



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Campus de São José do Rio Preto

THIAGO SALEM PANÇONATO TEIXEIRA

CARACTERIZAÇÃO BIOFÍSICA DA INTERAÇÃO ENTRE A
PROTEÍNA M2-1 DO VÍRUS SINCICIAL RESPIRATÓRIO
HUMANO (HRSV) COM QUERCETINA.

SÃO JOSÉ DO RIO PRETO - SP
2016

Thiago Salem Pançonato Teixeira

CARACTERIZAÇÃO BIOFÍSICA DA INTERAÇÃO ENTRE A
PROTEÍNA M2-1 DO VÍRUS SINCICIAL RESPIRATÓRIO
HUMANO (HRSV) COM QUERCETINA.

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Microbiologia, junto ao Programa de Pós-Graduação em Microbiologia, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Orientadora: Prof^a. Dr^a Fátima Pereira de
Souza

SÃO JOSÉ DO RIO PRETO - SP
2016

Teixeira, Thiago Salem Pançonato.

Caracterização biofísica da interação entre a proteína M2-1 do Vírus Sincicial Respiratório Humano (HRSV) com quercetina / Thiago Salem Pançonato Teixeira. -- São José do Rio Preto, 2016
98 f. : il., tabs.

Orientador: Fátima Pereira de Souza

Tese (doutorado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas

1. Virologia. 2. Vírus Sincicial Respiratório Humano. 3. Dinâmica molecular. 4. Biofísica. 5. Quercetina. I. Souza, Fátima Pereira de. II. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU – 576.858

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Câmpus de São José do Rio Preto

Thiago Salem Pançonato Teixeira

CARACTERIZAÇÃO BIOFÍSICA DA INTERAÇÃO ENTRE A
PROTEÍNA M2-1 DO VÍRUS SINCICIAL RESPIRATÓRIO
HUMANO (HRSV) COM QUERCETINA.

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Microbiologia, junto ao Programa de Pós-Graduação em Microbiologia, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Comissão Examinadora

Prof^a. Dr^a. Fátima Pereira de Souza
UNESP – São José do Rio Preto
Orientadora

Prof^a. Dr^a. Alessandra Vidotto
FAMERP

Dr. Fábio Rogério de Moraes
UNESP – São José do Rio Preto

Prof^a. Dr^a. Natália Bueno Leite Slade
UNILAGO – São José do Rio Preto

Prof. Dr. Marcio José Tiera
UNESP – São José do Rio Preto

SÃO JOSÉ DO RIO PRETO - SP
2016

AGRADECIMENTOS

Agradeço a Deus pela oportunidade de vivenciar a realidade do pesquisador, concedendo-me força, atenção e paciência, principalmente quando o cansaço e a saudade de casa sobrepunham meus ânimos.

À minha professora e orientadora Dr^a. Fátima Pereira de Souza que acreditou em mim para desenvolver esse trabalho e, me ensinou a ser visionário e ter esperanças e forças ao visualizar o final e não o meio.

Ao professor Dr. Marcelo Andrés Fossey pelos ensinamentos, atenção e ajuda nos experimentos de fluorescência e nos tratamentos dos resultados dos mesmos.

Ao professor Dr. Mario Sérgio Palma por conceder a seu laboratório para a realização do sequenciamento e à sua orientanda Ma. Nathalia Baptista Dias por dispor do seu tempo para a realização do sequenciamento da proteína.

Ao professor Dr. Altair Benedito Moreira por ceder o espectrofluorímetro, essencial para a realização desse trabalho. Ao professor Dr. Márcio Francisco Colombo por ceder o sistema de purificação ÄKTA, sem o qual esse trabalho não teria o mesmo dinamismo.

À FAPESP pelo apoio financeiro para o desenvolvimento desta pesquisa, processo: 2013/2435502.

À minha mãe Maria Vitória por ser o sopro que me move, o coração a compartilhar a dor e a alegria. A amiga que esteve presente com todo amor, mesmo a quilômetros de distância.

Ao meu pai Antônio Pançonato que me ensinou a resignar e ser paciente diante às dificuldades que às vezes o destino impõe.

A minha tia Gabriela, minha segunda mãe, que com seu amor acompanhou meus primeiros passos me ensinando a rir das dificuldades.

Aos meus irmãos Henrique e Hamilton que são exemplos e reflexões que sempre me surpreendem. Mais que irmãos e amigos, são um destino a ser acompanhado.

Ao meu namorado Tiago que sempre me apoiou nessa jornada, tornando o retorno ao lar um momento de alegria.

À minha amiga Deriane que me acolheu em amor fraternal e me tutelou no início nessa nova jornada.

À minha amiga Gabriela que em sua mansuetude me acalentou em momentos turbulentos, além de dividir espaço de trabalho e conhecimentos comigo.

Aos demais amigos e colegas Ana Karla, Anwar, Carlos, Cíntia, Daiane, Danúbia, Fernanda, Hana, Hêmily, Ícaro, Isabella, Karoline, Luciana, Mariana, Mônica, Natalia, Paulo, Pedro, Raphael, Ricardo, Taísa, Vitor e Yuri que colaboraram ora no crescimento profissional ora no pessoal.

Aos funcionários do laboratório e do departamento, que cada qual com sua função, foi uma mão a mais nesse trabalho, seja na limpeza do chão, no preparo de uma solução ou no preparo do café.

Sumário

1.Introdução	9
1.1. Morfologia	10
1.2. Função das proteínas do hRSV	10
1.2.1. Estrutura e função da proteína M2-1.	13
1.3. Terapêutica da infecção por hRSV	16
1.3.1. Quercetina.	18
2.Objetivos	20
3.Materiais e Métodos	21
3.1. Clonagem do gene M2-1 do hRSV com sítios de restrição para as enzimas EcoRI e BamHI.	21
3.1.1. Reação de PCR e eletroforese	22
3.2. Clonagem com o TOPO-XL PCR Cloning kit (Invitrogen®).....	23
3.2.1. Purificação dos produtos de PCR.....	23
3.2.2. Reação de ligação TOPO®	24
3.2.3. Transformação nas células competentes (transformação química)	24
3.2.4. Purificação dos plasmídeos para confirmação da clonagem.	25
3.2.5. Confirmação da clonagem por co-digestão enzimática.	25
3.3. Subclonagem no vetor pGEX4T-1.	25
3.4. Transformação em E. coli linhagem BL21 Gold (DE3).....	28
3.5. Sequenciamento da proteína M2-1.	28
3.6. Indução da expressão da proteína M2-1.....	28
3.7. Purificação da proteína M2-1 fusionada com GST.....	28
3.8. Medida da concentração da proteína.	29
3.9. Preparo das amostras.	29
3.10. Análise estrutura secundária por dicroísmo circular.	29
3.11. Supressão de fluorescência.....	30
3.11.1. Medidas espectrofotométricas e equipamentos	30
3.12. Análise de Bioinformática.....	33
3.12.1. Estrutura molecular.....	33
3.12.2. Dinâmica molecular.	33

3.12.3. Docking molecular.	36
4.Resultados e Discussão	38
4.1. Dados da proteína M2-1 obtida pelo servidor ProtParam.	38
4.2. Clonagem.....	38
4.2.1. Clonagem do gene M2-1 com sítio de restrição para EcoRI e BamHI. 38	
4.3. Expressão.	39
4.4. Purificação	39
4.4.1. Purificação com o sistema ÄKTA da proteína M2-1 fusionada com GST. 39	
4.5. Sequenciamento da estrutura primária da proteína M2-1.	41
4.6. Análise por espectroscopia de dicroísmo circular.	41
4.7. Supressão de fluorescência	43
4.8. Bioinformática	50
4.8.1. Predição estrutural.....	50
4.8.2. Dinâmica molecular.	55
4.8.3. Docking molecular.	61
5.Conclusões.....	68
6.Referências	70
Anexo I	79
Anexo II	81
Anexo III	82
Anexo IV.....	83
Apêndice	84

Lista de Figuras

Figura 1: Mapa gênico do hRSV.	10
Figura 2: Ciclo de infecção do hRSV.....	11
Figura 3: Estrutura cristalina da proteína M2-1 (PDB - 4C3B).....	15
Figura 4: Estrutura básica dos flavonoides (A) e da Quercetina (B).....	18
Figura 5: Mapa do vetor pCR-XL-TOPO (Invitrogen).	23
Figura 06: Mapa gênico do vetor pGEX-4T-1 (GE Healthcare Life Sciences)..	26
Figura 7: Reação de PCR do gene M2-1.	38
Figura 8: SDS-PAGE 15% corado com Comassie Brilliant Blue G-250 referente ao teste de indução dos clones da <i>E. coli</i> BL21 God + pGEX4T-1 + M2-1.	39
Figura 9: Cromatograma da purificação da proteína M2-1 no sistema ÄKTA (GE HEALTHCARE LIFE SCIENCES) com a resina Superdex 200 10/300 GL.	40
Figura 10: Resultado da purificação da proteína M2-1 em SDS-PAGE 15% corado com Coomassie Blue G-250.....	40
Figura 11: Sequenciamento da proteína M2-1.	41
Figura 12: Espectro de dicroísmo circular da proteína M2-1.....	42
Figura 13: Desnaturação térmica da proteína M2-1.....	43
Figura 14: Decaimento da fluorescência da proteína M2-1 com o aumento da temperatura.....	44
Figura 15: Espectros de emissão de fluorescência referente à titulação de M2-1 na concentração de 10 μ M.....	45
Figura 16: Fração de variação da fluorescência observada (fobs) em função da razão entre as concentrações de Quercetina e da M2-1.....	46
Figura 17: Gráfico de Scatchard relativo à interação da M2-1 com a Quercetina.	48
Figura 18: Gráfico de Van'Hoff referente à interação da proteína M2-1 com a Quercetina.....	49
Figura 19: Dependência de ΔG , ΔH e $T\Delta S$ com a temperatura, mostrando a diferença de interação entre a Quercetina e a M2-1 para os sítios do tipo 1 e tipo 2.	50
Figura 20: Modelos da proteína M2-1 gerado pelo servidor I-Tasser.....	52
Figura 21: Modelos dos oligômeros da proteína M2-1 geradas pelo servidor SymmDock.....	54

Figura 22: RMSD (A) e raio de giro (B) do <i>backbone</i> da proteína M2-1 monomérica em função do tempo à 313K.....	55
Figura 23: RMSD (A) e raio de giro (B) do <i>backbone</i> da proteína M2-1 dimérica em função do tempo à 313K.	56
Figura 24: RMSD (A) e raio de giro (B) do <i>backbone</i> da proteína M2-1 trimérica em função do tempo à 313K.	57
Figura 25: RMSD (A) e raio de giro (B) do <i>backbone</i> da proteína M2-1 tetramérica em função do tempo à 313K.	57
Figura 26: RMSF da proteína M2-1 monomérica (A), dimérica (B), trimérica (C) e tetramérica (D) em função do tempo à 313K.....	58
Figura 27: Gráfico padrão da estrutura secundária da proteína M2-1 monomérica (A) e, cadeias A da proteína dimérica (B), trimérica (C) e tetramérica (D).dimérica no solvente água.....	60
Figura 28: Sobreposição da estrutura obtida pela dinâmica molecular com a estrutura cristalina da proteína M2-1.....	61
Figura 29: Domínios da proteína M2-1.....	62
Figura 30: Docking molecular da Quercetina na proteína M2-1 monomérica. Quercetina foi docada no sítio 1 (A) (domínio de ligação ao RNA e à proteína P e resíduo N-terminal) e no sítio 2 (B) (domínio dedo de zinco e de oligomerização).	63
Figura 31: <i>Docking</i> molecular da Quercetina na proteína M2-1 tetramérica. ...	65

Lista de Tabelas

Tabela 1: Sequências utilizadas para alinhamento e desenho de <i>primers</i> do gene M2-1	21
Tabela 2: Enzimas de restrição e seus respectivos sítios de reconhecimento. 21	
Tabela 3: Primers para os genes RSV tipo A.....	22
Tabela 4: Reação de co-digestão do plasmídeo pCR-XL-TOPO + M2-1 com as enzimas BamHI e EcoRI.	25
Tabela 5: Protocolo de reação de ligação do plasmídeo pGEX-4T-1 com o inseto M2-1 na razão 3:1.....	27
Tabela 6: Protocolo de reação de ligação do plasmídeo pGEX-4T-1 com o inseto M2-1 na razão 3:1.....	27
Tabela 7: Valores de K1, constante de associação do sítio do tipo 1 e K2, constante de associação do sítio do tipo 2, nas três temperaturas estudadas. 48	
Tabela 8: Valores de ΔG , ΔH e $T\Delta S$ referentes à interação da proteína M2-1 com a Quercetina para os dois tipos de sítios.....	50
Tabela 9: Valores da energia de interação (kcal/mol) da Quercetina nos sítios 1, 2 e 3 do monômero e do tetrâmero.	66

Lista de abreviações

Å	Angström ($1 \times 10^{-10} \text{m}$)
CCA	Agente da coriza de chimpanzé (do inglês: chimpanzee coryza agent)
CD	Dicroísmo Circular
DSSP	Biblioteca de estrutura secundária (do inglês: database secondary structure assignments)
ϵ	Coeficiente de extinção molar
GE	gene de parada (do inglês: gene end)
GS	gene de iniciação (do inglês; gene start)
GST	Glutathione-S-transferase
HBeAG	Antígeno e do vírus da hepatite B
HBsAG	Antígeno de superfície do vírus da hepatite B
hRSV	Vírus sincicial respiratório humano (do inglês: human respiratory syncytial vírus)
INF	Interferon
IPTG	Isopropyl β -D-1-thiogalactopyranoside
kDa	Kilo daltons
K _{sv}	Constante de Stern-Volmer
m.A.U.	Mili unidades de absorbância
MD	Dinâmica Molecular
mRNA	RNA mensageiro
nm	Nanômetro ($1 \times 10^{-9} \text{m}$)
pb	Pares de bases
PDB	Protein Data Bank
PME	Particle Mesh Ewald
RMSD	Desvio Quadrático Médio
RMSF	Flutuação média quadrática
STAT2	Sinal de tradução e ativador da transcrição 2 (do inglês: signal transducer and activator of transcription 2)

Resumo

O Vírus Sincicial Respiratório Humano (hRSV) é o principal agente causador de doenças respiratórias agudas severas, como pneumonia e bronquiolite. Trata-se de um vírus da família *Paramyxoviridae*, com simetria helicoidal e envelope lipídico. Possui um genoma de RNA fita simples, não segmentado, com 10 genes codificantes. Um dos fatores que contribuem para o sucesso na replicação viral é a interação da proteína M2-1 com as proteínas P, M e com o RNA. Tal proteína age através da interação com RNA, e assim, participa efetivamente no processo de replicação viral, apresentando atividade anti-parada da replicação sendo ativa quando na forma tetramérica. Deste modo, o ligante que realizar interação com a M2-1 é um potencial candidato a prevenir as interações M2-1 com RNA e demais proteínas, impedindo deste modo o sucesso da replicação viral. Flavonoides são compostos naturais com várias atividades como antioxidante, anticancerígena, cardio protetora, antibacteriana e antiviral. Quercetina é um flavonoide que apresenta atividade antiviral, inibindo o complexo de replicação. Com a finalidade de investigar a interação de um potencial candidato a inibição da atividade da M2-1, clonamos, expressamos e purificamos a proteína M2-1 do hRSV e, determinamos sua característica físico-química de interação com quercetina em diferentes condições de temperatura. Para clonagem e expressão do gene M2-1 foram escolhidos o vetor pGEX-4T-1 e a bactéria *Escherichia coli* da linhagem BL21 Gold (DE3). A proteína expressa foi purificada em cromatografia de afinidade e cromatografia de exclusão molecular. As análises de composição de estrutura secundária e estabilidade térmica foram realizadas por espectroscopia de dicroísmo circular e, a interação com Quercetina com espectroscopia de fluorescência. Para o mapeamento da interação com a Quercetina foram realizadas *docking molecular* e dinâmica molecular. A proteína M2-1 purificada apresenta 40% α -hélice, 10% folhas β e 19% alças e 31% *random coils*, sendo a temperatura de *melting* de 55,6°C. O teste de titulação com o flavonoide Quercetina mostrou que ela interage em dois pontos, no sítio 1 (constante de afinidade $\sim 1,5 \times 10^6 \text{ M}^{-1}$) estabelece uma interação do tipo iônica e no sítio 2 ($1,1 \times 10^5 \text{ M}^{-1}$) estabelece uma interação do tipo hidrofóbica. Uma molécula de Quercetina interage no sítio 1 (formada pelo tetrâmero) e quatro moléculas interagem no sítio 2 (uma por unidade proteica). O estudo de dinâmica molecular mostra que o domínio central (composto

de α -hélice) é estável e, a extremidade N-terminal adquire maior estabilidade quando a proteína está na forma tetramérica. O *docking* do ligante Quercetina mostra que ela pode interagir na face N-terminal com interação do tipo iônica e entre cadeias próxima aos domínios de oligomerização e ligação ao RNA e proteína P, com tipo de interação hidrofóbica. Os dados teóricos reforçam os dados experimentais, mostrando que a quercetina interage com a proteína M2-1, principalmente no poro central da proteína.

