X Tampão Ânodo 2M Tris pH=8,80	Reagentes	Quantidade (g)
	Tris	242

Tabela 9- Preparação de 1L da solução tampão Tris-Tricina.

10X Tris-Tricina SDS pH=8.3	Reagentes	Quantidade
	Tris	121,1g
	Tricina	179,2
SDS 20%	50mL	

7.4- Dados da MGL-CRD calculados pela ferramenta ExPASy

Número de Aminoácidos da sequência: 137

Massa Molecular: 15 726,15 Daltons

Ponto Isoelétrico teórico: 4,90

Coefficiente de Extinção Molar (280nm): 63 940 M⁻¹c

7.5- Espectros de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD na ausência e na presença dos monoglicopéptidos Tn-T3 e Tn-T15 estudados

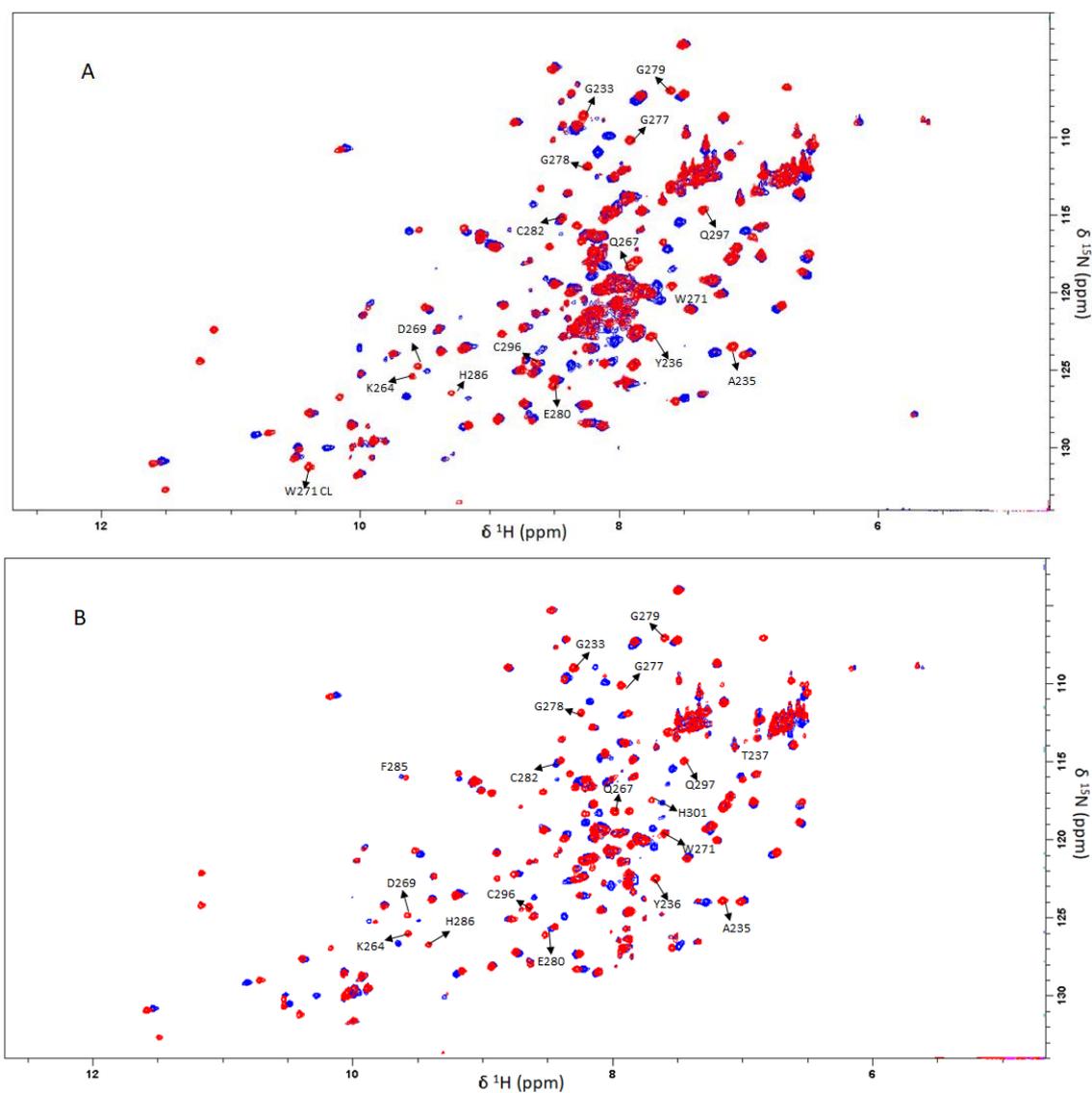


Figura 34: A- Sobreposição dos espectros de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD (azul) e da MGL-CRD na presença do ligando Tn-T3. B- Sobreposição dos espectros de $^1\text{H},^{15}\text{N}$ -HSQC da MGL-CRD (azul) e da MGL-CRD na presença do ligando Tn-T15. Os resíduos assinalados em ambos os espectros foram identificados a partir da análise de CSP. Os espectros foram adquiridos a 293K a pH 7,4.

7.6- CSPs da MGL-CRD na presença dos monoglicopéptidos Tn-T3 MUC1 e Tn-T15 MUC1

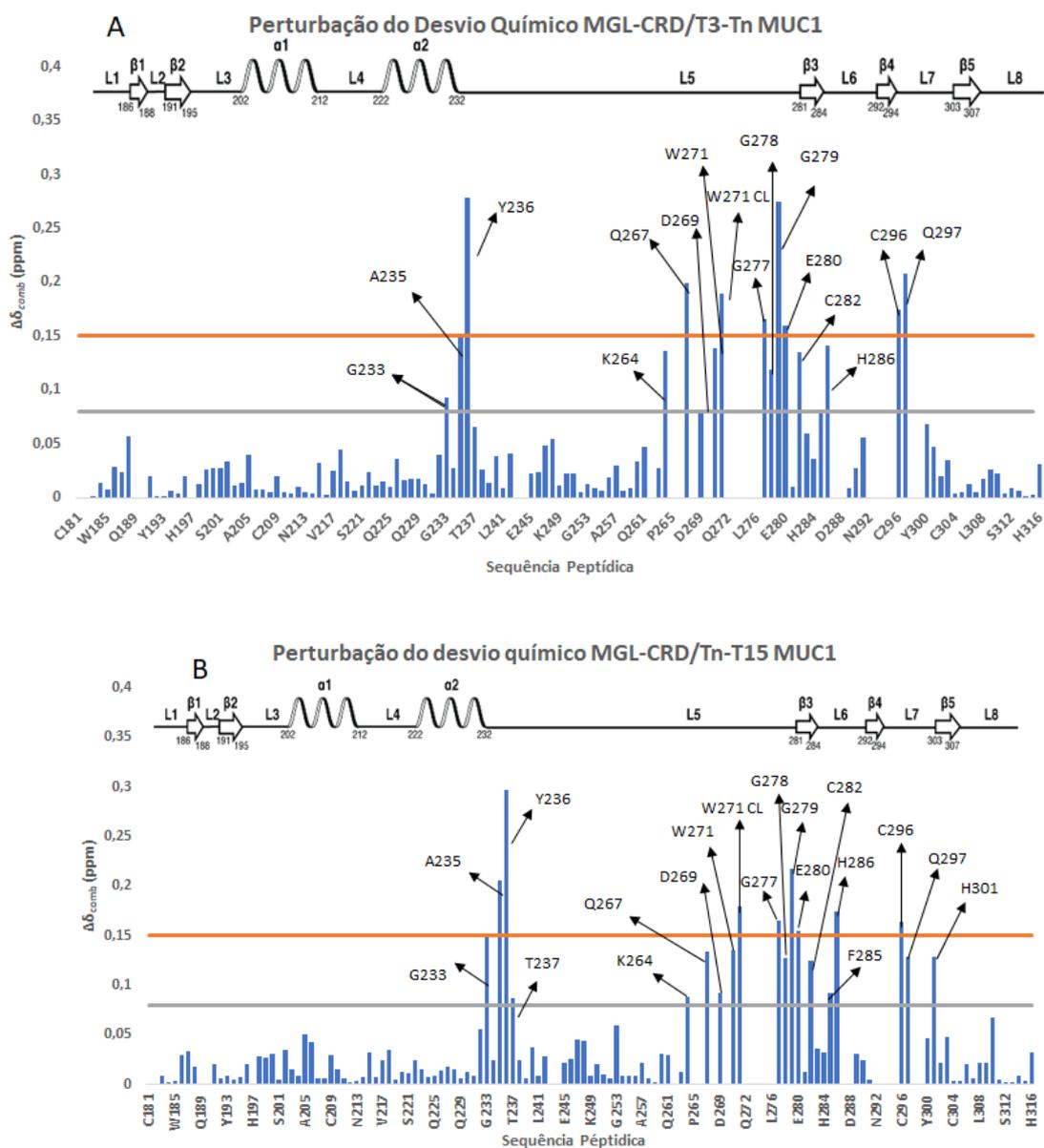


Figura 35: **A-** Histograma com a análise de CSP da MGL-CRD na presença do ligando Tn-T3 MUC1. **B-** Histograma com a análise de CSP da MGL-CRD na presença do ligando Tn-T15 MUC1. A linha a laranja e a linha a cinzento representam os valores de $\Delta\delta_{comb}$ considerados para os resíduos fortemente e moderadamente perturbados, respetivamente. Os resíduos identificados encontram-se assinalados na figura.

7.7- Espectro de $^1\text{H},^{15}\text{N}$ -HSQC do MGL-CRD na presença do diglicopéptido Tn-T3, T15 MUC1

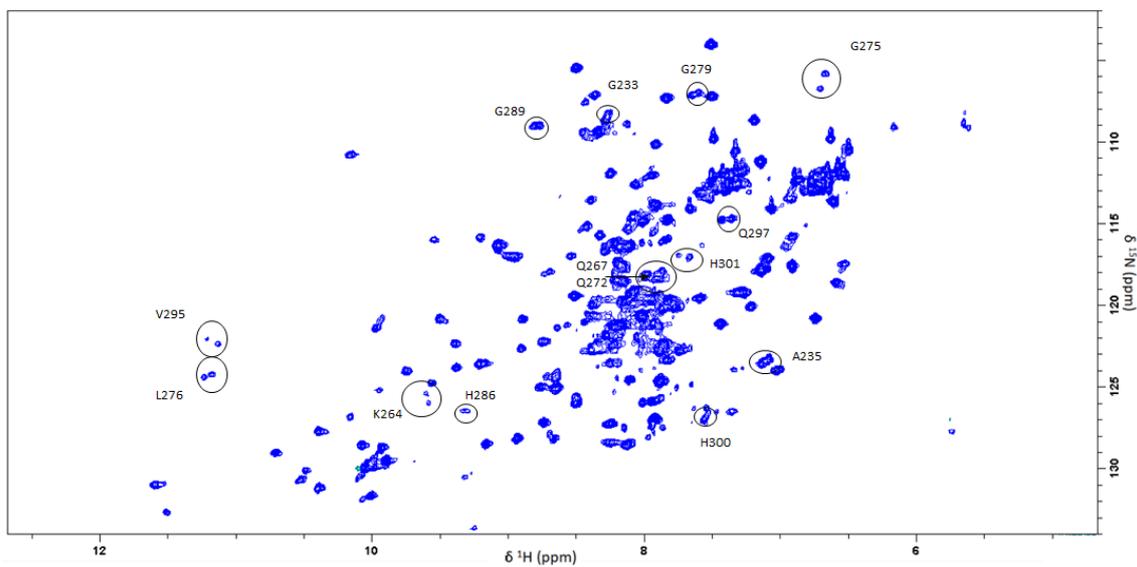


Figura 36: Espectro de $^1\text{H},^{15}\text{N}$ -HSQC do domínio MGL-CRD na presença do ligando Tn-T3, T15. Os resíduos assinalados correspondem aos resíduos identificados que apresentam duas ressonâncias correspondentes a cada um dos modos de união do domínio MGL-CRD existente em solução. O espectro foi adquirido a 293K a pH 7,4.