Palavra-chave: hRSV, M2-1, Quercetina, dinâmica molecular.

Abstract

Human Respiratory Syncytial Virus (hRSV) is the major causative agent of acute respiratory infections such as pneumonia and bronchiolitis. This is a virus of the *Paramyxoviridae* family with helical symmetry and lipid envelope. It has a single strand RNA genome, not segmented, with 10 coding genes. One of the factors that contribute to success in viral replication is the interaction of M2-1 protein with P, M protein and RNA viral. Such protein acts through the interaction with RNA, hence, participating in the viral replication process, having anti-stop activity in the replication process and, the protein is active only as a tetramer. Thus, a ligand that interacts with M2-1 is a potential candidate to prevent interaction of M2-1 with other proteins, thereby preventing the success of viral replication. The aim of the research is to clone, express and purify the M2-1 protein of hRSV and then determine its physicochemical characteristics of the interaction with Quercetin in different temperatures to investigate the interaction of a potential candidate to inhibit the activity of the M2-1. The M2-1 gene was cloned in vector pGEX-4T-1 and expressed in *Escherichia coli* strain BL21 Gold (DE3). The expressed protein was purified on affinity chromatography molecular exclusion chromatography, thermal stability and secondary structure were analyzed by circular dichroism and the interaction with Quercetin was analyzed by fluorescence spectroscopy. To understand the interaction mechanism was performed molecular dynamics and molecular docking. The purified protein has 40% α -helix, 10% β -sheet, 19% turns and 31% random coils, being the melting temperature of 55,6°C. The titration of M2-1 with Quercetin showed that it interacts in two points, on site 1 (the binding affinity is $\sim 1,5 \times 10^6 \text{ M}^{-1}$) and establishes an ionic interaction and, on site 2 ($1,1 \times 10^5 \text{ M}^{-1}$) establishes a hydrophobic interaction. One Quercetin molecule interacts on site 1 (formed by the tetramer) and four Quercetins interacts on site 2 (one per protein unit). The molecular dynamics study shows that the central domain (composed mainly by α -helix) is stable and, the N-terminal extremity gets more stabilization when the protein is in the tetrameric form. The molecular docking of Quercetin shows that it can interact on N-terminus face with ionic interaction pattern and between chains close to oligomerization and RNA and P protein interaction domains with a hydrophobic interaction pattern. Hence, this ligand could be inhibiting the protein activity by disrupting it.

Key-words: hRSV, M2-1, Quercetin, molecular dynamics.

1. Introdução

Em 1955, foi isolado de chimpanzés um vírus presente em secreções nasais mucopurulentas, que foi denominado agente da coriza do chimpanzé (Chimpazee coryza agent - CCA) (Blount et al., 1956). Posteriormente Robert J. Chanock isolou o CCA de dois bebês prematuros com bronquiolite e pneumonia. Esses vírus foram isolados em cultura de células e produziram características peculiares como células gigantes multinucleadas com grande sincício e, assim, o vírus passou a ser denominado vírus sincicial respiratório (respiratory syncytial virus – RSV) (Chanock et al., 1957, 1957b).

A transmissão do RSV ocorre primariamente pelo contato direto de secreções respiratórias com a mucosa da nasofaringe ou pela entrada de grandes gotículas respiratórias no nariz ou olho (Hall et al, 1975, 1981, 1981b). O vírus pode permanecer viável em superfícies sólidas por até 6 horas, em luvas de borracha por 90 minutos e na pele por 20 minutos (Hall et al., 1980). A constante higienização das mãos é necessária quando se está em contato com pacientes contaminados.

Após o contágio, a replicação começa no epitélio nasofaríngeo, mas se propaga para o epitélio bronquiolar onde a replicação é mais eficiente (Othumpangat et al., 2009). A replicação viral é seguida por necrose e edema da submucosa. A secreção de muco aumenta em quantidade e viscosidade. O resultado final é o bloqueio parcial das vias respiratórias e aumento da resistência respiratória causando o aprisionamento do ar (Aherne et al., 1970).

De acordo com estimativas da Organização Mundial da Saúde (OMS), o hRSV é responsável por mais de 160.000 mortes por ano. Só nos Estados Unidos da América, de 85.000 a 144.000 crianças são hospitalizadas anualmente com infecção por hRSV (Shay et al., 1999). No Brasil, no ano de 2005 na cidade de Ribeirão Preto, a prevalência de crianças infectadas com hRSV foi de 42% (Gagliardi, et al., 2013) e, no ano de 2004 a 2008 em Fortaleza – CE, a prevalência foi de 15,8% (Moura et al., 2013).

Infecções prévias com hRSV não garantem imunidade permanente (Walsh et al., 2004; Hall et al., 1991). O primeiro e segundo episódio da infecção

ocorre nos dois primeiros anos de vida e tendem ser mais severos (Glezen et al., 1986). A maioria das infecções subsequentes ocorre no trato respiratório superior e tem um curso moderado, embora a doença possa progredir para infecção do trato respiratório inferior com sintomas mais severos (Henderon et al., 1979; Hall et al., 2001).

1.1. Morfologia

O hRSV é classificado na ordem *Mononegavirales*, família *Paramyxoviridae*, subfamília *Paramyxovirinae*, gênero *Pneumovirus*. São vírus envelopados com nucleocapsídeo helicoidal (Santos et al., 2008). O genoma viral é constituído de RNA fita simples, linear, de polaridade negativa, não segmentado de aproximadamente 15kb, que contém 10 genes na seguinte ordem 3'-NS1-NS2-N-P-M-SH-G-F-M2-L-5', mas produz onze proteínas devido à presença de duas seqüências de leitura aberta do gene da proteína de matriz M-2 (M2) que produz as proteínas M2-1 e M2-2 (Figura 1). Dessas, duas proteínas não estruturais (NS1 e NS2) são responsáveis pela evasão da resposta imune; das oito proteínas estruturais, cinco (N, P, L, M2-1 e M2-2) formam o complexo nucleocapsídeo envolvido na replicação e transcrição viral e, quatro (SH, M, G e F) estão associadas à membrana, sendo que as proteínas G e F são glicoproteínas envolvidas na adesão e fusão do vírus com a célula hospedeira (Peter et al., 2008).

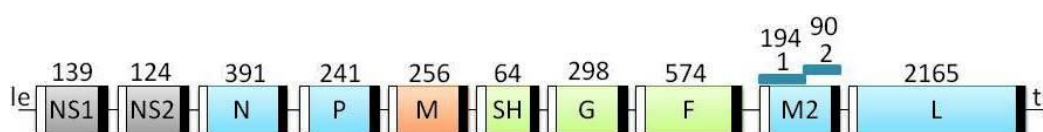


Figura 1: Mapa gênico do hRSV. O genoma sentido negativo é representado fora de escala. Os genes que codificam as proteínas responsáveis pela evasão imune estão em cinza, as proteínas do complexo polimerase em azul, as proteínas de superfície em verde e a proteína do capsídeo em laranja. Os códons de iniciação (GS) estão representados pelas caixas brancas e os códons de parada (GS) em preto. Na extremidade 3' temos a seqüência líder (le) e na região 5' a seqüência trailer (tr). Acima dos genes está representado o comprimento de aminoácidos da proteína. Acima da seqüência do gene M2 está representado a seqüência de leitura aberta (ORF) da proteína M2. (Adaptado de Collins e Melero, 2011).

1.2. Função das proteínas do hRSV

As proteínas NS1 e NS2 são as primeiras decodificadas e expressas em grande quantidade e são duas proteínas pequenas (NS1 – 139 resíduos; NS2 – 124 resíduos) sem homologia significativa entre si e com outras proteínas

celulares. Elas antagonizam a resposta antiviral celular e a indução de produção de interferon (INF) (Schlender et al., 2000; Bossert e Conzelmann, 2002; Bossert et al., 2003; Spann et al., 2004). IFN- α/β são secretados pela maioria de células eucarióticas como primeira linha de defesa contra infecção viral. O interferon ativa as células vizinhas pela via Jak-Stat, levando as células a um estado de resistência à infecção viral. Há evidências que as proteínas NS1 e NS2 são importantes para degradação de STAT2 (Lo et al., 2005; Ramaswamy et al., 2004, 2005) e inibição do Fator-3 Regulador de Interferon (IRF-3) (Spann et al., 2005).

No processo de ancoramento e adsorção do vírus, duas proteínas em particular são importantes: proteína de fusão (F) e proteína G de ancoramento; essas duas proteínas são os alvos primários dos anticorpos do hospedeiro. A glicoproteína G medeia o ancoramento do vírus na célula do hospedeiro (Figura 2), em seguida a proteína F promove a fusão do vírus com a célula do hospedeiro e também promove a agregação de células multinucleadas (produzindo o sincício pela qual são conhecidos (Peter et al., 2008) permitindo a transmissão de vírus de célula a célula.

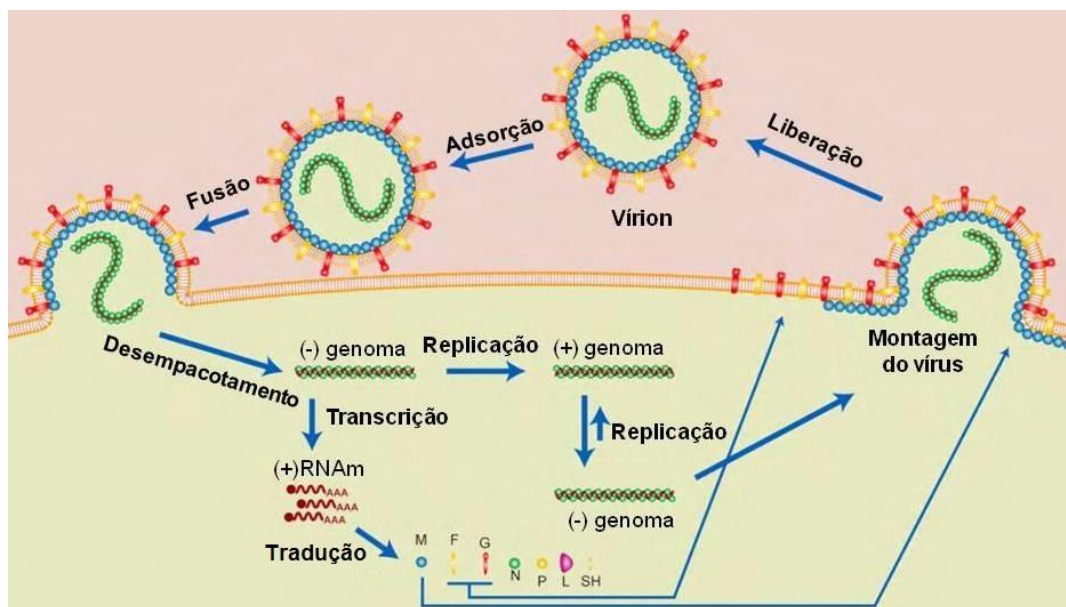


Figura 2: Ciclo de infecção do hRSV. Ao se aproximar da célula do hospedeiro, a proteína G ancora nos receptores e a proteína F promove a fusão do vírus com a célula. Há o descompactamento com a liberação do gene viral. O genoma com sentido negativo é replicado para genoma com sentido positivo e é transcrito os RNAs e tradução das proteínas. Então a proteína M junto à proteína M2-1 promove a formação do capsídeo e há a liberação de novos vírus (Adaptado de (Harrison, Sakaguchi, & Schmitt, 2010)).

A proteína SH (small hydrophobic protein) é descrita como não essencial para a replicação viral, embora na sua ausência a replicação diminua (Whitehead et al., 1999; Jin et al., 2000), mas é importante para a virulência, pois altera a permeabilidade da membrana celular (Perez et al., 1997). É uma proteína multimérica, formando pentâmeros ou hexômeros com um poro central de 1,9nm e 2,6nm, respectivamente, permitindo a passagem de pequenas moléculas (Carter et al., 2010).

Na face interna do envelope viral está associada à proteína de matriz M que se localiza no núcleo nos estágios iniciais da infecção e, é deslocada para o citoplasma nos estágios finais da infecção, onde se associa com as proteínas N, P e M2-1 (Ghildyal et al., 2002). A presença da proteína M nessas inclusões tem o papel de interromper a transcrição viral como um prelúdio para a montagem e liberação do vírus, associando o nucleocapsídeo com as glicoproteínas do envelope do RSV (Figura 2). Esta associação se dá através da porção N-terminal da proteína M com a proteína M2-1 (Ghildyal et al., 2005; Dongsheng et al., 2008).

A nucleoproteína N forma um complexo helicoidal que envolve o RNA viral (protegendo-o de ribonucleases), chamado nucleocapsídeo que se associa a outras proteínas (L, P e M2) para formar um complexo envolvido na replicação e transcrição (Blumberg et al., 1983; Meric et al., 1994; Lamb e Parks, 2006). Demonstrou-se que a região C-terminal da proteína N, essencial para a replicação, também é sítio de ligação à proteína P (Stokes et al., 2003). Diferentes trabalhos mostram que a fosforilação/desfosforilação da proteína P contribui para a interação com outras proteínas, auxiliando a transcrição: a fosforilação da Ser232 e Ser237 é essencial para ativar a transcrição viral (Barik et al., 1995); a interação da região central da proteína P com a proteína M2-1 é independente de fosforilação e, essa interação é importante para uma transcrição efetiva (Stephen et al., 2003); a presença de aminoácidos ácidos e hidrofóbicos na porção C-terminal da proteína P é crítico para a sua interação com a proteína N e não depende de fosforilação (Stephen et al., 2003; Tran et al., 2007); a fosforilação na Ser54 é crucial para o desnudamento da partícula viral no início da infecção para iniciar a transcrição e sua desfosforilação

promove associação com a proteína M, o que seria crucial para o processo de brotamento da partícula viral (Ansejo et al., 2005).

Acredita-se que em paramyxovirus a proteína P exerça duas principais funções: posicionar a proteína L no complexo N-RNA e, servir de chaperona para a proteína N recém-sintetizada (Curran et al., 1995). Há também um debate sobre a oligomerização da proteína P, o papel da fosforilação na oligomerização e qual a verdadeira função desses oligômeros. Castagne, et al. (2004), observou que a proteína P oligomeriza e essa associação é independente de fosforilação, sendo que a fosforilação é desnecessária para a interação P-P, P-N e P-M2-1. A região que compreende os resíduos 100-120 é importante para interagir com M2-1 e a região entre os resíduos 220-240 importante para a interação com a proteína N (Mason et al., 2003).

A proteína M2-1 é uma fosfoproteína básica que funciona na transcrição como um fator de processamento e antiparada (Collins et al., 1996; Hardy et al., 1998) e também auxilia na interação da proteína M com o nucleocapsídeo para a liberação da partícula viral (Li et al., 2008).

A proteína M2-2 é responsável pela mudança de fase da atividade viral e quando a concentração de mRNA atinge determinado nível, ela atua mudando a atividade do nucleocapsídeo para iniciar a replicação viral (Birmingham et al., 1999). A expressão elevada de M2-2 inibe completamente a replicação de RSV (Cheng et al., 2005).

1.2.1. Estrutura e função da proteína M2-1.

Quatro das onze proteínas (N, P, L e M2-1) formam o complexo RNA-polimerase dependente de RNA (RpRn) que está envolvido na replicação e transcrição do genoma. A transcrição é iniciada em uma região promotora de transcrição de 9 a 10 nucleotídeos conservados (gene start – GS) (Figura 1) e termina em uma região sinalizadora para poliadenilação e liberação, de 12 a 13 nucleotídeos conservados (gene end – GE) (Kuo et al., 1996). Caso a polimerase dissocie-se do RNA molde prematuramente, a transcrição dos genes a jusante são interrompidos o que leva a uma redução na taxa de transcrição dos genes finais do genoma (Fearn e Collins, 1999).

A proteína M2-1 é um fator que evita a parada prematura da transcrição, sendo assim, importante para a transcrição de todos os mRNA (Collins et al., 1996). O mecanismo realizado para evitar a parada da transcrição não está elucidado ainda, mas algumas hipóteses têm sido postuladas.

A primeira hipótese é que a proteína M2-1 pode se ligar ao mRNA nascente facilitando a elongação do transcrito, talvez prevenindo que o mRNA de hibridizar com o molde ou formar estruturas secundárias que podem desestabilizar o complexo de transcrição. Essa hipótese é sustentada por achados de que o mRNA do RSV co-precipita com M2-1 de células infectadas com RSV (Cartee e Wertz, 2001).

A segunda hipótese seria que o efeito do aumento da processividade da M2-1 poderia ser devido ao aumento da afinidade da polimerase pelo RNA genômico molde de uma maneira não específica (Cartee e Wertz, 2001).

E a terceira hipótese é que a proteína M2-1 poderia reconhecer sequências GE(3'-UCA(A/U)U(-/U/G)(A/U)(A/U)(U/A)UUUU) tanto no mRNA como no molde e evitar a liberação do complexo polimerase do molde, favorecendo a reiniciação da sequência GS (3'-CCCCGUUUUAU/C) a jusante (Kuo Jearns e Collins, 1997).

Estruturalmente, a proteína M2-1 é composta em sua maioria por α -hélice (Dubosclard et al., 2011; Tran et al., 2009) e possui quatro regiões distintas (Figura 3): os 32 resíduos iniciais da região N-terminal que se ligam ao zinco ("zinc finger") corroborando na transcrição (é indispensável para a função da proteína, mas sua função exata é desconhecida) (Hardy e Wertz, 2000; Tang et al., 2001; Zhou, Cheng e Jin, 2003); um domínio responsável pela oligomerização (formando tetrâmeros) situado entre os resíduos His33 e Gly62; um domínio globular composto por α -hélices (resíduos 63 a 170) capaz de se ligar à proteína P e RNA; e uma região C-terminal com baixo grau de estruturação (Tran et al., 2009).

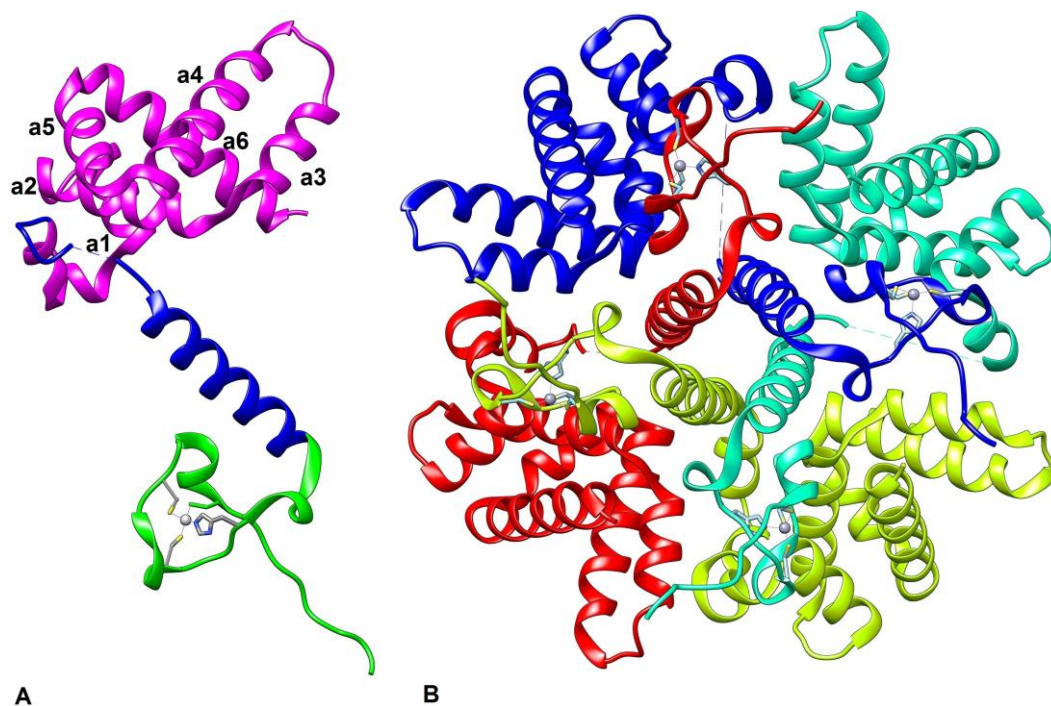


Figura 3: Estrutura cristalina da proteína M2-1 (PDB - 4C3B). Uma unidade monomérica da proteína (A) é composta por um domínio de ligação ao zinco (verde) um domínio de oligomerização (azul), um domínio de ligação à proteína P e ao RNA (magenta) e um domínio não estruturado (não mostrado). A forma ativa da proteína é na forma tetramérica (B) – cadeia A (vermelho), cadeia B (azul escuro), cadeia C (azul turquesa), cadeia D (amarela).

O domínio globular é formado por seis hélices: $\alpha 1$ (Gly75-Gly85), $\alpha 2$ (Lys92-Glu105), $\alpha 3$ (Ser108-Asp117), $\alpha 4$ (Lys124-Lys140), $\alpha 5$ (Lys143-Arg151) e $\alpha 6$ (Asp155-Ile171). Esse domínio apresenta duas faces carregadas e opostas. A face positiva contém um grande conjunto básico delimitado pelas hélices $\alpha 2$, $\alpha 5$ e $\alpha 6$ que contribuem para a interação com RNA (Blondot et al., 2012).

Observou-se que o domínio M2-1₁₅₈₋₁₇₇ interage com a proteína P na proporção 1:1 com Kd de $\approx 3\mu\text{M}$. A proteína P se liga numa região próxima à região de ancoramento de RNA. As hélices $\alpha 5$ e $\alpha 6$ podem interagir tanto com a proteína P como com RNA. Já $\alpha 4$ e $\alpha 2$ parecem estar envolvidas na ligação específica da proteína P e RNA, respectivamente (Blondot et al., 2012).

Análise de mutação sítio dirigida mostra que a substituição de Lys92 (hélice $\alpha 2$ – ligação com RNA) e Arg126 e Thr130 ($\alpha 4$ – ligação com P) por aspartato, diminuem a transcrição. Similarmente, mutações nos resíduos Lys150, Arg151, Thr160 e Asn163 nas hélices $\alpha 5$ e $\alpha 6$ que pertencem à ligação de RNA/P, tiveram o mesmo efeito (Blondot et al., 2012).

O estudo do domínio globular que compreende os resíduos Ser58-Lys177 apresenta alta afinidade por fitas RNA sentido negativo rica em uracila (Blondot

et al., 2012) e por RNA sentido positivo, ricas em purina, especialmente adenina, cuja constante de dissociação dessa interação é de 19,1nM (Tanner et al., 2014).

A análise de raio-X da proteína M2-1, mostrou que ela tetrameriza (Tanner et al., 2014), e que essa forma é a mais estável (Krissinel e Henrick, 2007). A tetramerização é principalmente mediada pela hélice $\alpha 1$, mas também é formada pelo domínio ‘dedo de zinco’ e a face N-terminal do domínio central. Essa superfície de contato extensa do tetrâmero ($14.369\text{\AA}^2$) provê uma mudança da energia livre de solvatação de -44kcal/mol , sugerindo um tetrâmero altamente estável (Tanner et al., 2014).

A face N-terminal do tetrâmero é altamente positiva, enquanto a face C-terminal também apresenta alta positividade, mas com quatro pontos de carga negativa. Entre a hélice de tetramerização e o núcleo, há uma região (resíduos Lys52-Asp67) na qual se encontra os resíduos Ser58 e Ser61 que são sítios de fosforilação (Tanner et al., 2014).

O grau de fosforilação da proteína M2-1 afeta sua atividade como pode ser observado para outras proteínas. A fosforilação do resíduo Thr108 evita a associação da proteína M2-1 ao complexo do nucleocapsídeo e este passa a ter atividade replicativa do RNA viral. Por outro lado, a desfosforilação do resíduo Thr108 leva a associação da proteína M2-1 ao nucleocapsídeo conferindo atividade transcricional (Ansejo et al., 2006). Ainda, foi observado que a fosforilação dos resíduos Ser58 e Ser61 e, a desfosforilação dos resíduos Thr56 e Ser58, reduzem a afinidade da proteína ao RNA, o que afeta a sua função (Cuesta et al, 2000; Tanner et al., 2014).

A análise de mutação sítio dirigida, mostra que os resíduos Ser58, Ser61, Lys92, Lys92, Lys150, Arg151, Arg151, Lys150, Arg151 e Lys159, exercem papel fundamental na interação do RNA à proteína e, que os resíduos Arg3, Arg4, Ser58, Ser61, Lys150, Arg151 e Lys159 são importantes para a atividade antiterminação da proteína (Tanner et al., 2014).

1.3. Terapêutica da infecção por hRSV

O tratamento primário para a infecção por hRSV é sintomático, exceto em pacientes imunocomprometidos. A administração de oxigênio para manter uma

saturação $\geq 93\%$ e a reposição de líquidos é suficiente na maioria dos casos de infecção em crianças. Mesmo assim, a mortalidade pelo vírus é alta e, estima-se que mesmo em países com baixa taxa de mortalidade, o custo anual devido à hospitalização de crianças com infecção por hRSV é em torno de US\$500 milhões (Pelletier et al., 2006).

A administração de corticosteroides, broncodilatadores e outros medicamentos são ineficientes em reduzir a taxa de hospitalização de crianças com infecção por hRSV e diminuir o período de internação desses pacientes (Flores et al., 1997; Dobson et al., 1998; American Academic 2006; Corneli et al., 2007). Uma das terapias disponíveis é o anticorpo monoclonal humanizado Synagis® (palivizumab) que se liga à proteína F facilitando a opsonização, neutralização e *clearance* do vírus. Estudos clínicos mostraram que essa terapia é eficiente em reduzir a admissão hospitalar de crianças infectadas por hRSV. O seu principal problema é o custo elevado, o que tem desencorajado o seu uso (Frenton et al., 2004; Wu et al., 2005; American Academy, 2009).

Outro fármaco licenciado, Virazole® (ribavirina) tem ampla ação antiviral, pois reduz a formação de mRNA e a polimerase viral (Sidwell et al., 2006). Mas testes mostram que não há melhora dos pacientes em relação aos pacientes do grupo controle. É uma antiviral de custo elevado e já foi demonstrado o surgimento de mutantes resistente à droga (Marquardt et al., 1995; Crotty et al., 2002a, 2002b). Inibidores da proteína F estão em desenvolvimento, mas apresentam efeitos adversos como inibição da síntese de interleucinas e baixa segurança (Andries et al., 2003; Wyde et al., 2003; Luttick et al., 2007).

As proteínas G e F são os principais alvos de estudo da inibição da infecção viral. Alguns fármacos pesquisados interagem com domínios dessas proteínas inibindo ligação do vírus à célula e a sua fusão. Um grupo de compostos tem ação semelhante, ligando à Tyr198 da proteína F, e inibindo a fusão vírus-célula. Esses compostos são: BMS-433771 (Cianci, Meanwell e Krystal, 2005), VP-14637 (Douglas et al., 2005), TMC353121 (Bonfanti et al., 2008).

O inibidor da proteína G MBX-300 apresenta potente inibição do RSV e é menos tóxico que a ribavirina, e está em fase clínica de estudo (Kimura et al., 2000; Douglas et al., 2004). RSV-604 é um inibidor da proteína N que se

mostrou ter boa absorção e atividade em faixas submicromolares (Henderson et al., 2007; Chapman et al., 2007) e, na fase 2 de estudos clínicos, mostrou diminuir a infectividade e sintomas clínicos da infecção por hRSV (Marty et al., 2007). Este composto se liga à proteína N inibindo a síntese do RNA, mas não altera a localização da proteína N na célula (Challa, Scott e Yuzhakov, 2015).

1.3.1. Quercetina.

Muitos compostos antivirais são derivados de compostos naturais (Butler et al., 2005; Harvey et al., 2008). Alguns estudos mostram que produtos naturais têm propriedades farmacológicas melhores que compostos gerados por química combinatorial (Feher et al., 2003; Koch et al., 2005; Wetzell et al., 2007). Os flavonoides, são metabólitos secundários, amplamente distribuído entre os produtos naturais (Simões et al., 2004) que apresentam atividade antitumoral, antioxidante, antiviral e anti-inflamatória (Simões et al., 2004; Cazarolli et al., 2008).

Os flavonoides são substâncias de baixo peso molecular e, estruturalmente, consistem de um esqueleto de difenil-propano ($C_6C_3C_6$) com dois anéis benzênicos (A e B) ligados a um anel pirano (C) (BEHLING et al., 2004) conforme apresentado na Figura 4-A. As várias classes de flavonoides se devem às modificações que o composto pode sofrer como hidroxilação, metilação, acilação, glicosilação, entre outras (Almeida, et al., 2004).

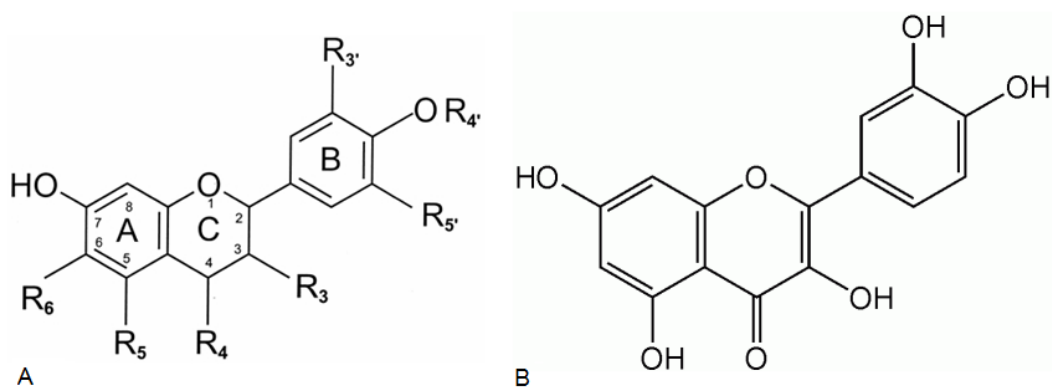


Figura 4: Estrutura básica dos flavonoides (A) e da Quercetina (B) Na estrutura básica, os anéis A e B são anéis benzênicos e C anel pirano e, as várias classes de flavonoides originam da modificação nos grupos “Rs”. (Fonte: Patel et al., 2008)

O flavonol Quercetina (Figura 4-B) é um flavonoide antioxidante que é encontrado em uma grande variedade de plantas. É um excelente acceptor de

radicais e, quelante de metais. É um dos mais potentes bioflavonoides estudado e possui uma variedade de atividades farmacológicas incluindo anti-inflamatória, antineoplásica, cardio protetora e atividade anticancerígena (Wiegand, et al., 2009; Binsack et al., 2001). Tem uma biodisponibilidade alta e, é estimado que a ingesta diária de Quercetina numa dieta é de 5 a 40mg/dia (Hertog et al., 1995), mas esses valores podem aumentar até 200 a 500mg/dia em indivíduos que consomem alta quantidade de frutas e vegetais ricos em flavonoides (maçã, cebola, tomates, etc.) (Harwood et al., 2007).

Apesar da descrição da Quercetina ter atividade anticâncer (Murakami, Ashida, Terao, 2008; Xavier et al., 2011; Dajas, 2012; Murphy et al., 2011), acredita-se que esse efeito é devido ao efeito protetor antioxidante. Por outro lado, Quercetina tem uma boa atividade antiviral. Observou-se que ela inibi a secreção de HBsAg e HBeAg em células infectadas com HBV (Wu et al., 2007); inibe a replicação de HCV (Pisonero-Vaquero et al., 2014); inibi a HIV-1 transcriptase reversa (Yu et al., 2007); e inibe a transcriptase reversa do vírus da mieloblastose aviária, vírus-2 associado a Rous e vírus da leucemia murina de Maloney (Spedding, Ratty e Middleton, 2009).

Ahmed-Belkacem e colaboradores (2014) mostraram que flavonas, especialmente a Quercetina, são capazes de inibir aproximadamente 80% da RNA-polimerase-RNA-dependente (RpRd) de HCV. Esta atividade se deve à ligação dos flavonoides ao túnel de entrada do RNA molde. Em 1985, Kaul et al., observaram que a Quercetina era capaz de inibir a infecção e replicação de RSV. Então, provavelmente, Quercetina se liga à alguma proteína do complexo RpRd do RSV evitando a replicação ou transcrição.