7.8- Dados de RMN das experiências de HetNOE da MGL-CRD livre e na presença dos ligandos estudados

Tabela 10- Valores obtidos da experiência de relaxação HetNOE e respectivos erros do domínio MGL-CRD na forma livre e na presença dos complexos MGL-CRD/1-5.

	MGL-CRD		MGL-CRD/1		MGL-CRD/2		MGL-CRD/3		MGL-CRD/4		MGL-CRD/5	
	HetNOE	Erro (\pm)	HetNOE	Erro (\pm)	HetNOE	Erro (\pm)	HetNOE	Erro (\pm)	HetNOE	Erro (\pm)	HetNOE	Erro (\pm)
C181												
P182												
V183	0,73		0,85		0,88		0,88				0,88	
N184	0,85	0,04	0,64	0,27	0,89	0,19	0,62	0,18	0,71	0,50	0,62	0,18
W185	0,79	0,11	0,98	0,28	1,00	0,27	0,68	0,19	0,95	0,25	0,64	0,19
V186	0,85	0,09	0,71	0,23		0,37	0,68	0,24	0,95	0,24	0,68	0,23
E187	0,50	0,14	0,55	0,25	0,61	0,54	1,02	0,25	1,18	0,40	0,97	0,25
H188	0,72	0,13	0,89	0,27	0,86	0,19	0,72	0,32	1,07	0,40	0,58	0,30
Q189		0,13		0,11		0,23		0,17		0,33		0,20
D190												
S191	0,90		0,64		0,86		0,89		0,43		0,96	
C192	0,87	0,09	1,05	0,16	0,68	0,20	0,97	0,17	1,08	0,23	1,13	0,17
Y193	0,94	0,16	0,84	0,44	1,03	0,28	1,05	0,52	0,75	0,49	1,05	0,52
W194	0,71	0,11	0,86	0,26		0,32	0,77	0,32		0,27	0,76	0,32
F195	0,95	0,06	1,16	0,15		0,34	0,61	0,19	0,79	0,35	0,63	0,18
S196	0,96	0,10	0,81	0,18	0,62	0,32	0,74	0,20	0,56	0,23	0,58	0,21
H197		0,05		0,15		0,12		0,13		0,17		0,09
S198	0,75		0,83				0,65		0,93		0,65	

G199	0,98	0,06	0,76	0,10	0,74	0,21	0,82	0,12	0,98	0,21	0,86	0,12
M200	0,75	0,06	0,69	0,19	0,84	0,14	0,93	0,14	1,05	0,28	0,93	0,14
S201	0,91	0,05	0,77	0,12	0,63	0,17	0,56	0,15	0,71	0,22	0,79	0,14
W202	0,68	0,04	0,93	0,10	0,68	0,11	1,09	0,16	0,86	0,11	1,09	0,10
A203	0,87	0,05	0,80	0,14	0,61	0,15	0,76	0,21	0,93	0,20	0,76	0,21
E204	1,06	0,04	0,78	0,09	0,81	0,09	0,77	0,09	0,88	0,12	0,77	0,09
A205	0,76	0,04	0,94	0,11	0,78	0,14	0,92	0,11	0,70	0,15	0,92	0,11
E206	0,33	0,04		0,14	0,63	0,14	0,86	0,13	0,82	0,13	0,31	0,13
K207	0,96	0,05	0,68	0,17	0,83	0,16	0,83	0,15	0,74	0,21	0,83	0,15
Y208	0,78	0,05	0,77	0,12	0,83	0,14	0,86	0,24	0,62	0,14	0,76	0,24
C209	1,05	0,04	0,75	0,08		0,08	0,88	0,10	0,64	0,09	0,95	0,07
Q210		0,08		0,14		0,26		0,21		0,17		0,23
L211	1,03	0,12	0,90	0,28	0,94	0,58	0,76	0,21	0,98	0,50	0,76	0,30
K212	0,76	0,05	0,79	0,17	0,77	0,20	0,80	0,16	0,46	0,18	0,80	0,16
N213	0,83	0,05	0,78	0,13	0,73	0,12	0,81	0,11	0,85	0,14	0,81	0,11
A214	0,65	0,05	0,77	0,12	1,06	0,12	0,88	0,14	0,76	0,15	0,88	0,13
H215	0,76	0,05	0,66	0,11	0,63	0,17	0,81	0,14	0,99	0,10	0,81	0,14
L216	1,22	0,08	0,97	0,21	0,90	0,17	0,98	0,26	0,79	0,22	0,98	0,26
V217	0,74	0,16	0,76	0,25	0,91	0,34	0,68	0,40	0,91	0,31	0,68	0,39
V218	0,55	0,09	1,00	0,17	1,02	0,28	1,02	0,17		0,31	1,02	0,17
I219	0,77	0,10	0,89	0,26	0,66	0,30	0,73	0,20	0,74	0,35	0,78	0,20
N220	0,83	0,06	0,79	0,17	0,58	0,13	0,68	0,12	1,03	0,16	0,69	0,12
S221	0,69	0,05	0,77	0,12	0,88	0,15	0,82	0,13	0,97	0,23	0,82	0,14
R222	0,97	0,03	0,79	0,07	0,89	0,10	0,57	0,09		0,13	0,57	0,09
E223	0,90	0,05	0,91	0,17	0,76	0,17	0,94	0,11	0,79	0,32	0,94	0,11
E224	1,00	0,04	0,85	0,12	0,79	0,13	0,94	0,12	0,94	0,14	0,94	0,12
Q225	0,74	0,06	0,89	0,15	0,96	0,15	0,91	0,14	0,96	0,18	0,91	0,14

N226	0,76	0,04	0,80	0,10	0,74	0,13	0,87	0,18	1,04	0,24	0,87	0,18
F227	0,79	0,04	0,85	0,09	0,87	0,11	0,90	0,16	0,63	0,26	0,90	0,16
V228	0,85	0,05	0,90	0,22	0,87	0,15	0,78	0,15	0,71	0,12	0,78	0,15
Q229	0,86	0,05	0,66	0,12	0,91	0,17	0,71	0,14	0,94	0,18	0,71	0,14
K230	0,78	0,05	0,70	0,11	0,66	0,14	0,74	0,12	0,93	0,14	0,74	0,12
Y231	0,90	0,05	0,53	0,10	0,83	0,10	0,90	0,11	0,77	0,14	0,90	0,11
L232	0,77	0,07	0,71	0,12	0,64	0,16	1,01	0,19	1,00	0,19	1,01	0,19
G233	0,72	0,05	0,86	0,18		0,14	0,67	0,17	0,62	0,24	0,67	0,17
S234	0,91	0,06	0,76	0,13	1,02	0,24	0,79	0,11	0,57	0,14	0,79	0,11
A235	0,48	0,04	0,61	0,23	0,49	0,21	0,35	0,10	0,79	0,10	0,44	0,10
Y236	0,87	0,06	0,57	0,08	0,43	0,11	1,09	0,09	0,72	0,18	1,09	0,09
T237	1,10	0,09	0,71	0,20	0,83	0,14	1,07	0,30	1,25	0,25	1,07	0,30
W238	0,81	0,07	0,93	0,18	0,60	0,22	0,88	0,28		0,37	0,88	0,28
M239	0,85	0,06	1,10	0,12	0,87	0,13	0,81	0,20	0,68	0,29	0,81	0,20
G240	0,92	0,07	1,06	0,28	0,74	0,18	0,90	0,18	0,65	0,21	0,90	0,18
L241	1,03	0,09	0,70	0,26	0,93	0,27	0,74	0,19	0,80	0,17	0,74	0,19
S242	0,91	0,04	0,71	0,11	0,95	0,14	0,75	0,10	0,88	0,15	0,78	0,10
D243	0,85	0,05	0,70	0,10	0,84	0,15	0,80	0,10	0,74	0,13	0,84	0,10
P244				0,06		0,08		0,08		0,08		0,06
E245	0,77		0,57		0,46		0,68		0,69		0,68	
G246	0,65	0,06	0,75	0,11	0,59	0,08	0,51	0,09	0,74	0,14	0,51	0,09
A247	0,61	0,08	0,74	0,08	0,76	0,08	0,60	0,06	0,76	0,11	0,60	0,06
W248	0,84	0,10	0,83	0,05	0,76	0,06	0,68	0,06	0,73	0,07	0,68	0,06
K249	0,64	0,10	0,82	0,09		0,09	0,81	0,08	0,90	0,10	0,78	0,08
W250	0,75	0,08	0,78	0,14	0,67	0,27	1,08	0,14	0,95	0,18	1,08	0,14
V251	0,86	0,08	0,83	0,15	0,94	0,13	1,06	0,20	0,53	0,22	1,06	0,20
D252	0,90	0,04	0,79	0,15	1,00	0,18	1,02	0,20	1,03	0,12	1,02	0,20

G253	0,73	0,04	0,85	0,08	0,84	0,12	0,73	0,13	0,87	0,15	0,73	0,13
T254	0,78	0,04	1,02	0,10	0,76	0,11	0,82	0,10	0,57	0,12	0,99	0,10
D255	0,70	0,04	0,78	0,08	1,09	0,08	1,00	0,08	0,61	0,09	1,00	0,10
Y256	0,82	0,05	0,85	0,17	0,61	0,17	0,63	0,16	0,94	0,13	0,63	0,16
A257	0,78	0,05	0,74	0,16	0,79	0,11	0,80	0,11	0,83	0,26	0,80	0,11
T258	0,80	0,03	0,79	0,08	0,79	0,10	0,70	0,08	0,62	0,10	0,70	0,08
G259	0,85	0,04	0,92	0,09	0,76	0,11	0,92	0,08	1,03	0,12	0,92	0,08
F260	0,84	0,03	0,74	0,08	0,95	0,09	0,84	0,10	0,92	0,12	0,84	0,10
Q261	0,94	0,03	0,88	0,06	0,84	0,22	0,95	0,06	0,84	0,11	0,90	0,06
N262		0,03		0,09		0,11		0,15		0,14		0,15
W263	0,75		0,87		0,85		0,86		0,74		0,86	
K264	0,63	0,03	0,82	0,09	1,04	0,10	0,79	0,12	0,53	0,17	0,93	0,12
P265		0,09		0,14		0,34		0,26		0,51		0,33
G266												
Q267	0,89	0,12	1,00	0,47	0,77		0,80	0,27	0,92	0,38	0,80	0,21
P268				0,07		0,17		0,34		0,27		0,31
D269	0,76		0,85		0,68		0,67		0,70		0,67	
D270		0,18	0,82	0,13	0,59	0,16	0,70	0,14	1,05	0,15	0,70	0,14
W271	0,81		0,79	0,12	0,54	0,20	0,93	0,14	0,95	0,18	0,93	0,14
Q272		0,15	0,85	0,10	0,60	0,13	0,98	0,11	0,68	0,20	1,04	0,11
G273				0,23		0,37		0,09		0,13		0,09
H274												
G275			0,81	0,16	0,81	0,24	0,86	0,14	1,04	0,38	0,88	0,14
L276			0,92	0,13	0,92	0,21	0,97	0,20	0,99	0,24	0,97	0,20
G277	0,74	0,08	0,81	0,11	0,71	0,12	0,66	0,08	0,75	0,10	0,66	0,08
G278	0,51	0,10	0,50	0,12	0,72	0,14	0,71	0,09	0,58	0,12	0,71	0,09
G279	0,64	0,06	0,78	0,09	0,86	0,20	0,70	0,10	0,69	0,12	0,72	0,10