2. Objetivos

A proteína M2-1 é uma proteína importante na infecção pelo hRSV, pois intermedeia a interação das proteínas envolvidas na replicação e transcrição. Assim, inibidores da proteína M2-1 estariam inibindo tanto a replicação como a transcrição. A sequência primária dessa proteína mostra que não há homologia com proteínas humanas. Com isso pretende-se:

- Clonar e expressar a proteína M2-1 em bactéria *Escherichia coli*;
- Purificar a proteína M2-1;
- Analisar a estrutura secundária por Dicroísmo Circular e observar sua estabilidade em diferentes condições de temperatura;
- Analisar a interação proteína/ligante por métodos espectroscópicos;
- Analisar a estrutura monomérica, dimérica, trimérica e tetramérica da proteína M2-1 por dinâmica molecular;
- Realizar o *docking* molecular da proteína M2-1 com o flavonoide Quercetina para inferir o mecanismo de interação.

3. Materiais e Métodos

3.1. Clonagem do gene M2-1 do hRSV com sítios de restrição para as enzimas EcoRI e BamHI.

Para o desenho dos oligonucleotídeos iniciadores (*primers*), primeiramente foi realizada uma busca de sequências para cada gene do hRSV do tipo A no banco de dados GeneBank (Tabela 1). As sequências gênicas foram alinhadas utilizando-se o programa BioEdit (Hall, 1999). De posse dos alinhamentos, identificou-se as regiões gênicas que representam os domínios amino e carboxi-terminal das proteínas por eles codificadas de forma a desenhar *primers* que possibilitassem à amplificação dos genes completos, incluindo os códons de iniciação e terminação.

Tabela 1: Sequências utilizadas para alinhamento e desenho de *primers* do gene M2.1

Número de Acesso	Descrição
AY911262	Human respiratory syncytial virus strain ATCC VR-26, complete genome
FJ614813	Respiratory syncytial virus strain 19, complete genome
U39661	RSU39661 Respiratory syncytial virus, complete genome
AY911262	Human respiratory syncytial virus strain ATCC VR-26, complete genome
FJ614813	Respiratory syncytial virus strain 19, complete genome

As sequências foram rastreadas para identificar regiões reconhecidas por enzimas de restrição. Foram realizados rastreamentos para verificar as enzimas de restrição que clivassem os genes estudados, sendo selecionadas as que não clivavam o gene para evitar que a região gênica pudesse sofrer digestão nos passos seguintes. Em cada *primer* foi inserido um sítio de restrição e de ancoragem para a enzima de restrição escolhida. As enzimas utilizadas foram BamHI e EcoRI. Foram ainda inseridas as regiões de ancoragem correspondentes às suas respectivas enzimas (Tabela 2).

Tabela 2: Enzimas de restrição e seus respectivos sítios de reconhecimento

Enzima de Restrição	Sítio de Reconhecimento
BamHI	GGATCC
EcoRI	GAATTC

Os *primers* da proteína M2-1 foram desenhados para a amplificação de seus genes completos e consequentemente para a expressão de suas proteínas completas. Após o desenho de cada *primer* este foi submetido a uma análise no programa Oligo Analyser (Integrated DNA Technologies) para verificação de parâmetros como temperatura de anelamento e ocorrência de formação de dímeros e grampos (hairpins). *Primers* que não apresentaram características satisfatórias foram refeitos e resubmetidos à análise. Os *primers* desenhados e sintetizados encontram-se na tabela 03.

Tabela 3: Primers para os genes RSV tipo A

Gene	Primers	Sequência	Tamanho do primer	Enzima de restrição	Tamanho do fragmento
Matrix (M2-1) pGEX-4T-1	RSVA M21+	GTACATGGATCCATGTCACGAAGG	24pb	EcoRI	603pb
	RSVA M21-	GTCCGTGAATTCTCAGGTAGTATC	24pb	BamHI	

3.1.1. Reação de PCR e eletroforese

Após a obtenção dos pares de oligonucleotídeos iniciadores a reação de PCR para amplificação do gene M2-1 foi padronizada. Para amplificação do gene M2-1, utilizou-se reação composta por 1,25U de Taq DNA Polymerase (recombinant) (Fermentas®), 1X Taq Buffer com KCl, 1,5mM de MgCl₂, 0,5µM de cada *primer* (Invitrogen®), 0,2mM de dNTP mix (Fermentas®), 100ng de cDNA, completando o volume de 25µL de reação com água. A reação foi então transferida para o termociclador PCR Genemate Series e aplicada as seguintes termociclagens:

- Etapa inicial de desnaturação
 - 94°C por 5 minutos
- 35 ciclos
 - Desnaturação: 94°C por 1minuto
 - Anelamento: 51°C por 1 minuto
 - Extensão: 72°C por 1 minuto e 30 segundos
- Extensão final
 - 72°C por 10minutos

Os produtos amplificados foram aplicados em gel de agarose 0,8% em tampão Tris-Borato-EDTA (Tris 890mM, EDTA 20mM e 890mM ácido bórico) com adição de 5µg/mL de Brometo Etídeo. O gel foi submetido à uma voltagem de 3V/cm por 1 hora. Após esse período, o gel foi visualizado em transiluminador ultra-violeta Vilber Lourmat.

3.2. Clonagem com o TOPO-XL PCR Cloning kit (Invitrogen®)

Após a amplificação e purificação do gene amplificado, procedeu-se a clonagem dos fragmentos com o TOPO-XL PCR Cloning Kit (Invitrogen®) (Anexo I). O vetor pCR-XL-TOPO é fornecido linearizado e possui sítios politimina nos quais os insertos se ligam, uma vez que as enzimas DNA polimerases comerciais possuem atividade transferase independente de molde que adiciona uma cauda poli-adenina ao final da extensão. Assim, o fragmento recém-sintetizado liga-se ao vetor sem a necessidade de digestões prévias de ambos. O plasmídeo pCR-XL-TOPO (Figura 05) apresenta 3,5kb de e um gene de resistência à canamicina como marca de seleção.

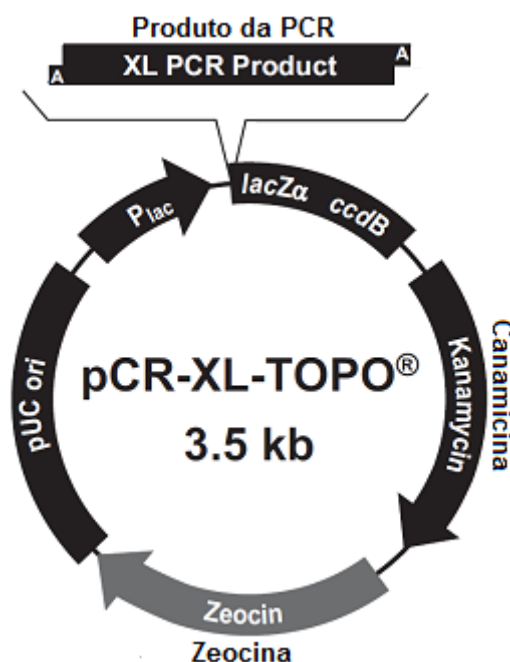


Figura 5: Mapa do vetor pCR-XL-TOPO (Invitrogen). O vetor apresenta genes de resistência aos antibióticos canamicina e zeomicina.

3.2.1. Purificação dos produtos de PCR.

Primeiramente preparou-se um gel de agarose de 1% em Tampão Tris-Acetato-EDTA (Tris 242g, ácido acético 57,1 ml, EDTA 100ml a 0,5M –H=8, água destilada q.s.p. 1 L) adicionado de 1,2µg/mL de Cristal Violeta. Às amostras foi acrescido tampão de corrida Cristal Violeta 1X e o gel foi submetido a uma voltagem de 80V, 90mA durante 45 minutos. Esse procedimento visa reduzir os danos ao DNA causados pelo brometo de etídeo e a exposição à luz ultra-violeta. Após esse período, as bandas foram visualizadas, com corante azul-violeta sob luz comum e excisadas, sendo purificadas em colunas de sílica fornecidas pelo fabricante. As amostras foram eluídas 40µL em água e armazenados a -20°C.

3.2.2. Reação de ligação TOPO®

Foi estabelecida a seguinte reação de clonagem TOPO®:

Produto de PCR purificado do gel	4µL
Vetor pCR-XL-TOPO	1µL
Volume final	5µL

A reação foi gentilmente homogeneizada e incubada a temperatura ambiente (25°C) por 5 minutos. Adicionou-se, então, 1µL da solução de parada (6X TOPO Stop Cloning) e agitando por dois segundos a temperatura ambiente. A reação foi centrifugada por 40 segundos a 2000 x g e colocada em gelo para posterior transformação.

3.2.3. Transformação nas células competentes (transformação química)

Adicionou-se 2µL da reação TOPO Cloning (produto da ligação) ao tubo com células *Escherichia coli* linhagem TOP10 quimicamente competentes. Incubou-se em gelo por 30 minutos. Em seguida, procedeu-se o choque térmico por 30 segundos a 42°C. Imediatamente os tubos foram transferidos para o gelo e incubados por 2 minutos. Adicionou-se 250µL de meio S.O.C. (fornecido pelo fabricante) à temperatura ambiente. Incubou-se horizontalmente no *shaker* por uma hora a 37°C. Em seguida, 50µL da cultura foi plaqueada em meio LB sólido contendo canamicina (50µg/mL). As placas foram incubadas *overnight* a

37°C. Uma colônia foi selecionada e crescida em meio líquido contendo antibiótico para ser utilizada para purificação do plasmídeo.

3.2.4. Purificação dos plasmídeos para confirmação da clonagem.

Os plasmídios pCR-XL-TOPO – M2-1 foram purificados utilizando-se o kit QIAprep® Spin Miniprep Kit (Qiagen®), seguindo recomendações do fabricante (Anexo II) sendo eluídos em 40µL de água.

3.2.5. Confirmação da clonagem por co-digestão enzimática.

Os plasmídeos purificados, foram co-digeridos com as enzimas BamHI e EcoRI (Tabela 4). A reação foi de 6 horas à 37°C e, 20 minutos à 65°C para inativação da enzima.

Tabela 4: Reação de co-digestão do plasmídeo pCR-XL-TOPO + M2-1 com as enzimas BamHI e EcoRI.

Reagente	Volume da reação
Tampão Tango 10X	2µL
Enzima EcoRI	2µL
Enzima BamHI	2µL
Água	4µL
Plasmídeo	10µL
Total	20µL

O produto da reação de co-digestão enzimática foi aplicado em gel de agarose 0,8% com 5µg/mL de Brometo Etídeo e feito a eletroforese à 80V, 90mA por 45 minutos. Em seguida o gel foi visualizado com o transluminador ultra-violeta Vilber Lourmat. O produto da digestão foi então purificado para ser utilizado na etapa seguinte

3.3. Subclonagem no vetor pGEX4T-1.

O plasmídeo pGEX-4T-1 possui aproximadamente 4.900pb nos quais encontram-se um gene de resistência à ampicilina (marca de seleção natural), uma região promotora T7 para a expressão de proteínas heterólogas e um *Tag* de Glutathione-S-Transferase na região N-terminal ao inserto, que permite que a proteína expressa seja purificada por meio de resinas específicas.

O sistema de fusão gênica Glutathione-S-Transferase (GST) (GE HealthCare) é um sistema composto com um promotor *tac* para alta indução de

proteínas recombinantes fusionada à GST, possui um gene *lacI^q* para uso em qualquer linhagem de *E. coli* e apresenta sítio de clivagem para trombina (Figura 06).

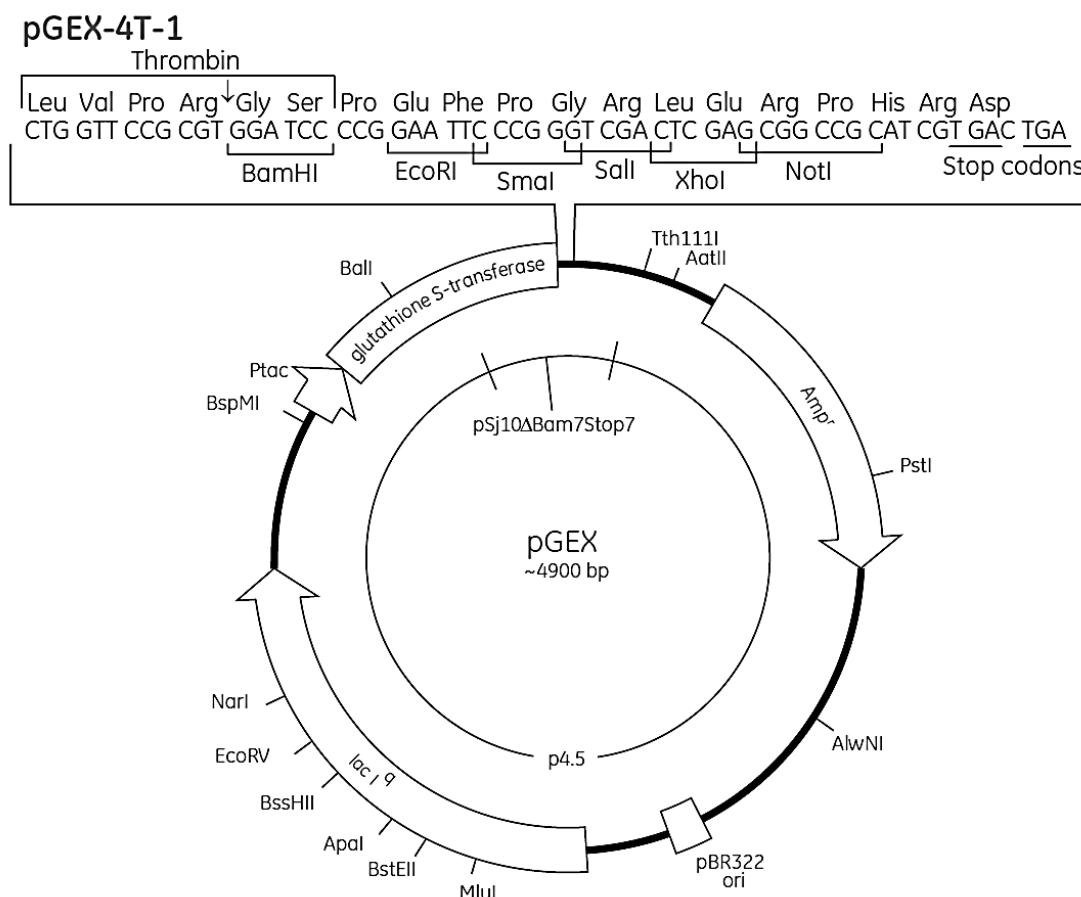


Figura 06: Mapa gênico do vetor pGEX-4T-1 (GE Healthcare Life Sciences). A proteína expressa apresenta uma cauda GST, e o vetor apresenta gene de resistência ao antibiótico ampicilina.

Para a construção dos plasmídeos recombinantes no pGEX-4T-1 para expressão foram digeridos pelas enzimas BamHI e EcoRI (FERMENTAS) tanto os plasmídeos recombinantes pCR-XL-TOPO – M21, quanto o vetor pGEX-4T-1 (GE HealthCare). Estes ensaios de restrição empregaram 2,0 µL de Buffer R (10 mM Tris-HCl pH 8.5, 10 mM MgCl₂, 100mM KCl and 0.1 mg/ml BSA.) (Fermentas), 1µL de enzima EcoRI e BamHI, aproximadamente 1,0 µg de DNA do inserto ou do vetor e água Milli-Q para um volume final de reação de 20,0 µL. A incubação aconteceu durante uma hora e meia a 37° C. O produto da reação foi visualizado por eletroforese em gel de agarose 1%. Para a purificação dos fragmentos do vetor foi usado o DNA Gel Extraction Kit (NORGEN, Canadá), de acordo com as orientações do fabricante (Anexo III).

Após a reação de digestão, foi determinada a concentração do fragmento M2-1 gerado e do vetor pGEX-4T-1 digerido. A reação de ligação de ambos foi feita com a enzima T4 DNA Ligase (1 U/μL) (INVITROGEN, USA), numa reação de razão molar 3:1 (Tabela 5) e 6:1 (Tabela 6). Os volumes dos reagentes foram calculados conforme a Equação 01. A reação de ligação foi realizada à 16°C por 16 horas.

$$\frac{\text{Quantidade de plasmídeo (ng)} \times \text{Tamanho do inserto (Kb)}}{\text{Tamanho do plasmídeo (Kb)}} \times \frac{\text{Razão molar do inseto}}{\text{Plasmídeo}} = \text{Quantidade de inserto (ng)} \quad (1)$$

Tabela 5: Protocolo de reação de ligação do plasmídeo pGEX-4T-1 com o inseto M2-1 na razão 3:1.

Reagente	Volume da reação
Plasmídeo pGEX-4T-1	11,1uL
Inserto M2-1	2,6uL
Água	13,3uL
Tampão T4 DNA ligase 10X(INVITROGEN, USA)	2uL
T4 DNA ligase (INVITROGEN, USA)	1uL
Total	20uL

Tabela 6: Protocolo de reação de ligação do plasmídeo pGEX-4T-1 com o inseto M2-1 na razão 3:1

Reagente	Volume da reação
Plasmídeo pGEX-4T-1	11,1uL
Inserto M2-1	5,3uL
Água	10,6uL
Tampão T4 DNA ligase 10X (INVITROGEN, USA)	2uL
T4 DNA ligase (INVITROGEN, USA)	1uL
Total	20uL

A *E. coli* linhagem DH5α foi transformada por choque térmico com os produtos da ligação e plaqueadas em meio LB sólido acrescido de ampicilina (100μg/ml) e incubado na estufa à 37°C overnight.

Para a confirmação da correta ligação, os plasmídeos foram extraídos usando o quite QIAprep® Spin Miniprep (Quiagen) e co-digeridos com EcoRI e BamHI (Tabela 04).

3.4. Transformação em *E. coli* linhagem BL21 Gold (DE3)

A linhagem BL21 Gold (DE3) (Agilent®) é uma variação melhorada da linhagem BL21. Essa linhagem é ideal para sistemas de expressão que usam o promotor T7, não possuem a protease Lon e OmpT, evitando a degradação de proteínas recombinantes e, não possuem o gene *endA* que codifica a enzima endonuclease I que degrada plasmídeos. Essa linhagem foi transformada com pGEX-4T-1 – M2-1 por choque térmico.

3.5. Sequenciamento da proteína M2-1.

A proteína M2-1 purificada foi enviada para a Unesp - Centro de Estudos de Insetos Sociais (CEIS) para determinação da sequência primária pelo método de Shotgun (Gillet et al., 2012). O experimento foi realizado pela doutoranda Nathalia Baptista Dias, orientada do Prof. Dr. Mário Sérgio Palma.

3.6. Indução da expressão da proteína M2-1.

Os clones foram cultivados em meio LB líquido acrescido de ampicilina (100µg/ml) e tetraciclina (50µg/ml) por 6 horas à 37°C e, então, acrescido IPTG (Isopropil-β-D-1-tiogalactopiranosídeo) para uma concentração final de 0,4mM e sulfato de zinco para uma concentração final de 50µM, e induzido à 28°C por 16 horas.

3.7. Purificação da proteína M2-1 fusionada com GST.

A cultura foi centrifugada em 54.000xg a 4°C por 40 minutos. O sobrenadante foi descartado e o *pellet* ressuspendido em tampão de lise (Tris-HCl 50mM, NaCl 1M, Tween 20 2% v/v, 2-Mercaptoetanol 1mM, CaCl₂ 1mM, MgSO₄ 1mM, lisozima 1mg/ml, pH=7,4) acrescido de inibidor de protease. A amostra ficou encubada por 30 minutos em gelo e foi sonicada com 15 pulsos (36W) de 10 segundos, com 50 segundos em repouso em gelo.

A amostra foi centrifugada à 54.000xg por 40 minutos, e o sobrenadante foi aplicado na resina GSTrap 4B previamente equilibrada com o tampão de lise. Após 3 horas de incubação, a coluna foi lavada com 2 volumes de coluna de

tampão de lise e depois lavada com 5 volumes de coluna de PBS (NaCl 137mM, KCl 2,7mM, Na₂HPO₄ 10mM, KH₂PO₄ 1,8mM pH = 7,4).

Foi aplicado 100 unidades de trombina, e a resina ficou sob agitação constante por 16 horas à 25°C *overnight*. Então, a amostra foi eluída com 3 volumes de coluna com tampão PBS, concentrada até um volume final de 500µL e aplicado na resina Superdex 200 10/300 GL equilibrada com tampão PBS e eluída com o mesmo tampão.

3.8. Medida da concentração da proteína.

Para a medida da concentração da proteína, foi utilizado o espectrofotômetro Cary-3E spectrophotometer (Varian, Palo Alto, CA). A concentração foi determinada pela Lei de Lambert Beer, na qual foi feito a medida de absorbância da amostra em 280nm e usado o coeficiente de extinção molar de 13.200 mol L⁻¹ cm⁻¹ (determinado pelo servidor ProtParam).

3.9. Preparo das amostras.

As amostras das proteínas foram purificadas e mantidas em tampão PBS pH 7,4, em gelo na geladeira. As soluções estoque dos flavonoides Quercetina (MM=302,24 Da) (3,3',4',5,6-Pentahydroxyflavone) foi preparado em etanol 90%v/v.

3.10. Análise estrutura secundária por dicroísmo circular.

As medidas de CD foram realizadas em espectropolarímetro Jasco J-810, a temperatura ambiente, em cubeta de 0,5cm de caminho óptico. O espectro entre 200 e 260nm foi acumulado a uma velocidade de 20nm/min, com resposta de 2s e abertura de fenda de 1,0nm, realizando-se 10 acumulações.

As medidas de desnaturação térmica foram realizadas em espectropolarímetro Jasco J-810 acoplado a um sistema peltier, no intervalo de temperatura de 20°C a 80°C, com velocidade de aquecimento de 1°C/min e leituras a cada 5°C. Foi utilizada cubeta de 0,5cm de caminho óptico, com leitura em 222nm, com resposta de 2s e abertura de fenda de 1,0nm, realizando-se 3 acumulações.

3.11. Supressão de fluorescência.

A análise da supressão de fluorescência foi realizada em espectrofluorímetro ISS PC1 de estado estacionário (Champaign, IL, USA). Para a titulação da proteína com o flavonoide, primeiro foi determinado a concentração da proteína para usar uma razão ligante:proteína de 0,1 a 2. A titulação da proteína com os flavonoides foi feito com controle de temperatura.

3.11.1. Medidas espectrofotométricas e equipamentos

Os espectros de UV/Vis foram obtidos com espectrofotômetro Cary-3E spectrophotometer (Varian, Palo Alto, CA) em cubetas de quartzo de 1 cm de caminho óptico na temperatura ambiente, nas seguintes condições: velocidade de varredura 600 nm/min, intervalo de 1 nm, abertura de fenda 2,0 nm.

As medidas de fluorescência foram realizadas com espectrofluorímetro ISS PC1 de estado estacionário (Champaign, IL, USA), equipado com sistema de controle de temperatura Neslab RTE-221, em cubetas de quartzo de 1 cm de caminho óptico, na temperatura de 25°C. A abertura das fendas de excitação e de emissão foi ajustada em 8,0 nm.

As medidas de UV/Vis foram utilizadas para se determinar as concentrações estoque das amostras da proteína e o ligante Quercetina. As soluções estoque de Quercetina (MM=338,3Da) foi preparadas em 10% de água Milli-Q e 90% de etanol PA e as proteínas M2-1 foi preparada em solução tampão PBS pH=7.4. Na determinação da concentração da solução estoque da Quercetina foi utilizando o coeficiente de extinção molar de $\epsilon_{375}=21.880 \text{ M}^{-1}\text{cm}^{-1}$ em 375nm (Merck) (Pradu, et al, 2006), e o cálculo das concentrações das proteínas M2-1 foi utilizado o valor do coeficiente de extinção molar $\epsilon_{280}=13.200 \text{ cm}^{-1}\text{M}^{-1}$ em 280nm.

3.11.1.1. Princípios da supressão da fluorescência (quenching)

As medidas da interação da proteína M2-1 com a Quercetina foi realizada medindo-se os espectros de fluorescência do triptofano das proteínas e da

diminuição (supressão) da intensidade dos mesmos após a adição de ligante na solução.

A intensidade de fluorescência de uma macromolécula pode diminuir devido a uma variedade de interações com outros compostos moleculares: rearranjo molecular, transferência de energia, supressão devido a colisões, dentre outros (Lakowicz, 1999). Tal decréscimo da fluorescência é denominado supressão ou “*quenching*” e é descrita pela relação de Stern–Volmer, através da seguinte equação:

$$\frac{F_0}{F} = 1 + K_{SV}[L] = 1 + k_q \tau_0 [L] \quad (2)$$

onde F_0 e F são respectivamente as intensidades de fluorescência antes e após a adição de ligante. K_q , K_{SV} , τ_0 , e $[L]$ são a constante de supressão da biomolécula, a constante de Stern–Volmer que representa a constante dinâmica de supressão, o tempo de vida média de emissão da fluorescência da biomolécula na ausência de ligante (da ordem de $\tau_0 = 10^{-8}$ s) e a concentração de ligante (supressor), respectivamente.

Quando for observada relação linear entre a razão de fluorescências F_0/F e a concentração de ligante $[L]$, é um forte indício de que a interação entre a proteína e o ligante é uma ligação estática, não colisional, pois neste tipo, o gráfico de Stern-Volmer apresentaria uma concavidade positiva, assim, a concentração de ligantes está relacionada à razão F_0/F pela seguinte relação (Lakowicz, 1999):

$$\frac{F_0}{F} = (1 + K_D[L])(1 + K_S[L]) \quad (3)$$

ou seja:

$$\frac{F_0}{F} = 1 + (K_D + K_S)[L] + K_D K_S [L]^2 \quad (3')$$

onde K_D e K_S são as constantes dinâmicas e estáticas respectivamente sendo prevista uma dependência quadrática com a concentração de ligante, resultando numa curva com concavidade positiva no gráfico F_0/F versus $[L]$.

3.11.1.1.1. Cálculo dos parâmetros de interação.

No caso do mecanismo de supressão estática, onde o ligante ligado suprime totalmente a fluorescência da proteína, a suposição adotada é que os sítios de ligação na proteína são iguais e independentes. Este modelo permite escrever a equação de equilíbrio de reação como:



onde $[P]$ é a concentração de proteína, $[L]$ é a concentração de ligante e $[L_nP]$ é a concentração de complexo Proteína-ligante com constante de ligação K_b , dada por:

$$K_b = \frac{[L_nP]}{[L]^n[P]} \quad (5)$$

Assumindo um modelo de dois estados no qual a proteína na solução encontra-se complexada com o ligante ou livre, ou seja, $[P_t] = [L_nP] + [P_f]$, onde $[P_f]$ é a concentração de proteína não complexada com ligante ou proteína livre, de tal forma que a razão entre a intensidade de fluorescência da solução que contém proteína livre (sem a presença de ligante) F_0 e a intensidade de fluorescência após a adição de ligante na solução F é igual à razão entre a concentração de proteína livre e proteína ligada:

$$\frac{[P_f]}{[L_nP]} = \frac{F_0}{F} \quad (6)$$

Fazendo $[P_b] = [L_nP]$, temos:

$$[P_t] = [P_b] + [P_f] \quad (7)$$

Substituindo as equações (6) e (7) em (5), temos:

$$K_b = \frac{[P_t] + [P_f]}{[L]^n[P_f]} = \frac{[P_t] + [P_f]}{[L]^n[P_f]}, \text{ onde}$$

$$\frac{[P_t] + [P_f]}{[P_f]} = \frac{F_0 - F}{F}, \text{ portanto:}$$

$$\frac{F_0 - F}{F} = K_b + [L_f]^n \quad (8)$$

Admitindo que a concentração de ligante na solução é maior que a concentração de proteína, portanto $[L_t] \cong [L_f] = [L]$, pode-se então relacionar a fração de fluorescência com a concentração de ligantes através da seguinte relação:

$$\log\left(\frac{F_0 - F}{F}\right) = \log(K_b) + n \log([L]) \quad (9)$$

Onde, F_0 e F são as fluorescências antes e após a adição de ligante $[L]$ (Quercetina) à proteína M2-1.

Através do gráfico $\log[(F_0 - F)/F]$ versus $\log([L])$, onde se observe dependência linear, pode-se obter os valores relativos ao número n de sítios de ligação da proteína e a constante de associação K_b relativa à interação dos ligantes com as proteínas.

3.12. Análise de Bioinformática

3.12.1. Estrutura molecular

As estruturas iniciais da proteína M2-1 foram obtidas pelo servidor I-Tasser (Yang et al., 2015). A partir da sequência primária da proteína, o servidor gerou cinco estruturas, na qual a primeira apresentou melhor confiança na qualidade, de acordo com os parâmetros do próprio servidor. Em seguida foi construído os modelos dímero trímero e tetrâmero da proteína M2-1 pelo servidor SymmDock (Schneiden-Duhovny et al., 2005). Os modelos escolhidos foram aqueles que apresentaram maior contato entre os domínios de oligomerização.

3.12.2. Dinâmica molecular.

A dinâmica molecular é uma ferramenta computacional que utiliza a mecânica molecular para simular a evolução temporal de um sistema. Assim,

são calculados a posição atômica de cada átomo constituinte do sistema em distintos e consecutivos instantes de tempo. Esse conjunto de coordenadas por tempo descreve a trajetória do sistema (Hansson, Oostenbrink e van Gunsteren, 2002).

A dinâmica molecular tem por base um modelo simples de interações que leva em consideração as contribuições de processos como o alongamento das ligações químicas entre dois átomos, a variação dos ângulos e a rotação sobre ligações simples. À primeira vista pode parecer um modelo muito simples, porém seus resultados em prever estruturas podem ser comparados aos resultados obtidos por métodos *ab initio*, sendo produzidos em um tempo computacional bem menor. Entretanto não é possível obter resultados referentes à distribuição eletrônica (Leach, 1996).

3.12.2.1. Campo de Força: GROMOS96

Para calcular a posição de cada núcleo em cada instante, é necessário conhecer a força que atua sobre ele, e a força pode ser calculada através da função gradiente da energia total do sistema para sistemas clássicos. Essa é a base para o tratamento de dinâmica molecular. Os potenciais clássicos são representados por funções e seus parâmetros são constantes obtidas de dados experimentais ou mesmo por cálculos *ab initio*. Cada campo de força utiliza a energia total ou hamiltoniano de um sistema clássico é do tipo:

$$H = E_{cinética} + V_{potencial} \quad (10)$$

A energia cinética é, normalmente, independente da posição da coordenada r , só depende do momento de cada partícula, p , e da massa das N partículas do sistema. A massa de cada átomo está parametrizada no campo de força, portanto a energia cinética será a velocidade naquele instante ao quadrado vezes a massa da partícula dividido por dois, para todos os N átomos do sistema. Note que no campo de força GROMOS96 os átomos de hidrogênio ligados a átomos de carbono alifáticos são tratados apenas como um átomo, o que diminui os graus de liberdade do sistema, facilitando os cálculos (Guerra 2014).

O termo referente à energia potencial pode ser dividido em:

$$V_{potencial} = V_{ligado} + V_{nãoligado} + V_{especiais} \quad (11)$$

onde o termo V_{ligado} se refere aos termos com ligações covalentes, e $V_{nãoligado}$ são referentes às forças que agem a longa e curta distancias para átomos não ligados covalentemente, e $V_{especiais}$ se referem às restrições de posições e vínculos em comprimentos e ângulos de ligação.

3.12.2.2. Parâmetros da Dinâmica Molecular usados.

Para o estudo de dinâmica molecular, utilizou-se o programa GROMACS (Berendsen, van der Spoel e van Drunen, 1995, Lindahl, Hess e van der Spoel, 2001). A proteína M2-1 foi solvatada com solvente explícito modelo SPC, em uma caixa cúbica com espaço mínimo entre a borda e o soluto de 15 nm entre a extremidade do soluto e, 8, 16, 24 e 32 íons Cl^- foram adicionados para neutralizar o sistema com o monômero, dímero, trímero e tetrâmero, respectivamente. Foram feitos três passos de minimização de energia do sistema: usando o método *steepest descent* para minimizar primeiro a proteína e depois o solvente e, o método de gradiente conjugado. A temperatura de referência foi fixada em 310 K. O campo de força GROMOS96 53a6 foi usado para o cálculo da energia em um tempo de 100 ns. As coordenadas da trajetória foram salvas a cada 2 ps (Anexo IV). As ferramentas implementadas no GROMACS foram usadas para a análise dos resultados.

3.12.2.3. Análises

Os resultados oriundos da dinâmica molecular foram analisados utilizando-se as ferramentas do pacote computacional GROMACS que permite escolher entre vários campos de força, modelos de água, de acordo com a necessidades da dinâmica.