E280	1,01	0,14	0,75	0,12		0,39	0,84	0,12	1,04	0,17	0,84	0,12
D281	0,84	0,04	0,87	0,07	1,01	0,10	0,69	0,10	1,00	0,12	0,69	0,10
C282	0,98	0,26	0,83	0,08	0,97	0,15	0,81	0,10	0,85	0,13	0,81	0,10
A283	0,80	0,06	0,70	0,12	0,76	0,14	0,61	0,09	0,75	0,11	0,61	0,09
H284	0,96	0,08	0,90	0,13	0,75	0,16	0,82	0,13	0,86	0,14	0,82	0,13
F285	0,73	0,12	0,90	0,19	0,83	0,18	0,82	0,16	1,24	0,32	0,81	0,17
H286	1,10	0,22	0,92	0,17	0,58	0,18	0,61	0,16	1,25	0,31	0,61	0,16
P287												
D288			0,89	0,13	0,92	0,17	0,85	0,18	0,86	0,20	0,85	0,18
G289	0,98	0,05	0,91	0,09	0,65	0,13	0,97	0,20	0,83	0,18	0,97	0,20
R290	0,79	0,07	0,88	0,11	0,86	0,13	0,53	0,11	0,87	0,18	0,50	0,11
W291	0,78	0,05	0,80	0,12	1,05	0,24	0,86	0,15	1,14	0,27	0,86	0,14
N292			0,98	0,10	0,73	0,18	1,00	0,24	0,68	0,14	1,25	0,23
D293			0,94	0,10	0,79	0,13	0,98	0,18	0,84	0,14	0,97	0,18
D294			0,85	0,14	0,90	0,20	0,77	0,16	1,10	0,24	0,90	0,20
V295			0,85	0,14	0,65	0,31	0,67	0,16	0,81	0,23	0,67	0,16
C296	0,56	0,20	0,99	0,29	1,10	0,37	0,93	0,25	0,95	0,19	0,89	0,17
Q297	0,86	0,05	0,71	0,11	0,89	0,24	0,95	0,14	0,89	0,20	0,95	0,14
R298			0,83	0,08	0,67	0,12	0,71	0,09	1,03	0,14	0,71	0,09
P299												
Y300	0,66	0,04	1,10	0,13	0,82	0,14	0,85	0,12	0,95	0,17	0,85	0,12
H301	0,83	0,09	1,06	0,15	0,63	0,26	1,00	0,36	0,78	0,24	1,08	0,39
W302	0,79	0,06	1,02	0,22	0,64	0,20	1,09	0,21	0,86	0,18	1,09	0,20
V303	0,94	0,06	0,93	0,18	0,99	0,25	0,74	0,16	0,75	0,16	0,74	0,16
C304	1,09	0,08	0,64	0,15	0,93	0,25	0,72	0,16	0,76	0,18	0,72	0,16
E305	0,78	0,08	0,78	0,21	1,00	0,25	0,92	0,20	0,86	0,19	0,92	0,20
A306	1,04	0,18	0,77	0,44	0,64	0,28	0,58	0,26	0,29	0,32	0,88	0,30

G307	0,76	0,10	0,90	0,22	0,72	0,27	0,77	0,27	0,71	0,20	0,77	0,26
L308	0,64	0,10	0,57	0,21	0,82	0,24	0,89	0,24	0,68	0,27	0,89	0,23
G309	0,61	0,06	0,99	0,16	0,50	0,16	0,40	0,12	0,81	0,22	0,40	0,11
Q310	0,69	0,07	0,88	0,33	0,82	0,17	0,73	0,44	0,52	0,21	1,00	0,18
T311	0,79	0,05	0,39	0,21	0,64	0,24	0,30	0,14	0,46	0,21	0,38	0,10
S312		0,10		0,43		0,42		0,28		0,43		0,20
Q313	-0,47	0,04		0,13	0,77	0,17	0,17	0,09	0,57	0,18	0,16	0,09
E314	-0,22	0,02	-0,13	0,06	-0,28	0,05	-0,28	0,03	-0,42	0,08	-0,28	0,03
S315	-0,27	0,03		0,14	-0,63	0,11	-0,61	0,07	-0,73	0,20	-0,61	0,07
H316	-1,07	0,01	-0,74	0,03	-1,12	0,03	-1,17	0,03	-0,97	0,03	-1,17	0,03

7.9- Valores de R_1 obtidos para a MGL-CRD na forma livre e na presença dos ligandos estudados

Tabela 11- Valores obtidos da experiência de relaxação R_1 e respectivos erros do domínio MGL-CRD na forma livre e na presença dos complexos MGL-CRD/1-5.

	MGL-CRD		MGL-CRD/1		MGL-CRD/2		MGL-CRD/3		MGL-CRD/4		MGL-CRD/5	
	R_1 (s^{-1})	Erro (\pm)	R_1 (s^{-1})	Erro (\pm)	R_1 (s^{-1})	Erro (\pm)	R_1 (s^{-1})	Erro (\pm)	R_1 (s^{-1})	Erro (\pm)	R_1 (s^{-1})	Erro (\pm)
M180												
C181												
P182												
V183	1,25	0,09	1,12	0,03	1,37	0,19	1,25	0,06	1,31	0,11	1,34	0,05
N184	1,24	0,01	1,23	0,13	1,31	0,07	1,41	0,13	1,26	0,19	1,20	0,12
W185	1,14	0,03	1,08	0,13	1,12	0,04	1,22	0,11	0,98	0,07	1,09	0,05
V186	1,48	0,09	1,06	0,27	1,04	0,05	1,22	0,48	0,85	0,02	1,32	0,15
E187	1,22	0,09	1,36	0,10	1,22	0,04	1,37	0,10	0,94	0,09	1,02	0,06
H188	1,74	0,08	1,13	0,06	1,52	0,10	1,65	0,08	0,86	0,06	1,37	0,05
Q189												
D190												
S191	1,37	0,09	1,07	0,07	1,30	0,04	1,42	0,07	1,13	0,04	1,20	0,08
C192	0,78	0,21	1,15	0,25	1,21	0,29	0,95	0,07	1,11	0,06	0,72	0,18
Y193	1,22	0,12	1,21	0,10	1,01	0,15	1,00	0,25	0,98	0,08	1,33	0,19
W194	1,28	0,08	1,39	0,11	1,45	0,11	1,48	0,09	1,46	0,07	1,31	0,11
F195	1,45	0,25	1,14	0,08	0,67	0,02	0,85	0,14	1,10	0,08	1,07	0,02
S196	1,30	0,02	1,20	0,06	1,47	0,12	1,15	0,08	1,14	0,02	1,36	0,03
H197												
S198	1,60	0,01	1,25	0,07	1,42	0,09	1,20	0,06	1,30	0,10	1,41	0,07
G199	1,37	0,06	1,31	0,02	1,71	0,06	1,64	0,03	1,51	0,02	1,33	0,14

M200	1,38	0,12	1,48	0,21	1,26	0,06	1,29	0,08	1,54	0,03	1,40	0,09
S201	1,36	0,05	1,26	0,04	1,28	0,02	1,34	0,02	1,37	0,03	1,29	0,02
W202	1,31	0,08	1,22	0,15	1,29	0,13	1,32	0,17	1,31	0,14	1,36	0,01
A203	1,22	0,03	1,17	0,05	1,21	0,02	1,27	0,01	1,22	0,07	1,30	0,04
E204	1,23	0,02	1,19	0,05	1,25	0,02	1,17	0,04	1,26	0,07	1,17	0,03
A205	1,28	0,06	1,15	0,04	1,07	0,04	1,23	0,04	1,25	0,03	1,13	0,08
E206	1,31	0,19	1,23	0,17	1,26	0,11	1,18	0,04	1,60	0,39	1,58	0,16
K207	1,30	0,01	1,27	0,03	1,16	0,07	1,10	0,04	1,15	0,05	1,31	0,02
Y208	1,29	0,01	1,21	0,01	1,40	0,05	1,26	0,04	1,20	0,02	1,28	0,04
C209	0,95	0,06	1,25	0,06	1,15	0,12	1,20	0,10	1,21	0,07	1,22	0,07
Q210		1,50	0,86	0,14	0,90	0,09	0,85	0,40	1,37	0,10	1,54	0,14
L211	1,28	0,02	1,20	0,07	1,19	0,04	1,28	0,09	1,27	0,08	1,26	0,04
K212	1,20	0,03	1,18	0,02	1,24	0,03	1,18	0,02	1,21	0,01	1,10	0,07
N213	1,38	0,08	1,17	0,07	1,24	0,01	1,24	0,07	1,04	0,06	1,14	0,04
A214	1,30	0,09	1,28	0,04	1,21	0,06	1,15	0,03	1,21	0,05	1,36	0,02
H215	1,20	0,10	1,24	0,05	1,19	0,06	1,30	0,04	1,23	0,20	1,27	0,20
L216	1,37	0,03	1,26	0,35	1,00	0,28	1,19	0,13	1,60	0,28	1,08	0,08
V217	1,57	0,18	1,39	0,09	1,40	0,20	1,49	0,05	1,46	0,27	1,30	0,05
V218	1,40	0,06	1,17	0,06	1,09	0,05	1,06	0,04	0,99	0,13	1,07	0,13
I219	1,31	0,10	1,06	0,14	0,83	0,06	0,71	0,18	0,69	0,12	1,17	0,28
N220	1,17	0,04	1,08	0,04	1,23	0,05	1,30	0,11	1,08	0,10	1,25	0,06
S221	1,36	0,02	1,26	0,02	1,23	0,08	1,36	0,05	1,21	0,08	1,25	0,07
R222	1,32	0,05	1,08	0,11	1,55	0,07	1,46	0,09	1,57	0,19	1,53	0,10
E223	1,47	0,11	1,33	0,10	1,48	0,02	1,50	0,13	1,35	0,03	1,59	0,04
E224	1,45	0,00	1,29	0,09	1,08	0,03	0,82	0,02	1,35	0,11	1,27	0,07
Q225	1,33	0,02	1,37	0,05	1,24	0,04	1,28	0,15	1,32	0,07	1,27	0,02
N226	1,44	0,08	1,23	0,03	1,37	0,07	1,42	0,10	1,33	0,07	1,42	0,04