Para obter o desvio entre duas estruturas, ou as flutuações médias dos átomos durante o tempo de simulação utiliza-se o desvio quadrático médio (RMSD); esse método nos mostra como uma estrutura de proteína varia ao longo do tempo de simulação em relação à estrutura inicial, mostrando a influência do solvente. Ou ainda, permite a comparação de duas estruturas

iguais submetidas a diferentes processos. A equação utilizada para o cálculo do RMSD é dada por:

$$RMSD = \left[\frac{1}{N} \sum_{i=1}^N (\vec{x}_i - \vec{x}_i^0)^2 \right]^{\frac{1}{2}} \quad (12)$$

Sendo $\vec{x}_i$ e $\vec{x}_i^0$ os vetores de posição atômicos do átomo i em dois instantes distintos. Essa análise é feita diretamente pelo pacote computacional GROMACS através do comando `g_rms`.

Outra análise importante é o desvio padrão entre todas as estruturas durante um certo tempo de simulação. No caso das proteínas utilizamos essa ferramenta para calcular quanto variaram em média os resíduos de aminoácidos quando comparados à estrutura de referência, sendo que o comando utilizado é o `g_rmsf`.

Com o propósito de obter a estrutura que mais frequentemente ocorreu durante o tempo de simulação, utilizamos a facilidade denominada `g_cluster`; essa ferramenta permite indicar qual é a estrutura mais representativa da trajetória. Para obter esse dado calcula-se o desvio quadrático médio (RMSD) entre todas as estruturas da trajetória em relação à estrutura inicial, ou qualquer estrutura de escolha, então as estruturas com RMSD semelhantes vão sendo alocados em clusters. Os arquivos de saída nos mostram o tamanho dos cluster e qual o frame característico para o frame mais representativo de cada cluster.

A estrutura secundária da proteína foi obtida através de uma rotina computacional denominada `do_dssp`, que lê a trajetória da dinâmica e computa a estrutura secundária para cada frame e por resíduo. O arquivo de saída é uma figura mostrando o número de resíduos com cada tipo de estrutura secundária ao longo do tempo.

3.12.3. Docking molecular.

O estudo de *docking* molecular consiste na predição da conformação e orientação (*pose*) do ligante no sítio de ligação de um alvo e, do cálculo de sua energia de interação (*score*). O processo de *docking* inicia com o posicionamento do ligante no receptor, e através de funções de *scoring*, busca-

se uma conformação do ligante que seja complementar ao receptor de forma estrutural e energética. As características estruturais determinam os arranjos espaciais moleculares e, as características energéticas são as ligações não covalentes (Sinnokrot e Sherrill, 2003).

Todos os cálculos de *docking* foram efetuados usando o programa AutoDock 4.0 (Morris, et al. 2009). O GRID foi construído centralizado na proteína com uma caixa de 126x126x126 pontos com espaçamento de 0,500Å. Foi utilizado o algoritmo de busca Lamarckian Genetic Algorithm com o número de GA de 200 passos, tamanho da população de 300 e o número máximo de avaliações médio. Os demais parâmetros foram mantidos padrão. Os cálculos de *docking* foram realizados num computador Dell Vostro 1520 com processador Intel Core 2Duo, sistema operacional 64 bits com 4GB de memória RAM.

Os resultados obtidos foram analisados usando a própria interface gráfica do AutoDockTools (Sanner, 1999). Para a geração das representações gráficas foi utilizado o programa Chimera (Pettersen, et al. 2004).

4. Resultados e Discussão

4.1. Dados da proteína M2-1 obtida pelo servidor ProtParam.

Para início das análises, foi feita uma avaliação das propriedades físico-químicas da proteína M2-1 através do servidor ProtParam (GASTEIGER et al., 2005), sendo os parâmetros obtidos para a proteína é que esta é composta por 194 aminoácidos, massa molecular de 22,180 kDa e ponto isoelétrico teórico de 9,13. Este valor foi tomado como referencial para o preparo dos tampões para a purificação da proteína. Também, o servidor calculou o valor do coeficiente de extinção molar $\epsilon_{280} = 13.200 \text{ cm}^{-1}\text{M}^{-1}$, que foi utilizado nesse trabalho para calcular a concentração da proteína (também utilizado por (Esperante et al., 2013; Tran et al., 2009)).

Porém a proteína recombinante é expressa fusionada à proteína GST. Essa construção é composta por 420 aminoácidos, com peso molecular de 48,472 kDa, ponto isoelétrico teórico de 8,23.

4.2. Clonagem

4.2.1. Clonagem do gene M2-1 com sítio de restrição para EcoRI e BamHI.

O gene da proteína M2-1 clonado com os sítios de restrição para EcoRI e BamHI foi amplificada com tamanho aproximado de 603pb (Figura 7).

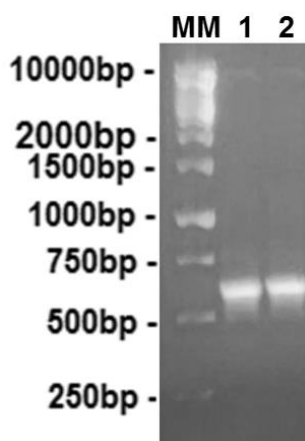


Figura 7: Reação de PCR do gene M2-1. A sequência gênica foi amplificada (~603pb) à partir de amostra de pacientes (colunas 1 e 2). MM – marcador de peso molecular.

Após a ligação do gene M2-1 ao vetor pGEX-4T-1, foi feita a co-digestão para confirmar a clonagem. Os clones obtidos do plasmídeo pGEX-4T-1-M2-1

foram digeridos e o resultado mostrou uma banda correspondente a 603pb mostrando a positividade da clonagem.

4.3. Expressão.

Dos clones transformados, cinco foram escolhidos para induzir a expressão da proteína M2-1 em 0,4mM de IPTG e 50µM de sulfato de zinco cultivados por 16 horas à 28°C. Dos cinco clones escolhidos, todos tiveram um padrão de expressão semelhante (Figura 8), podendo ser observado a banda correspondente à proteína M2-1 fusionada à GST de 48,47 kDa. Então foi selecionado um clone para induzir a expressão da proteína e realizar a purificação em larga escala.

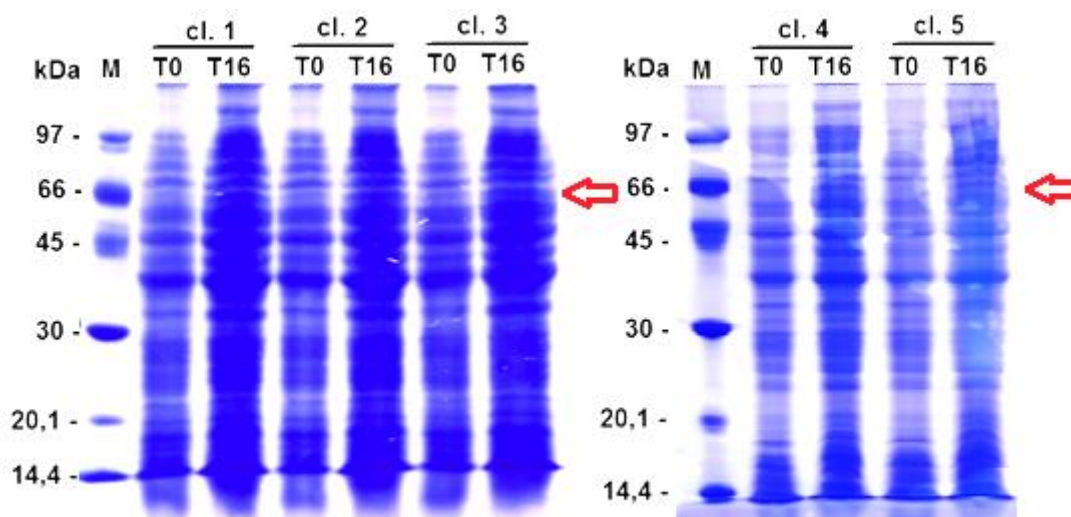


Figura 8: SDS-PAGE 15% corado com Coomassie Brilliant Blue G-250 referente ao teste de indução dos clones da *E. coli* BL21 God + pGEX4T-1 + M2-1. O tempo zero (T0) corresponde ao tempo antes da indução e o tempo 16 (T16) ao tempo após 16 horas de indução. A proteína M2-1 fusionada à proteína GST apresenta massa molar de aproximadamente 48,47kDa. (M – marcador de peso molecular). (Proteínas do marcador molecular: Fosforilase B – 97kDa; Albumina – 66kDa; Ovoalbumina – 45kDa; Anidrase carbônica – 30kDa; Inibidor de tripsina – 20,1kDa; Lactoalbumina-α – 14,4kDa).

4.4. Purificação

4.4.1. Purificação com o sistema ÄKTA da proteína M2-1 fusionada com GST.

O resultado do processo de purificação da proteína mostrando a separação da proteína M2-1 (22,18kDa) da proteína trombina (37 kDa) é mostrado na (Figura 9).

O resultado desta purificação pode ser visualizado na figura 10, na qual se observa a proteína pura resultante da segunda etapa da purificação das três frações de proteínas (Figura 10, colunas B2 e B4).

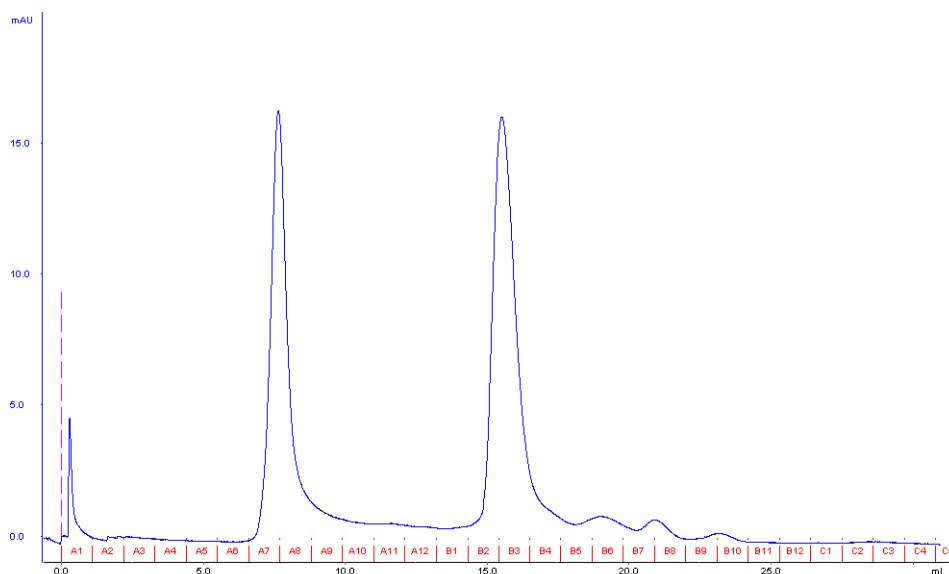


Figura 9: Cromatograma da purificação da proteína M2-1 no sistema ÄKTA (GE HEALTHCARE LIFE SCIENCES) com a resina Superdex 200 10/300 GL. Gráfico da purificação da proteína M2-1 com a resina Superdex 200 10/300 GL.

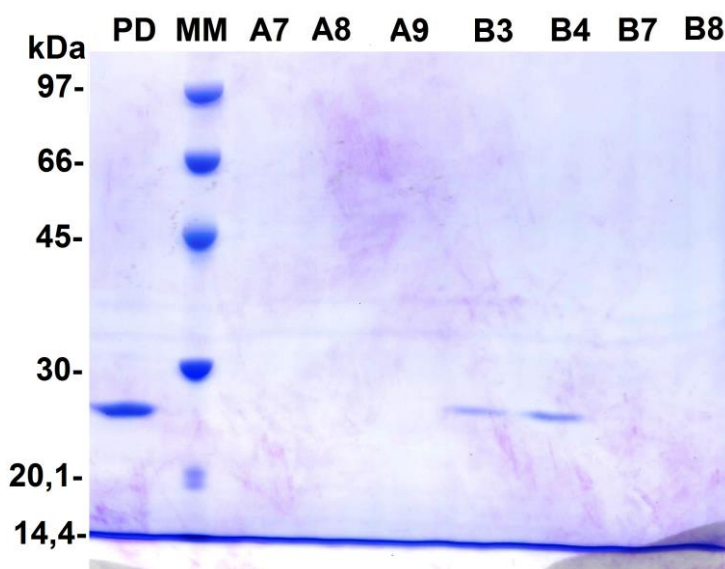


Figura 10: Resultado da purificação da proteína M2-1 em SDS-PAGE 15% corado com Coomassie Blue G-250. PD – Produto da digestão por trombina; A7 à B8 – frações coletadas pelo sistema de purificação ÄKTA; MM – Marcador de peso molecular. (Proteínas do marcador de mass molecular: Fosforilase B – 97kDa; Albumina – 66kDa; Ovoalbumina – 45kDa; Anidrase carbônica – 30kDa; Inibidor de tripsina – 20,1kDa; Lactoalbumina- α – 14,4kDa).

Devido à característica física dos ácidos nucleicos apresentarem carga negativa, eles podem exercer uma atração eletrostática sobre outras moléculas, mascarando resultados. Para garantir que não havia moléculas de ácidos nucleicos, foi feita a medida de absorbância em 260 nm (comprimento de absorção máxima de ácidos nucleicos) e a medida da absorbância em 280 nm (absorção máxima de proteínas) e calculado a relação A_{260}/A_{280} , que apresentou valor igual a 0,276, ou seja, existe aproximadamente 0,2 moléculas de ácido nucleico para cada 1 molécula de proteína. Isso garante que a concentração da proteína é maior que a de ácido nucleico, tendo mínima interferência nas análises.

4.5. Sequenciamento da estrutura primária da proteína M2-1.

A proteína M2-1 purificada foi enviada para sequenciamento da estrutura primária e, as sequências encontradas foram as que se observa na Figura 11. O alinhamento da sequência usando o banco de dados NCBI com a ferramenta BLASTp, resultou numa cobertura de 67% e identidade de 46% com a proteína de matriz 2-1 do hRSV (código de acesso – AAX23995). Este resultado confirma que a proteína obtida nos passos anteriores é a proteína M2-1, já que com uma cobertura de 19% é considerado estrutura semelhante.

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLP
YYIDGDVKLTQSMALIRYIADKHNMLGGCPKERAELSMLEGAVLDIRYGVSRAYS
KDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDEFMLYDÄLDVVLYMDPM
CLDAFPKLVCFKKRIEALPQIDKYLKSSKYIAWPLQGWQATFGGGDHPKSDLVPR
GSMSRRNPCKFEIRGHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRILKSMDKSI
DTLSEISGAELDRTEEYALGVGVLESYIGSINNITKQSACVAMSKLLTELNSDD
IKKLRDNEELNSPKIRVYNTVISYIESNRKNNKQTIHLLKRLPADVLKKTIKNTLD
IHKSTITNNPKELTVSDTNDHAKNNDTT

Figura 11: Sequenciamento da proteína M2-1. A sequência em negrito corresponde à proteína GST; a sequência em itálico corresponde à proteína M2-1; a sequência em negrito, itálico e sublinhado corresponde à sequência de reconhecimento pela trombina e as sequências sublinhadas correspondem às sequências encontradas pelo método de Shotgun.

4.6. Análise por espectroscopia de dicroísmo circular.

Uma vez obtida a proteína, fez-se necessário observar se a proteína estava estruturada com as condições de trabalho. Portanto, foi realizada a análise da

estrutura secundária pela espectroscopia de dicroísmo circular (CD). O resultado da análise de dicroísmo circular mostra que a proteína apresenta um padrão de estrutura secundária de aproximadamente 40% de alfa hélice, 10% de folhas beta e 19% de voltas e 31% de estrutura aleatórias (Figura 12). De acordo com Tran et al., 2009, a proteína M2-1 é composta por 38% alfa hélice, 12% folhas beta, 10% alças e 40% *random coils*. Nossos resultados mostram que a proteína M2-1 purificada apresenta o mesmo padrão de estrutura secundária encontrada por aquele grupo.

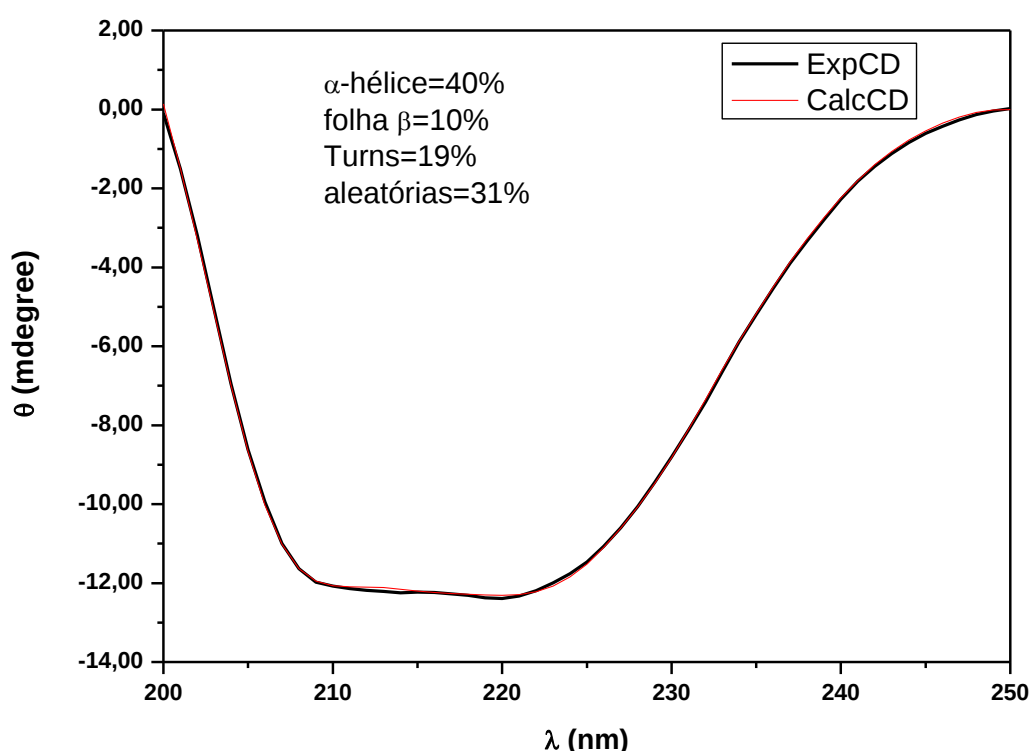


Figura 12: Espectro de dicroísmo circular da proteína M2-1. A linha vermelha representa o ajuste aos dados experimentais obtidos da deconvolução do espectro de CD.

A desnaturação térmica da proteína pela espectroscopia de dicroísmo circular mostra que ela apresenta um T_m (temperatura de *melting*) de aproximadamente 55,6°C (Figura 13), mostrando que ela é estável para trabalhar em temperatura ambiente e em temperatura semelhante ao sistema fisiológico (37°C). Pelo espectro de desnaturação (Figura 13) observamos que na faixa de temperatura biológica (aproximadamente 37°C), nas condições de experimento, houve 10% de desnaturação, não interferindo na titulação

isotérmica realizada para identificar o perfil de interação entre a proteína M2-1 e a Quercetina.

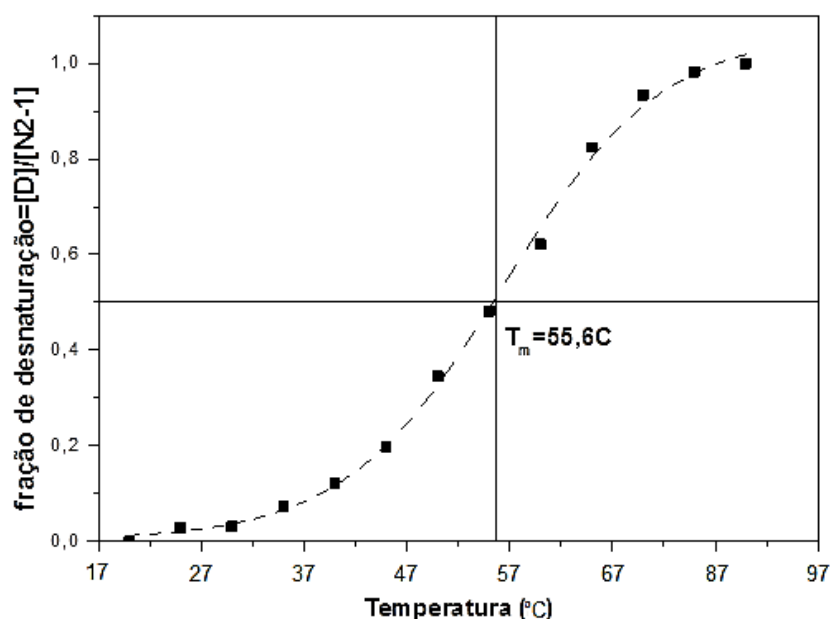


Figura 13: Desnaturalização térmica da proteína M2-1. A variação da fração de M1-2 desnaturalizada em função do aumento de temperatura, mostra que a temperatura na qual há equilíbrio entre a concentração de proteína no estado Nativo [N], e no estado desnaturalizado [D], correspondendo a uma $T_m=55,6^{\circ}\text{C}$.

4.7. Supressão de fluorescência

Para o estudo de supressão da fluorescência, fez-se necessário analisar primeiramente o decaimento da fluorescência como aumento da temperatura. Em baixas temperaturas a proteína tende a ficar numa conformação mais estável. Conforme há o aumento da temperatura, a proteína sofre alterações conformacionais que mudam o meio entorno do triptofano, levando a uma redução da emissão da fluorescência (Alcala, Gratton, & Prendergast, 1987a, 1987b). A proteína foi excitada em 280nm e feito a leitura da emissão na faixa de 290nm a 500nm em temperaturas de 25°C, 31°C e 37°C. Observamos que a cada grau aumentado há um decréscimo de aproximadamente 2,33 unidades de fluorescência (Figura 14).

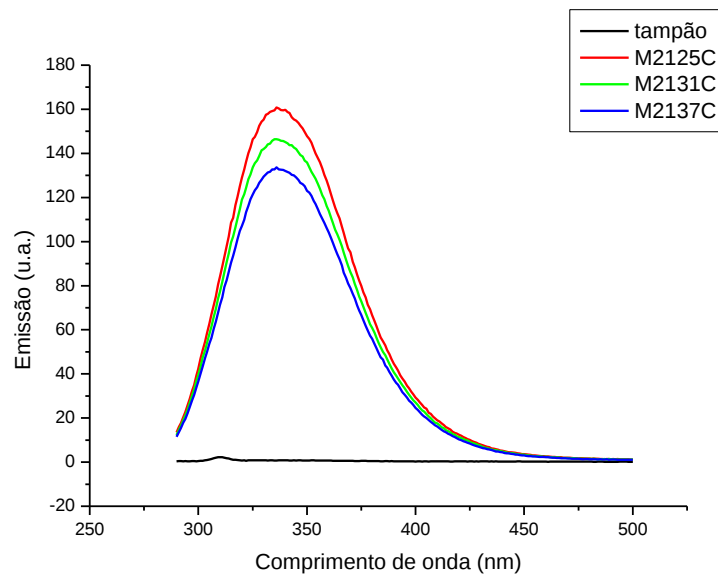


Figura 14: Decaimento da fluorescência da proteína M2-1 com o aumento da temperatura. Após a temperatura do sistema estabilizar, a amostra foi excitada em 280nm e a leitura da emissão foi feita na faixa de 295nm a 500nm. Linha vermelha – 25°C; linha verde – 31°C; linha azul – 37°C; linha preta – emissão do tampão sem amostra.

Então, procedeu-se a titulação da proteína M2-1 com o ligante Quercetina. O experimento foi realizado em três temperaturas diferentes: 25°C, 31°C e 37°C. A concentração da proteína era de aproximadamente 10 μM e a concentração final do ligante variou de 0 a 45,80 μM (Figura 14).

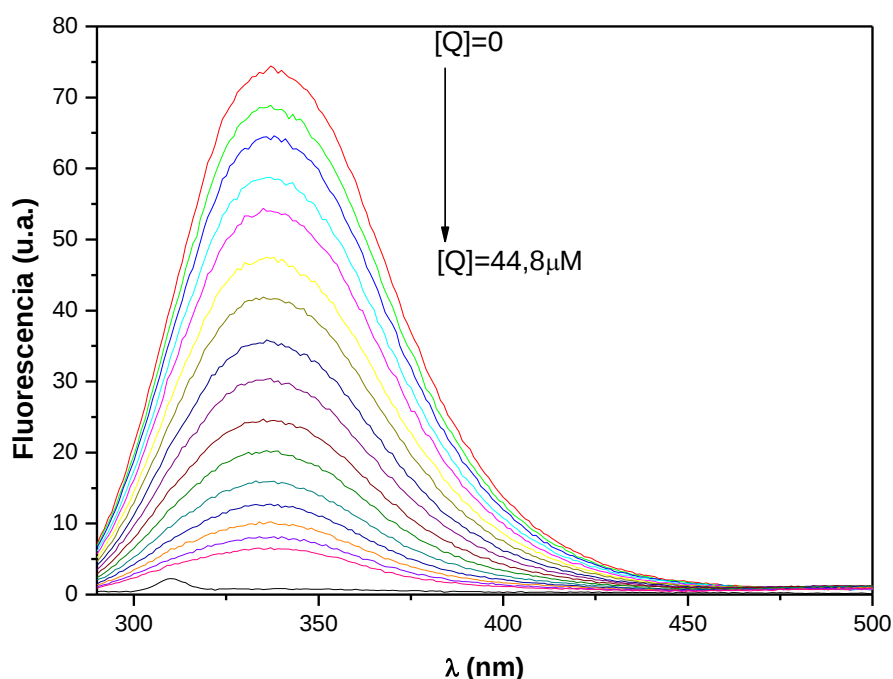


Figura 15: Espectros de emissão de fluorescência referente à titulação de M2-1 na concentração de $10\mu\text{M}$. A proteína foi titulada com quercetina com concentração de quercetina variando de 0 a $44,8\mu\text{M}$ na temperatura de 25°C , onde se verifica a supressão da fluorescência com o aumento da concentração de Quercetina na solução. Os espectros correspondente à 31°C e 37°C apresentam o mesmo perfil.

Ao ser excitado em 280 nm , a proteína apresenta uma emissão máxima em 337 nm . Esse ponto foi usado para acompanhar a supressão da fluorescência. Observou-se que conforme há o aumento da concentração do ligante, há uma redução na emissão da proteína (Figura 15), indicando que o ligante pode estar interagindo com a proteína. Também é possível observar que não há deslocamento do pico de fluorescência da proteína quando acrescida de ligantes, indicando que não há contribuição da fluorescência dos anéis aromáticos dos flavonoides neste comprimento de onda.

Inicialmente foram analisadas as variações de fluorescência da M2-1 com o acréscimo de ligante na solução de forma a se obter uma faixa de variação da razão Ligante/Proteína até chegar próximo à total supressão da fluorescência da proteína. No caso estudado os valores da razão variaram de 0,0 até 4,5, como mostrado no gráfico da Figura 16. O gráfico da variação percentual da fluorescência observada em função da razão Ligante/Proteína permite observar que, em todos os experimentos realizados, foi atingida a região de saturação com uma variação de aproximadamente 90% da

fluorescência inicial observada para M2-1. Verifica-se que este comportamento é semelhante para todas as temperaturas analisadas. A variação da fluorescência observada (f_{obs}) indica a fração da fluorescência inicial que foi acessível ao supressor e seu gráfico é uma modificação do gráfico de Stern Volmer (Lakowicz, 1999), sendo descrita por:

$$f_{obs} = \frac{F_0 - F}{F_0} \quad (13)$$

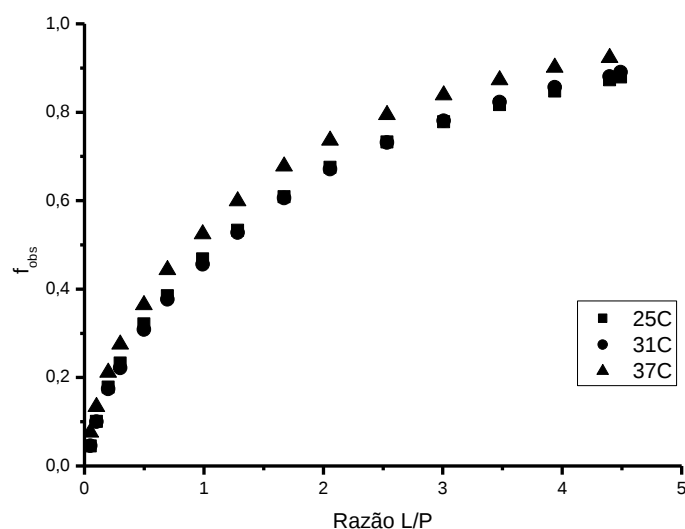


Figura 16: Fração de variação da fluorescência observada (f_{obs}) em função da razão entre as concentrações de Quercetina e da M2-1. Quadrado preto – sistema à 25°C; losangos vermelhos – sistema à 31°C; triângulos verdes – sistema à 37°C.

Nos experimentos realizados, embora a Quercetina tenha sido preparada em etanol 90%, o volume total de etanol na solução contendo a proteína e o ligante, após as titulações, não ultrapassou 4,1% do volume total da solução, não afetando desta forma a estrutura terciária da proteína, uma vez que concentrações acima de 10% de etanol são suficientes para causar discretas alterações na estrutura proteica (Taboada et al., 2007).

Na análise destes resultados é admitido que a interação da Quercetina com a M2-1 suprime totalmente a fluorescência da M2-1, ou seja, após cada adição de Quercetina a fluorescência remanescente corresponde à proteína que não interagiu com a Quercetina. Desta forma, pode-se calcular a fração de proteína complexada com o ligante (f_b) através da variação de fluorescência com o incremento de ligante:

$$f_b = \frac{F_0 - F}{F_0} = \frac{[PL]}{[P_T]}, \quad (14)$$

na qual, F_0 é a fluorescência da proteína M2-1 na ausência de ligante, e F é a fluorescência após adição de ligante, $[PL]$ é a concentração de proteína complexada com n ligantes e $[P_T]$ é a concentração total de proteína na solução.

Desta forma, pode-se construir o gráfico de Scatchard e obter a partir deste o valor da constante de associação e o número de ligantes ligados por proteína.

O resultado desta análise (Figura 17) se verifica que há dois tipos sítios de ligação independentes, sendo um com constante de associação forte ($\sim 1,5 \times 10^6 \text{ M}^{-1}$) sendo ocupado por 0,25 ligante e outro tipo de sítio com constante de associação fraco ($1,1 \times 10^5 \text{ M}^{-1}$) sendo ocupado por um ligante. Na literatura é descrita que a proteína M2-1 apresenta se como um tetrâmero, portanto, este resultado sugere a existência de um sítio de ligação forte formado pelas quatro unidades do tetrâmero e que é ocupado por um ligante (0,25 ligante/proteína=1 ligante/tetrâmero), e o segundo sítio de ligação é ocupado por um ligante (1 ligante/proteína = 4 ligantes/tetrâmero) e a ligação é não cooperativa o que significa que a ligação de uma Quercetina é independente da ligação das demais, não há aumento de afinidade após a primeira Quercetina se ligar, ou seja a interação nos diferentes sítios são independentes.

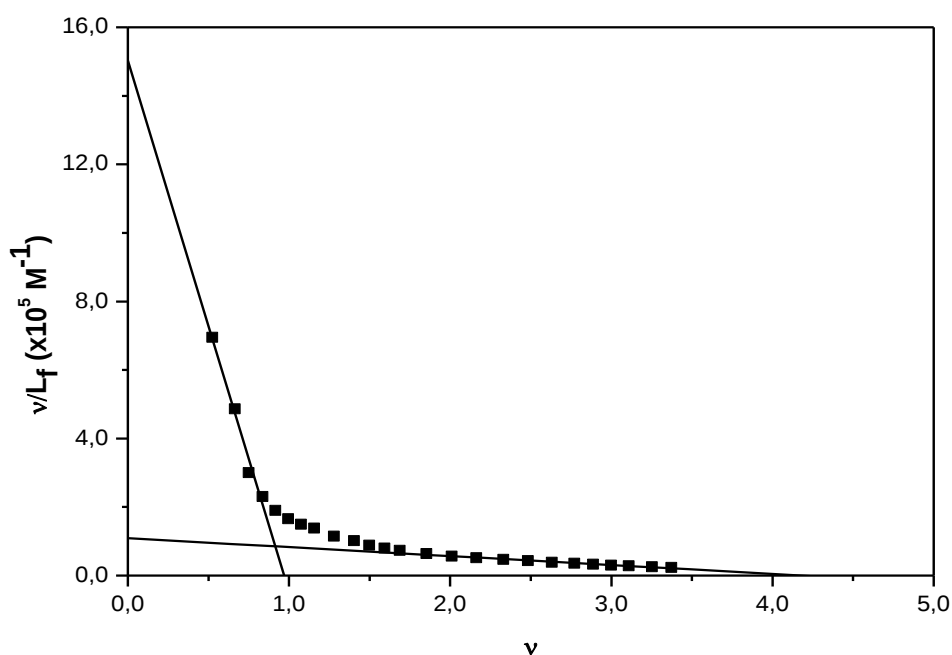


Figura 17: Gráfico de Scatchard relativo à interação da M2-1 com a Quercetina. Concentração da proteína M2-1 5 μ M, temperatura 31°C. Mostra a existência de dois tipos de sítios de ligação diferentes e não cooperativos. As linhas retas representam o melhor ajuste linear.