F227	1,22	0,06	1,29	0,13	1,37	0,13	1,33	0,06	1,10	0,13	1,13	0,06
V228	1,46	0,09	1,25	0,08	1,47	0,16	1,48	0,08	1,26	0,05	1,55	0,17
Q229	1,54	0,01	1,39	0,02	1,29	0,12	1,27	0,08	1,37	0,06	1,42	0,04
K230	1,38	0,01	1,36	0,01	1,27	0,08	1,33	0,05	1,21	0,08	1,35	0,04
Y231	1,27	0,07	1,21	0,03	1,38	0,02	1,37	0,06	1,09	0,10	1,25	0,18
L232	1,41	0,04	1,40	0,14	1,24	0,06	1,12	0,06	1,19	0,12	1,25	0,13
G233	1,51	0,04	1,29	0,03	1,42	0,08	1,36	0,03	1,28	0,09	1,54	0,10
S234	1,40	0,04	1,25	0,08	1,28	0,03	1,34	0,04	1,37	0,05	1,30	0,07
A235	1,38	0,10	1,04	0,09	1,28	0,06	0,99	0,06	1,25	0,06	1,19	0,07
Y236	1,19	0,11	1,15	0,08	1,28	0,12	1,33	0,16	1,07	0,08	1,43	0,10
T237	1,45	0,15	1,07	0,03	1,17	0,07	0,99	0,03	1,24	0,26	1,22	0,17
W238	1,35	0,03	1,23	0,05	1,31	0,05	1,66	0,03	1,48	0,03	1,29	0,06
M239	1,22	0,10	1,06	0,04	0,92	0,11	1,06	0,16	1,09	0,06	0,95	0,12
G240	1,24	0,13	1,06	0,04	1,29	0,14	1,59	0,09	1,22	0,11	1,42	0,11
L241	1,24	0,03	1,11	0,11	1,20	0,06	1,35	0,10	1,44	0,17	1,33	0,01
S242	1,46	0,03	1,17	0,03	1,24	0,03	1,25	0,02	1,13	0,09	1,34	0,04
D243	1,21	0,05	1,12	0,02	1,22	0,05	1,13	0,05	1,16	0,02	1,13	0,02
P244												
E245	1,29	0,06	1,17	0,09	1,34	0,04	1,30	0,02	1,26	0,03	1,28	0,03
G246	1,40	0,06	1,15	0,04	1,30	0,05	1,22	0,06	1,18	0,05	1,18	0,08
A247	1,13	0,05	1,08	0,04	1,16	0,02	1,08	0,01	1,12	0,04	1,12	0,03
W248	1,24	0,13	1,11	0,05	1,13	0,07	1,21	0,01	1,24	0,03	1,15	0,05
K249	1,31	0,13	1,02	0,08	1,23	0,04	1,36	0,04	1,00	0,07	1,33	0,01
W250	1,48	0,04	1,23	0,02	1,17	0,10	1,20	0,02	1,05	0,03	1,24	0,14
V251	1,22	0,01	1,25	0,05	1,26	0,04	1,34	0,05	1,20	0,09	1,37	0,15
D252	1,49	0,05	1,27	0,04	1,27	0,06	1,19	0,06	1,21	0,10	1,26	0,04
G253	1,36	0,06	1,36	0,02	1,31	0,05	1,43	0,06	1,23	0,07	1,18	0,03

T254	1,35	0,02	1,22	0,04	1,41	0,08	1,16	0,05	1,19	0,02	1,28	0,01
D255	1,32	0,10	1,18	0,06	1,43	0,05	1,18	0,03	1,33	0,04	1,28	0,10
Y256	1,31	0,05	1,10	0,05	1,46	0,11	1,49	0,12	1,35	0,14	1,31	0,07
A257	1,44	0,03	1,27	0,04	1,39	0,07	1,42	0,01	1,32	0,05	1,42	0,05
T258	1,40	0,03	1,17	0,02	1,46	0,02	1,44	0,13	1,29	0,04	1,51	0,03
G259	1,31	0,03	1,01	0,03	1,40	0,04	1,37	0,08	1,19	0,01	1,36	0,04
F260	1,21	0,01	1,10	0,04	1,20	0,05	1,12	0,02	1,17	0,04	1,13	0,04
Q261	1,36	0,04	1,25	0,05	1,17	0,04	1,29	0,04	1,18	0,03	1,14	0,04
N262												
W263	1,42	0,05	1,22	0,05	1,31	0,12	1,31	0,08	1,31	0,06	1,12	0,02
K264	1,50	0,04	1,20	0,09	1,17	0,16	1,62	0,04	0,98	0,03	1,56	0,16
P265												
G266		0,58		0,49		12,57		1,67		0,06		0,62
Q267	1,22	0,02	1,12	0,12	1,39	0,05	1,04	0,06	1,44	0,09	1,09	0,01
P268												
D269	1,17	0,18	1,17	0,02	1,25	0,17	1,45	0,04	1,33	0,09	1,42	0,12
D270			1,13	0,11	1,55	0,05	1,67	0,19	1,08	0,12	1,55	0,05
W271	1,14	0,10	1,22	0,05	1,30	0,16	1,45	0,11	1,27	0,04	1,30	0,04
Q272			1,26	0,17	1,23	0,18	1,23	0,21	1,25	0,12	1,28	0,05
G273												
H274				0,04		4,06		0,59		6,52		11,79
G275			1,08	0,11	1,11	0,06		0,11	1,02	0,11	1,36	0,10
L276			1,18		1,04	0,09	0,57	0,09	1,03	0,03	1,06	0,10
G277	1,41	0,13	1,17	0,06	1,20	0,05	1,19	0,05	1,21	0,03	1,20	0,03
G278	1,29	0,13	1,00	0,09	1,06	0,09	0,89	0,06	1,06	0,03	1,22	0,06
G279	1,47	0,06	1,05	0,04	1,29	0,10	1,15	0,07	1,12	0,10	1,22	0,05
E280	1,40	0,16	1,15	0,08	1,04	0,08	1,43	0,09	1,29	0,05	1,31	0,03

D281	1,06	0,03	1,03	0,05	1,19	0,02	1,11	0,03	0,95	0,07	1,28	0,04
C282	1,23	0,19	1,10	0,05	1,21	0,12	1,34	0,08	1,07	0,08	1,26	0,01
A283	1,33	0,02	1,11	0,07	1,22	0,06	1,39	0,02	1,21	0,04	1,16	0,10
H284	1,16	0,18	1,18	0,11	1,31	0,24	1,37	0,06	0,86	0,13	1,22	0,14
F285	1,06	0,08	1,14	0,13	1,01	0,12	1,24	0,06	1,57	0,13	1,29	0,14
H286	1,14	0,18	1,36	0,10	1,16	0,11	1,00	0,10	1,35	0,16	1,35	0,21
P287												
D288			1,27	0,06	1,12	0,08	1,14	0,06	1,20	0,04	1,37	0,10
G289	1,45	0,11	1,38	0,10	1,37	0,10	1,25	0,12	1,02	0,24	1,56	0,14
R290	1,41	0,01	1,09	0,04	0,82	0,08	1,56	0,05	1,48	0,09	1,41	0,16
W291	1,32	0,04	1,30	0,13	1,30	0,05	1,30	0,14	1,44	0,09	1,32	0,09
N292			1,04	0,05	1,39	0,08	1,45	0,26	1,13	0,22	1,21	0,07
D293			1,12	0,05	1,18	0,10	1,00	0,08	1,06	0,05	1,44	0,15
D294			1,09	0,06	1,17	0,17	1,33	0,06	1,05	0,10	1,11	0,13
V295			1,34	0,10	1,22	0,04	1,60	0,07	1,38	0,14	1,47	0,07
C296	1,32	0,09	1,03	0,16	1,08	0,19	0,71	0,14	1,33	0,10	1,66	0,12
Q297	1,37	0,04	1,35	0,10	1,47	0,05	1,37	0,07	1,29	0,24	1,40	0,01
R298			1,26	0,07	1,41	0,09	1,75	0,07	1,39	0,04	1,45	0,06
P299												
Y300	1,33	0,08	1,15	0,16	1,08	0,17	1,30	0,14	1,45	0,16	1,44	0,07
H301	1,25	0,10	1,20	0,31	1,30	0,14	1,18	0,14	2,00	0,11	1,15	0,08
W302	1,37	0,10	1,26	0,04	1,34	0,19	1,49	0,03	1,29	0,13	1,36	0,10
V303	1,20	0,01	1,21	0,15	1,39	0,06	1,64	0,07	1,05	0,18	1,21	0,09
C304	1,25	0,09	1,45	0,06	0,98	0,16	0,90	0,03	1,03	0,19	1,19	0,12
E305	1,49	0,14	1,41	0,06	1,02	0,18	1,11	0,13	0,86	0,14	1,16	0,19
A306		0,42		0,23		0,06		0,01		0,26		0,43
G307	1,58	0,22	1,47	0,16	1,33	0,17	1,10	0,01	1,49	0,24	1,07	0,09

L308	1,57	0,14	1,42	0,04	1,45	0,08	1,55	0,12	1,12	0,09	1,49	0,02
G309	1,51	0,03	1,48	0,13	1,53	0,13	1,30	0,11	1,54	0,11	1,67	0,04
Q310	0,92	0,06	1,08	0,09	0,81	0,08	1,09	0,07	1,14	0,11	1,21	0,04
T311	1,00	0,35	1,13	0,13	1,54	0,47	0,87	0,09	1,47	0,12	1,21	0,05
S312		0,69		0,17		2,22		0,38		0,17		0,61
Q313	1,35	0,13	1,17	0,10	1,23	0,05	1,24	0,06	1,57	0,08	1,46	0,20
E314	1,44	0,04	1,30	0,03	1,46	0,09	0,06	0,13	1,55	0,02	1,54	0,01
S315	1,40	0,14	1,14	0,04	1,47	0,17	1,35	0,08	1,32	0,02	1,34	0,04
H316	1,21	0,01	1,01	0,00	1,28	0,02	1,33	0,01	1,18	0,01	1,25	0,01

7.10- Valores de R_2 obtidos para a MGL-CRD na forma livre e na presença dos ligandos estudados

Tabela 12- Valores obtidos da experiência de relaxação R_2 e respectivos erros do domínio MGL-CRD na forma livre e na presença dos complexos MGL-CRD/1-5.

	MGL-CRD		MGL-CRD/1		MGL-CRD/2		MGL-CRD/3		MGL-CRD/4		MGL-CRD/5	
	R_2 (s^{-1})	Erro (\pm)	R_2 (s^{-1})	Erro (\pm)	R_2 (s^{-1})	Erro (\pm)	R_2 (s^{-1})	Erro (\pm)	R_2 (s^{-1})	Erro (\pm)	R_2 (s^{-1})	Erro (\pm)
M180												
C181												
P182												
V183	14,40	0,14	16,84	0,56	17,04	0,92	17,91	0,16	16,09	0,91	15,86	0,07
N184	14,19	0,21	13,11	0,94	13,45	0,66	14,58	0,32	14,54	1,02	14,94	1,03
W185	13,60	0,31	10,88	1,93	15,47	0,74	12,39	0,96	15,50	0,22	16,23	0,54
V186	15,10	1,45	14,20	2,07	14,58	2,36	12,25	1,49	14,06	0,65	10,24	1,57
E187	12,96	0,35	13,21	0,42	11,66	0,56	11,10	0,24	12,92	1,26	12,11	0,48
H188	11,12	1,03	10,90	0,71	11,04	0,74	11,71	0,10	12,93	0,98	12,78	1,60
Q189												
D190												
S191	12,91	0,73	14,62	0,96	12,93	0,85	13,14	0,16	12,92	0,45	13,91	0,56
C192	12,88	0,59	9,73	0,78	11,45	1,15	11,28	0,21	12,62	0,63	9,93	1,25
Y193	12,78	1,06	10,92	0,17	12,30	0,53	12,40	0,94	13,00	0,41	12,24	0,82
W194	10,18	0,21	11,65	0,70	11,76	0,14	10,82	0,54	9,70	0,66	10,15	0,70
F195	14,30	1,33	11,61	0,87	11,33	0,16	10,20	0,17	10,26	0,92	11,69	0,73
S196	13,68	0,32	10,86	0,94	12,79	0,09	11,25	0,45	14,56	0,38	11,73	0,49
H197												
S198	11,42	0,09	9,70	0,75	10,92	0,13	10,77	0,26	10,66	0,22	11,22	0,67

G199	11,52	0,10	11,83	0,76	10,74	0,92	10,63	0,18	12,29	0,74	11,83	0,54
M200	12,03	0,29	11,42	0,31	11,59	0,27	12,07	0,31	11,41	0,33	11,47	0,21
S201	10,92	0,31	9,32	0,23	10,26	0,12	9,92	0,15	10,44	0,29	9,94	0,17
W202	15,51	0,18	12,34	0,80	13,31	1,21	13,73	0,90	14,29	0,60	14,03	0,32
A203	13,51	0,07	13,49	0,26	13,03	0,13	13,07	0,24	13,65	0,19	12,91	0,09
E204	12,65	0,18	12,56	0,18	12,78	0,26	12,18	0,35	13,24	0,41	13,13	0,24
A205	15,02	0,10	14,98	0,08	14,64	0,10	13,96	0,53	14,05	0,12	15,13	0,36
E206	5,62	0,19	6,27	0,64	7,69	0,28	6,28	0,46	5,81	0,41	10,01	0,15
K207	15,32	0,54	14,50	0,34	13,43	0,32	14,09	0,48	13,94	0,31	14,18	0,48
Y208	15,35	0,08	14,44	0,21	16,51	0,23	16,62	0,25	15,84	0,29	14,74	0,16
C209	14,42	0,52	11,26	0,69	14,45	0,38	15,32	0,10	15,58	0,81	14,37	0,44
Q210		0,88		1,04		0,81		1,80		0,14		0,40
L211	14,29	0,35	14,29	0,73	11,31	0,22	12,45	0,14	14,31	0,40	12,31	0,10
K212	12,29	0,07	13,16	0,24	12,23	0,30	12,11	0,28	13,22	0,33	12,41	0,16
N213	11,98	0,25	10,73	0,26	12,49	0,37	11,44	0,21	11,31	0,46	11,38	0,35
A214	11,91	0,17	12,03	0,14	12,74	0,15	12,36	0,30	13,69	0,47	12,51	0,22
H215	12,41	0,18	9,89	0,29	11,29	0,50	11,91	0,72	13,50	0,48	12,25	0,33
L216	11,35	0,79	9,38	0,76	11,53	1,70	11,60	0,37	10,82	1,56	11,47	1,10
V217	14,75	0,30	14,77	0,77	12,86	0,61	11,50	0,53	14,39	0,32	11,61	0,33
V218	16,05	1,29	14,81	1,15	12,37	0,20	13,90	0,95	16,09	0,45	13,95	0,37
I219	14,24	0,55	15,09	2,38	14,64	1,58	12,75	0,75	12,98	0,69	11,93	1,34
N220	11,80	0,23	11,40	0,44	12,30	0,38	11,93	0,40	12,77	0,87	12,29	0,47
S221	11,14	0,16	11,06	0,24	11,41	0,29	11,03	0,16	11,85	0,10	11,59	0,05
R222	11,54	0,25	8,12	0,36	11,42	0,20	11,20	0,09	12,27	0,30	11,63	0,09
E223	11,93	0,17	10,89	0,50	12,45	0,18	12,00	0,13	13,30	0,38	12,00	0,25
E224	13,77	0,44	12,48	0,36	1,29	0,03	11,53	0,36	10,20	0,54	8,77	0,10
Q225	11,92	0,22	13,04	0,33	11,87	0,55	11,86	0,12	12,50	0,36	12,35	0,51

N226	11,74	0,11	12,38	0,28	12,40	0,23	11,78	0,36	12,78	0,17	12,47	0,45
F227	11,27	0,49	12,34	0,08	11,63	0,14	12,08	0,44	12,11	0,80	12,19	0,55
V228	11,87	0,18	12,48	0,43	11,19	0,26	12,14	0,34	9,95	0,42	11,15	0,29
Q229	10,86	0,17	12,01	0,33	11,98	0,20	11,61	0,18	11,52	0,25	12,13	0,44
K230	12,79	0,21	11,81	0,25	12,68	0,38	12,22	0,25	12,77	0,23	12,52	0,26
Y231	11,26	0,24	11,01	0,48	11,37	0,36	11,13	0,20	11,78	0,61	11,49	0,06
L232	12,00	0,12	11,72	0,27	12,57	0,59	11,92	0,34	12,05	0,75	11,57	0,22
G233	9,40	0,33	9,68	0,27	9,92	0,42	10,17	0,31	10,76	0,26	10,17	0,10
S234	10,90	0,10	8,12	0,47	10,39	0,34	9,99	0,38	10,56	0,36	10,07	0,16
A235	11,57	0,42	9,51	0,28	9,18	0,25	9,55	0,40	10,16	0,25	9,98	0,07
Y236	17,67	0,75	10,04	0,52	8,35	0,44	10,42	0,52	10,33	0,39	8,98	0,19
T237	12,77	0,36	11,08	0,88	10,53	0,54	10,60	0,86	11,63	1,92	10,71	0,25
W238	12,22	0,31	13,60	0,59	11,30	0,32	11,08	0,51	11,91	0,45	12,18	0,25
M239	12,80	0,37	11,50	0,60	13,45	0,39	13,46	0,37	14,23	0,35	14,59	0,83
G240	13,14	0,15	12,47	0,76	12,40	0,22	11,18	0,60	11,53	0,70	11,69	0,36
L241	12,41	0,24	13,47	0,96	12,44	0,10	11,23	0,35	13,73	0,40	11,31	0,04
S242	12,42	0,38	11,79	0,76	14,05	0,86	13,26	0,12	13,82	0,42	13,26	0,18
D243	11,53	0,37	10,13	0,33	11,27	0,19	11,18	0,09	12,33	0,09	11,09	0,09
P244												
E245	13,01	0,41	12,02	0,34	12,27	0,19	11,48	0,08	12,56	0,25	11,79	0,25
G246	11,88	0,06	10,79	0,34	10,98	0,11	10,67	0,14	12,11	0,41	10,76	0,19
A247	13,95	0,36	10,96	0,05	10,78	0,13	10,47	0,09	10,98	0,15	10,51	0,21
W248	11,47	0,14	11,12	0,09	11,49	0,09	11,38	0,17	11,95	0,07	11,29	0,16
K249	12,81	0,18	11,20	0,41	12,54	0,28	12,13	0,20	11,98	0,75	12,17	0,33
W250	11,88	0,37	9,38	0,45	10,91	0,18	10,75	0,21	11,83	0,80	11,32	0,41
V251	13,32	0,09	11,07	0,25	13,43	0,18	13,03	0,09	14,03	0,39	12,29	0,20
D252	13,00	0,38	12,85	0,24	13,41	0,34	12,41	0,17	13,81	0,25	13,16	0,01

G253	12,23	0,22	12,24	0,16	13,35	0,22	13,05	0,33	13,87	0,15	13,43	0,16
T254	15,45	0,48	14,49	0,28	15,36	0,40	14,52	0,31	15,86	0,26	15,42	0,29
D255	17,92	0,36	19,98	0,58	17,88	0,54	20,49	0,41	18,93	1,52	19,08	0,31
Y256	11,37	0,19	10,50	0,36	11,40	0,27	11,51	0,08	12,03	0,37	11,28	0,34
A257	10,79	0,14	11,82	0,18	11,97	0,08	11,53	0,17	11,62	0,24	11,48	0,19
T258	10,68	0,18	10,87	0,21	10,97	0,20	10,76	0,42	11,60	0,12	11,18	0,02
G259	11,17	0,29	10,94	0,35	11,49	0,28	11,55	0,19	12,11	0,30	11,37	0,12
F260	11,53	0,16	10,66	0,28	11,36	0,17	10,78	0,22	12,33	0,12	11,09	0,07
Q261	10,82	0,06	11,61	0,27	11,19	0,12	10,42	0,14	11,38	0,44	11,13	0,06
N262												
W263	11,17	0,05	11,08	0,15	11,39	0,32	11,62	0,17	12,53	0,61	11,72	0,26
K264	14,70	0,41	11,21	0,45	11,01	0,32	11,11	0,31	10,98	0,59	10,14	0,66
P265												
G266												
Q267	11,91	0,31	12,21	0,23	12,01	0,26	13,15	1,01	13,44	0,35	13,05	0,12
P268												
D269	20,12	2,14	11,70	0,12	11,05	0,60	12,49	0,35	12,80	1,49	11,82	0,41
D270			16,02	0,92	13,10	0,59	11,72	0,84	10,86	1,08	13,46	0,24
W271	13,76	0,41	13,41	0,31	14,12	0,26	12,05	0,29	13,44	0,71	13,04	0,49
Q272			13,97	0,54	14,60	0,68	12,86	0,28	10,86	1,12	12,79	0,31
G273												
H274												
G275			12,87	0,65	11,77	0,83		7,60	12,74	2,93	12,29	0,32
L276			12,19	0,70	11,09	0,47	10,26	0,86	11,35	0,59	10,43	0,35
G277	9,62	0,04	10,57	0,18	10,76	0,16	9,98	0,17	11,10	0,37	10,36	0,22
G278	15,33	0,51	9,21	0,34	10,56	0,03	10,50	0,15	10,79	0,07	10,15	0,12
G279	13,02	0,22	10,63	0,18	10,89	0,38	12,00	0,02	12,38	0,41	11,26	0,44

E280	20,91	0,44	9,85	0,41	9,38	0,54	11,17	0,25	11,74	0,11	10,83	0,28
D281	10,06	0,27	9,59	0,18	10,39	0,21	10,62	0,00	11,37	0,29	11,13	0,17
C282	15,36	1,90	13,86	1,06	11,90	0,31	12,74	0,79	13,60	0,36	12,77	0,35
A283	12,51	0,12	10,67	0,38	13,12	0,36	11,20	0,48	12,47	0,39	11,34	0,30
H284	15,59	0,68	12,48	0,94	12,46	0,54	13,42	0,26	15,79	0,59	12,81	0,22
F285	18,62	1,51	13,17	0,35	13,46	0,82	13,53	0,62	11,38	0,74	13,08	1,03
H286	16,63	0,86	11,20	0,27	10,42	0,66	10,66	0,49	12,77	0,33	10,55	0,98
P287												
D288			11,22	0,51	12,04	0,41	12,11	0,47	12,22	0,29		
G289	12,68	0,33	11,08	0,43	11,08	0,20	10,99	0,23	12,37	0,18	11,34	0,24
R290	11,86	0,25	10,35	0,48	12,01	0,22	11,98	0,24	7,35	0,25	12,33	0,29
W291	13,51	0,12	10,98	0,74	10,68	0,56	10,36	0,29	11,25	0,24	12,17	0,49
N292			12,07	0,72	13,27	0,34	12,36	0,25	12,64	0,25	12,90	0,40
D293			12,09	0,14	14,37	0,29	11,99	0,12	14,33	1,07	13,67	0,31
D294			14,58	1,04	12,55	0,60	12,84	0,43	16,32	0,84	12,18	0,47
V295			12,53	0,43	10,32	0,61	10,94	0,29	10,64	0,73	10,73	0,25
C296	16,73	0,35	20,90	2,70	15,60	1,64	16,67	1,54	11,80	0,23	11,42	0,31
Q297	13,11	0,09	13,62	1,24	13,98	0,47	10,99	0,73	13,44	0,76	11,93	0,25
R298			11,39	0,09	11,27	0,17	11,56	0,11	11,29	0,43	11,96	0,37
P299												
Y300	9,56	0,05	8,58	0,64	9,37	0,13	9,56	0,17	10,35	0,14	9,53	0,23
H301	15,95	0,49	12,69	1,50	13,83	0,81	14,72	0,65	16,67	0,51	14,25	0,99
W302	12,75	0,53	11,14	0,53	10,97	0,43	10,81	0,47	13,41	0,94	12,97	0,38
V303	12,64	0,61	13,25	0,63	12,40	0,88	12,34	0,34	12,63	0,33	14,27	0,13
C304	12,21	0,48	11,24	0,08	10,51	0,31	10,63	0,18	11,05	1,14	12,02	0,63
E305	13,50	0,19	11,79	0,48	11,29	0,97	11,85	0,13	11,89	0,56	13,01	0,55
A306	16,02	1,29	14,87	1,46	10,31	1,78	14,04	2,10	9,90	1,07	16,61	1,07

G307	10,96	0,59	9,84	0,41	10,70	0,34	10,84	0,86	11,42	0,27	11,64	0,34
L308	8,84	0,37	10,24	0,28	10,08	0,68	9,75	0,29	10,34	0,43	9,04	0,28
G309	9,19	0,10	7,10	0,58	8,59	0,16	7,66	0,24	8,40	0,44	8,02	0,39
Q310	8,09	0,07	11,61	0,45	10,55	0,34	13,51	0,47	12,00	0,16	7,29	0,07
T311	3,95	0,25	5,08	0,31	5,82	0,41	3,92	0,34	4,94	0,17	5,13	0,28
S312		0,37		0,65		0,26		1,69		0,16		0,13
Q313	2,63	0,24	7,45	0,56	5,30	0,55	10,88	0,19	2,70	0,09	2,66	0,14
E314	2,14	0,04	1,44	0,00	2,07	0,07	2,16	0,22	2,24	0,01	2,36	0,06
S315	3,03	0,13	1,66	0,09	1,97	0,20	6,81	1,64	1,68	0,06	2,15	0,05
H316	1,49	0,01	1,60	0,03	1,57	0,02	2,05	0,01	1,47	0,01	1,64	0,03

7.11- Valores de R_2/R_1 calculados para cada um dos resíduos na presença dos ligandos em estudo

Tabela 13- Valores calculados do rácio R_2/R_1 para cada resíduo do domínio MGL-CRD na forma livre e na presença dos complexos MGL-CRD/1-5.

	MGL-CRD	MGL-CRD/1	MGL-CRD/2	MGL-CRD/3	MGL-CRD/4	MGL-CRD/5
	R_2/R_1	R_2/R_1	R_2/R_1	R_2/R_1	R_2/R_1	R_2/R_1
M180						
C181						
P182						
V183	11,50	15,07	12,42	14,34	12,32	11,85
N184	11,43	10,67	10,31	10,30	11,52	12,43
W185	11,95	10,08	13,82	10,19	15,82	14,83
V186	10,17	13,43	13,98	10,05		7,77
E187			9,59	8,11	13,67	11,84
H188	6,41	9,63	7,24	7,09	14,97	9,34
Q189						
D190						
S191	9,45	13,70	9,95	9,25		11,57
C192		8,46	9,46	11,91	11,35	
Y193	10,48	9,05	12,16	12,39	13,32	9,22
W194	7,95	8,36	8,09	7,30	6,62	7,74
F195	9,86	10,18		11,94	9,37	10,91
S196	10,56	9,04	8,68	9,79		
H197						
S198	7,14	7,78	7,70	8,98	8,20	7,97
G199	8,41	9,03	6,28	6,47	8,12	8,93
M200	8,71	7,70	9,20	9,33	7,43	8,21
S201	8,03	7,39	8,01		7,65	7,70
W202	11,87	10,14	10,31	10,39	10,92	10,29
A203	11,07	11,53	10,75	10,29	11,15	9,91
E204	10,31	10,58	10,22	10,45	10,51	11,26
A205	11,75	12,99	13,66	11,35	11,24	13,42
E206	4,28	5,10	6,12	5,31	3,64	6,33
K207	11,78	11,39	11,54	12,78	12,09	10,84
Y208	11,92	11,91	11,79	13,22	13,25	11,47
C209		8,98	12,59	12,79	12,87	11,76
Q210						
L211	11,12	11,90	9,47	9,69	11,24	9,78
K212	10,24	11,15	9,89	10,27		11,24
N213	8,69	9,14	10,05	9,26	10,85	10,01
A214	9,15	9,41	10,57	10,76	11,30	9,17
H215	10,38	8,00	9,47	9,13	10,94	9,63

L216	8,30	7,47	11,58	9,71	6,77	10,60
V217	9,39	10,60	9,20	7,72	9,87	8,96
V218		12,61	11,35	13,11	16,30	13,03
I219	10,90	14,27				10,16
N220	10,07	10,53		9,18	11,78	9,83
S221	8,19	8,76	9,29	8,09	9,77	9,26
R222	8,74	7,52	7,37		7,81	
E223	8,10	8,19	8,41	8,01	9,87	7,56
E224	9,49	9,65	1,20		7,53	6,89
Q225	8,94	9,50	9,55	9,27	9,45	9,69
N226	8,13	10,06	9,04	8,33	9,62	8,75
F227	9,26	9,54	8,47	9,10	11,04	10,79
V228	8,13	9,95	7,61	8,19	7,90	7,20
Q229	7,07	8,62	9,27	9,14	8,39	8,53
K230	9,26	8,67	9,95	9,20	10,58	9,25
Y231	8,89	9,09	8,22	8,12	10,76	9,19
L232	8,48		10,13	10,62	10,12	9,30
G233	6,22	7,48	6,99	7,46	8,38	6,61
S234	7,77	6,51	8,11	7,47		7,76
A235		9,16			8,09	
Y236	14,87	8,72	6,50	7,83	9,67	6,26
T237	8,83	10,36	8,99	10,65	9,35	8,81
W238	9,02	11,03	8,65	6,66	8,04	9,44
M239	10,53	10,87		12,66	13,04	
G240	10,61	11,78	9,60	7,05	9,47	8,24
L241	10,04	12,17	10,35	8,32	9,53	8,47
S242	8,53	10,07	11,32	10,62	12,24	9,91
D243	9,49	9,03	9,25	9,87	10,64	9,83
P244						
E245	10,09			8,84	9,97	9,20
G246	8,47	9,37			10,31	9,12
A247	12,30	10,19	9,31	9,68	9,79	9,41
W248	9,28	10,01	10,17	9,42	9,61	9,86
K249	9,78	10,93	10,17	8,89	12,02	9,17
W250	8,05		9,31	8,98	11,24	9,13
V251	10,91	8,83	10,65	9,72		8,97
D252	8,74	10,10	10,57	10,43	11,40	10,48
G253	8,98	8,99	10,18	9,14	11,30	11,34
T254	11,41	11,88	10,90	12,50		12,00
D255	13,61	16,93	12,50			14,89
Y256	8,65	9,52	7,79	7,74	8,91	8,61
A257	7,48	9,34	8,64	8,10	8,79	8,07
T258	7,62	9,30	7,53	7,46	9,01	7,41
G259	8,51	10,83	8,19	8,41	10,18	8,38

F260	9,50	9,71	9,43	9,63	10,55	9,83
Q261	7,98	9,30	9,57	8,07	9,63	9,76
N262						
W263	7,86	9,09	8,71	8,90	9,56	10,42
K264	9,83	9,35	9,41	6,87		6,52
P265						
G266						
Q267	9,77	10,85	8,67	12,68	9,36	11,93
P268						
D269		9,98	8,85	8,61	9,62	8,34
D270		14,20		7,03	10,07	8,67
W271	12,03	10,95		8,32	10,55	10,03
Q272		11,07	11,89	10,45	8,69	9,99
G273						
H274						
G275		11,93	10,62		12,49	9,00
L276		10,38	10,69		10,99	9,87
G277	6,84	9,02	8,94	8,38	9,21	8,66
G278			9,98	11,82		8,34
G279	8,84	10,09	8,43	10,41	11,09	9,25
E280	14,95	8,55	9,05	7,81	9,11	8,29
D281	9,47	9,27	8,74	9,57	11,95	8,70
C282	12,44	12,57	9,85	9,48	12,76	10,11
A283	9,41	9,64	10,72	8,05	10,32	9,74
H284	13,38	10,56	9,51	9,79		10,53
F285		11,56	13,32	10,95	7,23	10,15
H286	14,54	8,26		10,61	9,47	7,80
P287						
D288		8,85	10,71	10,58	10,15	
G289	8,74	8,04	8,06	8,77	12,12	7,29
R290	8,41	9,46			4,95	
W291	10,27	8,42	8,19	7,95	7,84	9,22
N292		11,62	9,58	8,54	11,19	10,69
D293		10,77	12,16	11,97	13,58	9,47
D294		13,38	10,74	9,62	15,59	10,97
V295		9,37	8,45	6,83	7,72	7,32
C296		20,22	14,40		8,87	6,90
Q297	9,60	10,07	9,50	8,00	10,44	8,51
R298		9,03	7,99	6,60	8,10	8,28
P299						
Y300	7,20	7,44	8,65	7,33	7,14	6,63
H301	12,77	10,54	10,65	12,49	8,32	12,35
W302	9,33	8,83	8,19	7,26	10,41	9,52
V303	10,51	10,95	8,90	7,54	11,99	11,79

C304	9,77	7,78	10,70	11,86	10,69	10,13
E305	9,09	8,34	11,06	10,66	13,78	11,25
A306	7,59	15,12	5,60			
G307	6,93	6,71	8,07	9,84	7,64	10,87
L308	5,63		6,96	6,29	9,25	6,05
G309	6,10	4,79			5,47	
Q310	8,74	10,71		12,37		6,05
T311	3,96		3,77		3,36	
S312						
Q313			4,30			
E314						
S315						
H316						

7.12- Histogramas dos valores de relaxação HetNOE, R_1 e R_2 da MGL-CRD na forma livre e na presença dos vários ligandos em estudo

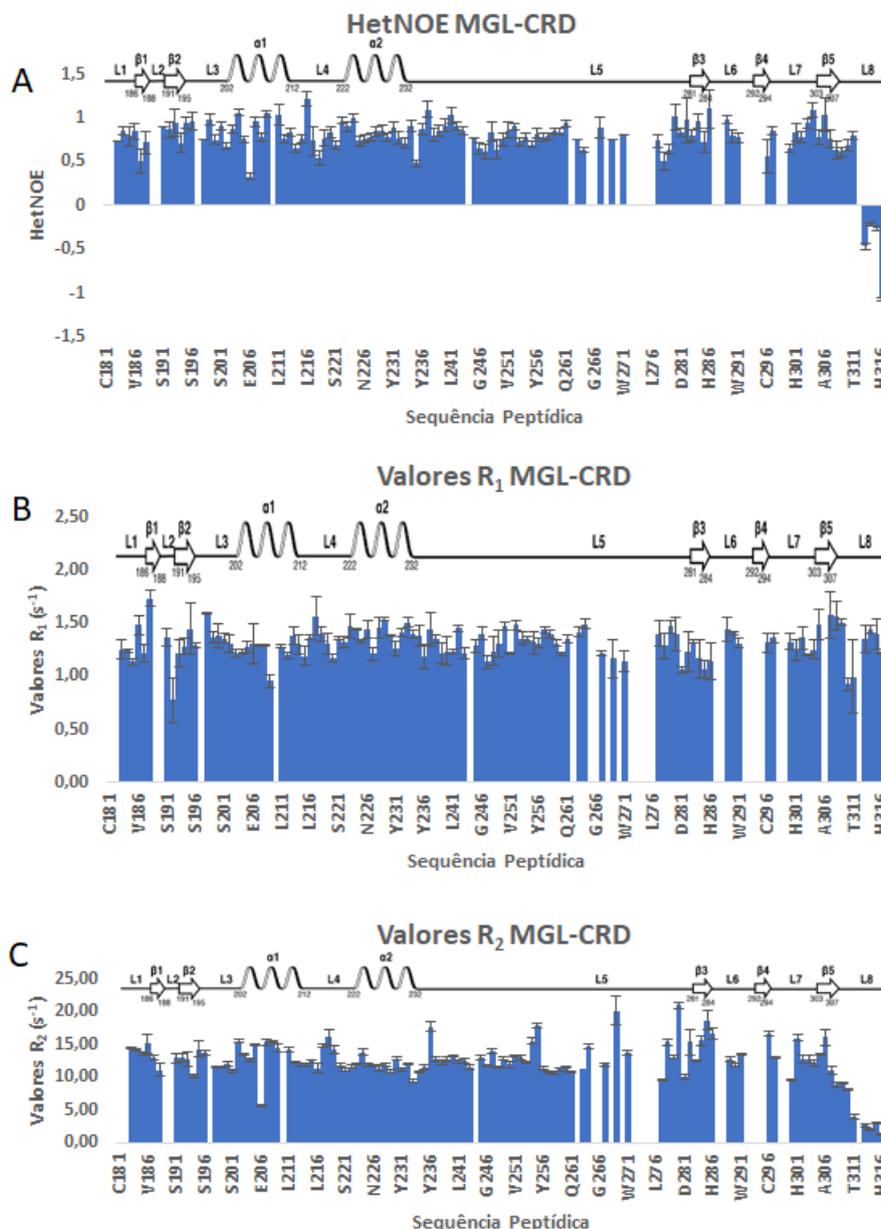


Figura 37: A- Histograma dos valores de HetNOE para cada um dos resíduos para o domínio MGL-CRD. B- Histograma dos valores de R_1 para cada um dos resíduos para o domínio MGL-CRD. C- Histograma dos valores de R_2 para cada um dos resíduos para o domínio MGL-CRD. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD.

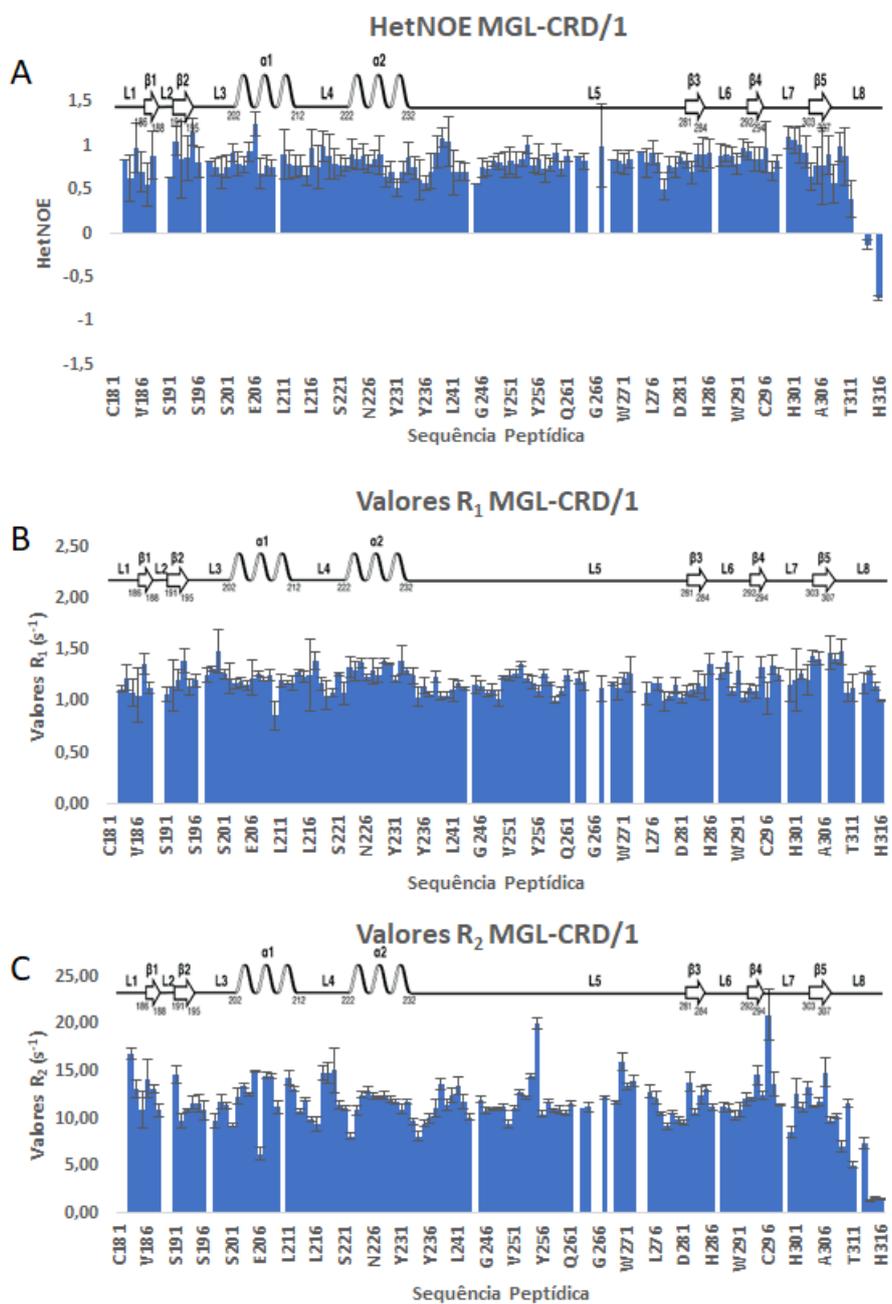


Figura 38: **A-** Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/1. **B-** Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/1. **C-** Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/1. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD.

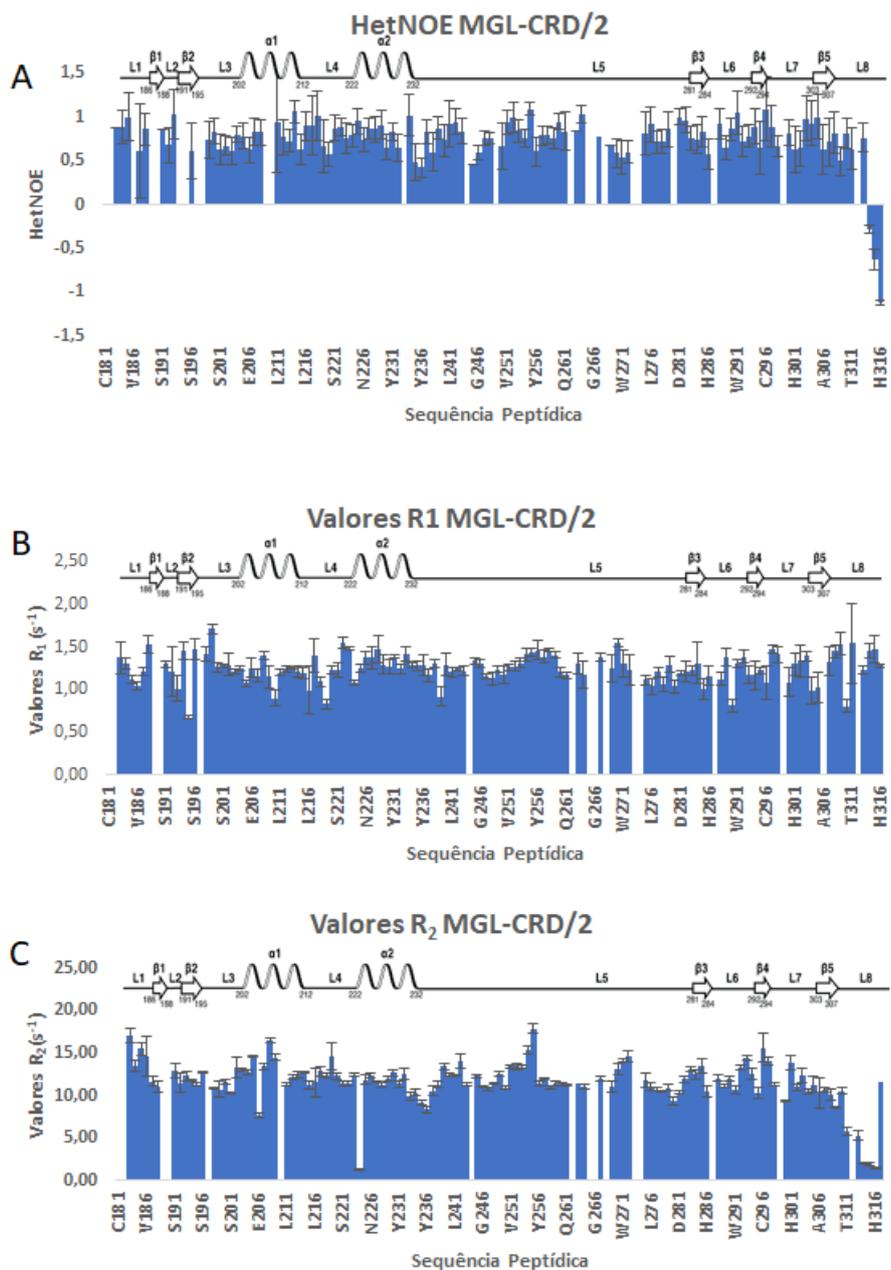


Figura 39: A- Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/2. B- Histograma dos valores de R₁ para cada um dos resíduos para o complexo MGL-CRD/2. C- Histograma dos valores de R₂ para cada um dos resíduos para o complexo MGL-CRD/2. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD.

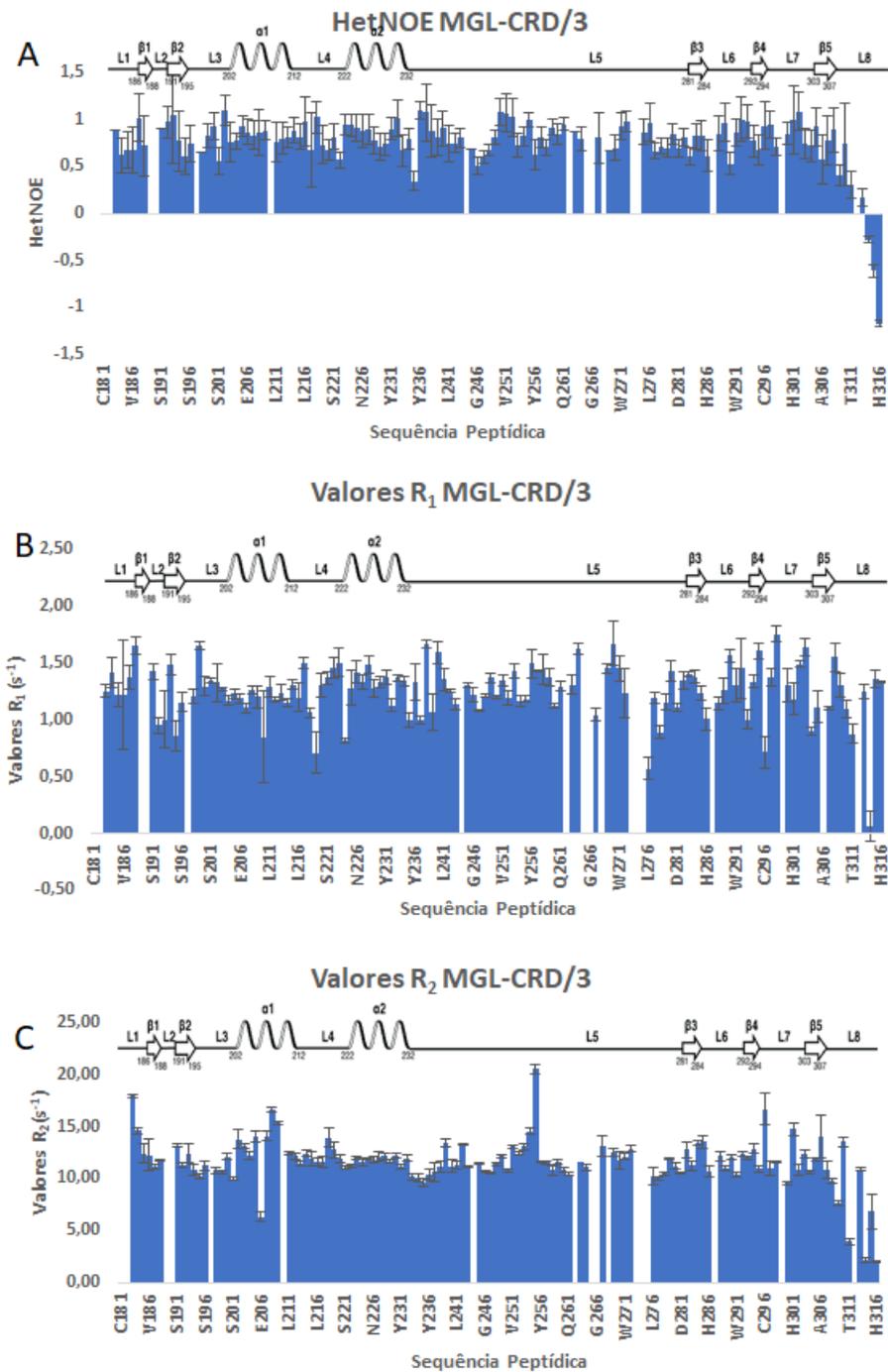


Figura 40: A- Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/3. B- Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/3. C- Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/3. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD.

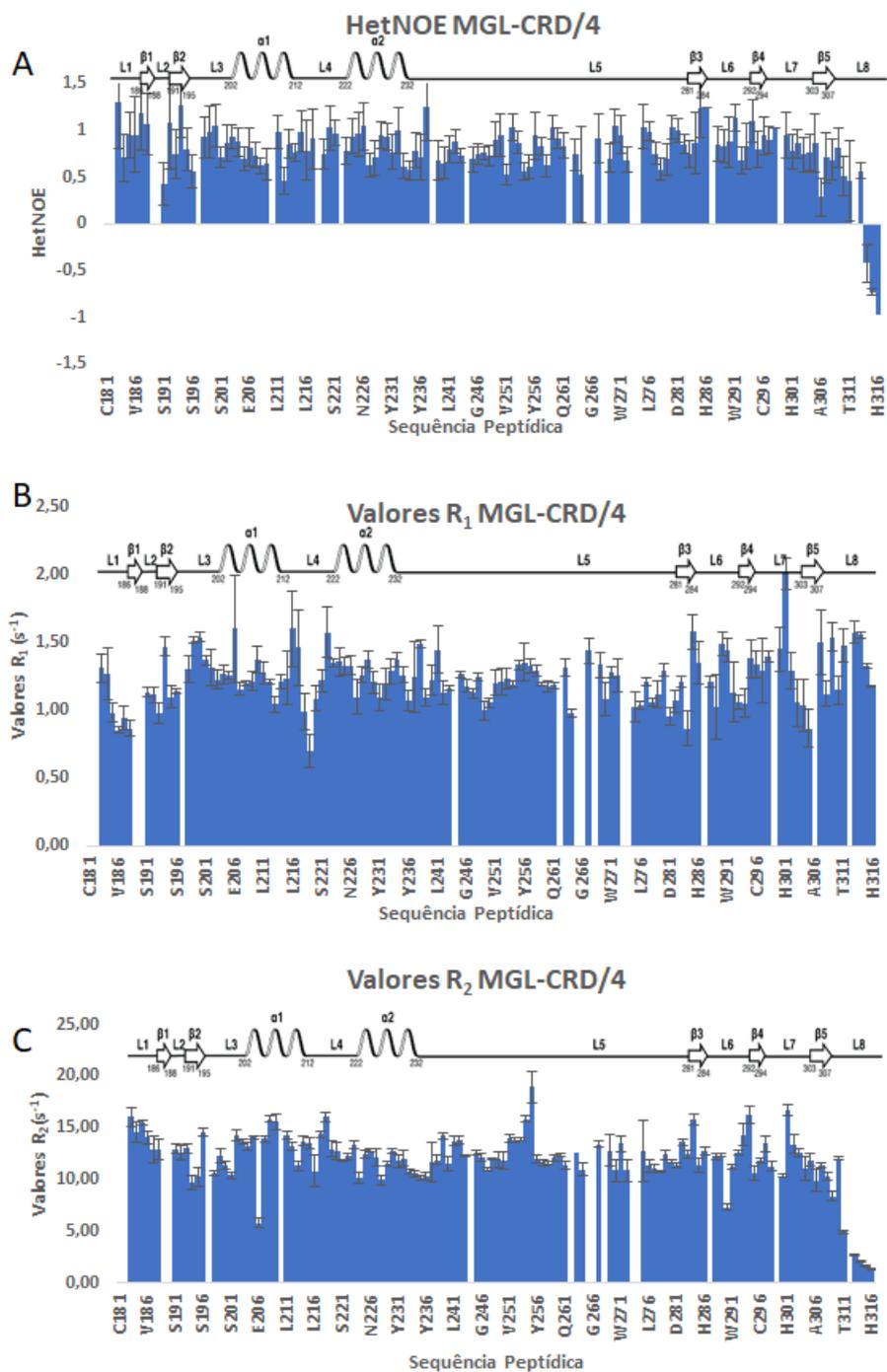


Figura 41: A- Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/4. B- Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/4. C- Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/4. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD.

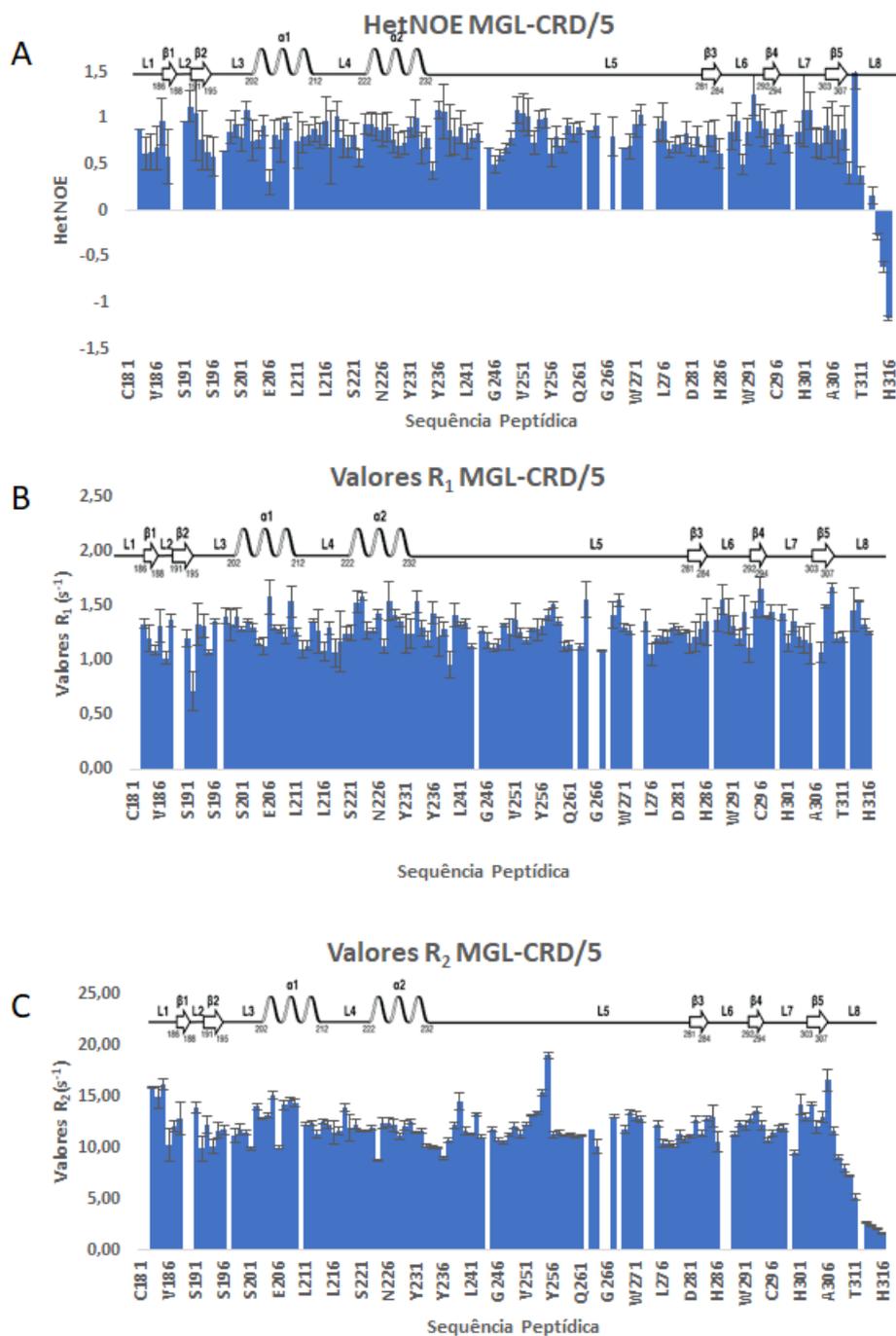


Figura 42: **A-** Histograma dos valores de HetNOE para cada um dos resíduos para o complexo MGL-CRD/5. **B-** Histograma dos valores de R_1 para cada um dos resíduos para o complexo MGL-CRD/5. **C-** Histograma dos valores de R_2 para cada um dos resíduos para o complexo MGL-CRD/5. Encontra-se representado nos histogramas o esquema da estrutura secundária do modelo de homologia do domínio MGL-CRD.