Com o objetivo de tentar identificar o tipo de interação entre a proteína M2-1 e a Quercetina com os dois tipos de sítio de interação, foram realizadas titulações em três diferentes temperaturas, 25°C, 31°C e 37°C.

Os resultados obtidos são mostrados na tabela 7, os valores das duas constantes de associação K1 (sítio forte) e K2 (sítio fraco) obtidos em três temperaturas diferentes. Pode-se observar que conforme há o aumento da temperatura, as constantes tem uma redução nos seus valores, indicando ligação do tipo estática (Lakowicz, 1999).

Tabela 7: Valores de K1, constante de associação do sítio do tipo 1 e K2, constante de associação do sítio do tipo 2, nas três temperaturas estudadas.

Constant	25°C	31°C	37°C
K1 ($\times 10^6 \text{ M}^{-1}$)	1,85 $\pm$ 0,04	1,60 $\pm$ 0,10	1,50 $\pm$ 0,10
K2 ($\times 10^5 \text{ M}^{-1}$)	0,82 $\pm$ 0,06	1,18 $\pm$ 0,02	1,09 $\pm$ 0,01

Com os valores das constantes pode-se construir um gráfico de Van'Hoff para se obter o valor da variação de entalpia da interação ΔH , que é dado pela declividade do gráfico de $-\ln(K)$ vs $(1/T)$, onde $R=1,98 \text{ cal/MolK}$. Na figura 18, é mostrado este gráfico para as duas constantes de associação, onde se verifica que o sítio forte apresenta variação de entalpia negativa (-3,6Kcal/mol) reação

exotérmica, mostrando que existe e o sítio fraco apresenta variação de entalpia positiva (4,4Kcal/mol) reação endotérmica.

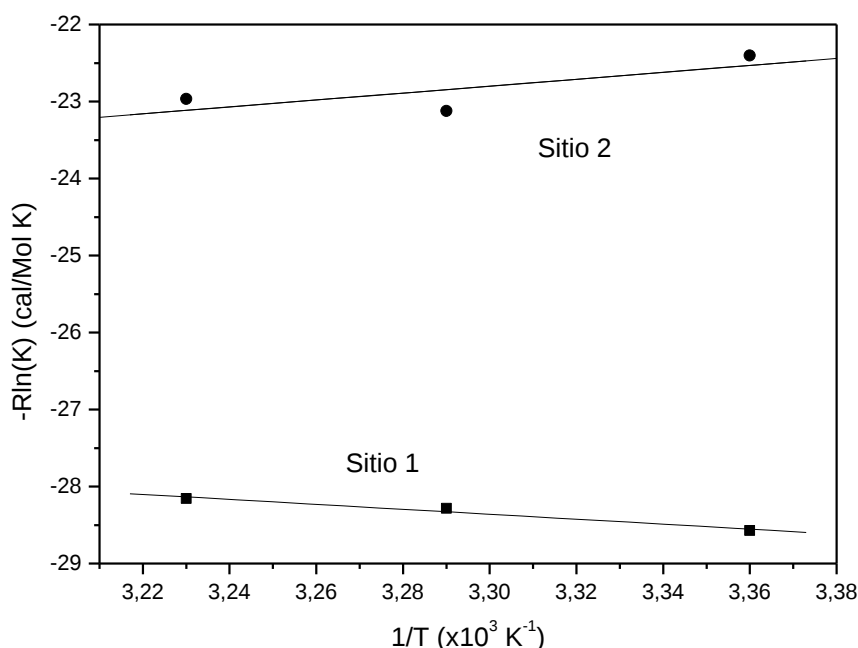


Figura 18: Gráfico de Van'Hoff referente à interação da proteína M2-1 com a Quercetina.

Em seguida foi calculada a energia livre de Gibb com a constante de associação K através da relação:

$$\Delta G = -RT\ln(K) = \Delta H - T\Delta S \quad (15)$$

Pode-se obter os valores da variação da energia livre de Gibb (ΔG), a variação de entalpia (ΔH) e a variação de entropia (ΔS), onde se verifica que a interação entre o ligante e a proteína é energeticamente favorável ($\Delta G < 0$).

O resultado da tabela 4 relaciona os parâmetros termodinâmicos obtidos nos experimentos de titulação espectrofluorimétrica e na figura 19 a variação destes na faixa de temperatura explorada. Esses resultados mostram que os valores de ΔH e ΔS referentes aos dois tipos de sítio, indicam que o sítio forte (sitio 1) estabelece uma interação do tipo iônica ($\Delta H < 0$ e $\Delta S > 0$) e o sítio fraco (sítio 2) estabelece uma interação do tipo hidrofóbica entre o ligante e a proteína ($\Delta H > 0$ e $\Delta S > 0$).

Tabela 8: Valores de ΔG , ΔH e $T\Delta S$ referentes à interação da proteína M2-1 com a Quercetina para os dois tipos de sítios.

Kcal/mol	25°C	31°C	37°C
$\Delta G1$	$-8,15 \pm 0,02$	$-8,60 \pm 0,05$	$-8,73 \pm 0,06$
$\Delta G2$	$-6,7 \pm 0,05$	$-7,00 \pm 0,01$	$-7,10 \pm 0,07$
$\Delta H1$	$-3,2 \pm 0,6$	$-3,2 \pm 0,6$	$-3,2 \pm 0,06$
$\Delta H2$	$4,4 \pm 0,4$	$4,4 \pm 0,4$	$4,4 \pm 0,4$
$T\Delta S1$	$5,3 \pm 0,9$	$5,4 \pm 0,9$	$5,5 \pm 0,9$
$T\Delta S2$	$11,1 \pm 0,8$	$11,4 \pm 0,2$	$11,5 \pm 0,1$

$\Delta G1$ = variação energia livre sítio do tipo 1, $\Delta G2$ =variação de energia livre sítio do tipo 2, $\Delta H1$ = variação de entalpia sítio do tipo 1, $\Delta H2$ = variação de entalpia sítio do tipo 2, $T\Delta S1$ = variação de entropia sítio do tipo 1, $T\Delta S2$ = variação de entropia sítio do tipo 2.

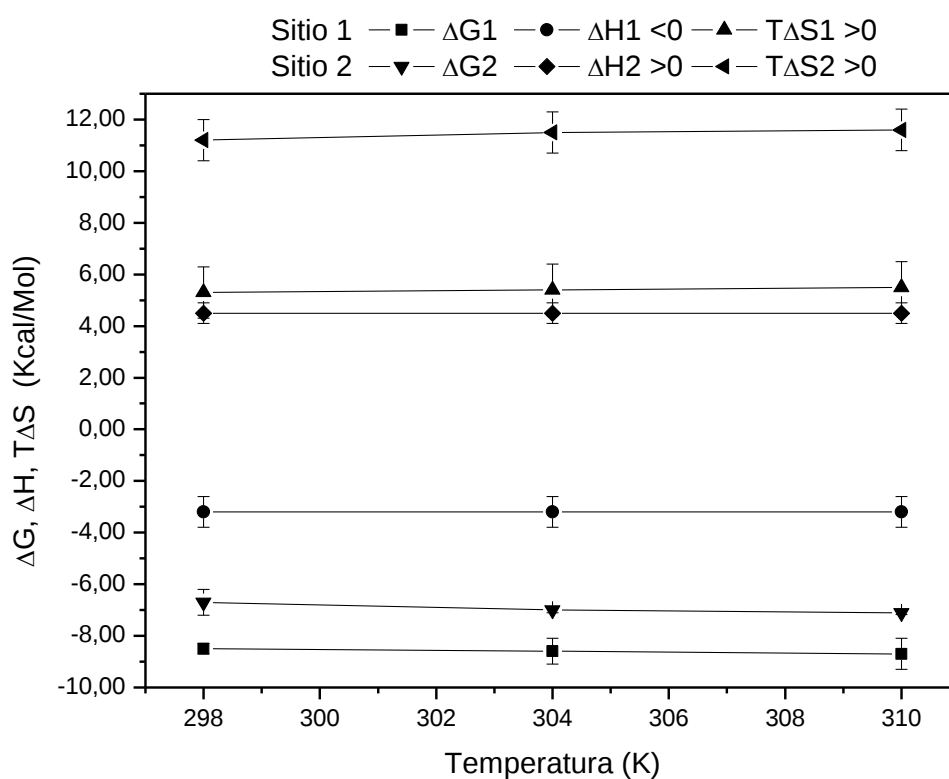


Figura 19: Dependência de ΔG , ΔH e $T\Delta S$ com a temperatura, mostrando a diferença de interação entre a Quercetina e a M2-1 para os sítios do tipo 1 e tipo 2.

4.8. Bioinformática

4.8.1. Predição estrutural

No início deste projeto, somente o domínio central da proteína M2-1 (domínio de ligação ao RNA e à proteína P) havia sido determinado por NMR

(Blondot et al., 2012). Como a estrutura da proteína estava incompleta, realizamos a predição da estrutura da proteína completa pelo servidor I-Tasser, que gerou cinco estruturas, sendo que a estrutura de número 1 (Figura 20 – A) apresentou maior C-score (0,31) e TM-score de $0,75 \pm 3,0$ (Zhang, 2008; Roy, Kucukural e Zhang 2010; Yang, et al., 2015). Além do modelo 1 apresentar parâmetros melhores (de acordo com as análises do próprio servidor) ele apresenta um RMSD com a estrutura cristalina de 0,254Å.

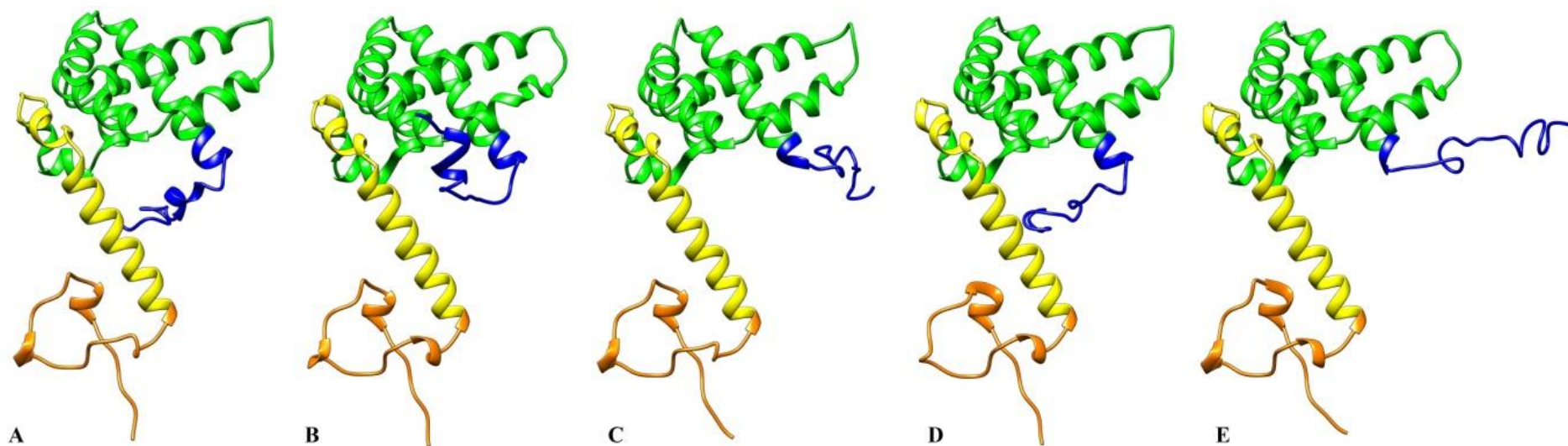


Figura 20: Modelos da proteína M2-1 gerado pelo servidor I-Tasser. O modelo 1 (A) apresentou um C-Score de 0,31, o modelo 2 (B) apresentou um C-Score de -0,54, o modelo 3 (C) de -2,67, o modelo 4 (D) 0,30 e o modelo 5 (E) -3,86. As estruturas estão coloridas por domínios: laranja – dedo de zinco; amarelo – oligomerização; verde – ligação à proteína P e RNA; azul – região desestruturada.

Como a proteína atua na forma tetramérica, usamos o servidor SymmDock para gerar o dímero, trímero e tetrâmero da proteína. O servidor gerou 100 estruturas para o dímero, trímero e tetrâmero. Os modelos escolhidos foram aquele que apresentaram maior contato entre os domínios de oligomerização, sendo que para o dímero foi escolhido o modelo 01 (score 19006), para o trímero o modelo 18 (score 6204) e para o tetrâmero o modelo 06 (score 7848) (Figura 21).

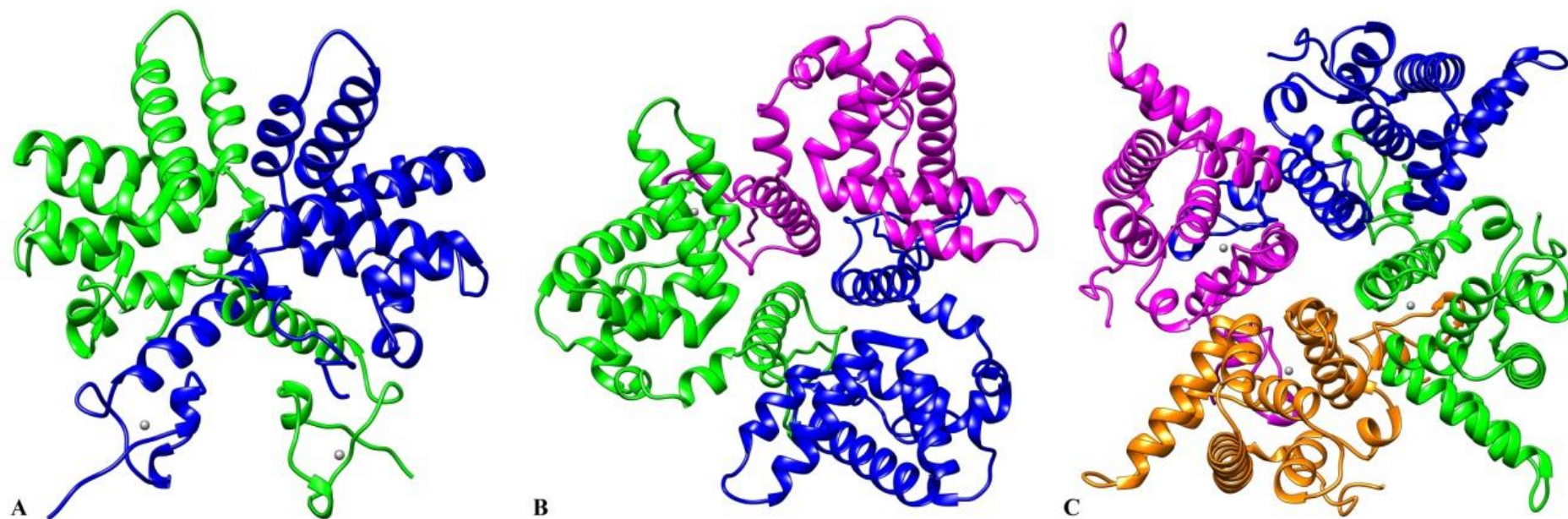


Figura 21: Modelos dos oligômeros da proteína M2-1 geradas pelo servidor SymmDock (dímero – A; trímero – B; tetrâmero – C). (verde – cadeia A, azul – cadeia B, magenta – cadeia C, laranja – cadeia D).

4.8.2. Dinâmica molecular.

Pela análise das estruturas da proteína M2-1, observamos que o domínio central (Gly62-Ile173) é estável, não apresentando variação estrutural no decorrer do tempo da simulação. As regiões que apresentam maiores movimentações estruturais são as regiões N-terminal e C-terminal.

A análise de RMSD (do inglês Root Mean Square Deviation) é geralmente usada como um indicativo da convergência da estrutura através de um estado de equilíbrio. Na forma monomérica da proteína M2-1, nos primeiros 10 ns de simulação, a variação estrutural foi de aproximadamente 0,7 nm (Figura 22 A). Nesse tempo, o domínio de oligomerização se aproxima do domínio central, fazendo com que a proteína sofra uma leve compactação (o raio de giro reduz em aproximadamente 0,3 nm) (Figura 22 B). No decorrer do tempo, a proteína mantém essa estrutura, sendo que os resíduos N-terminal (Met1-His25) e C-terminal (Asn175-Thr194) apresentam grande movimentação no decorrer do tempo (Figura 26 A). Também, a alça formada pelos resíduos Arg115-Asn129 apresenta movimentação no tempo de simulação.

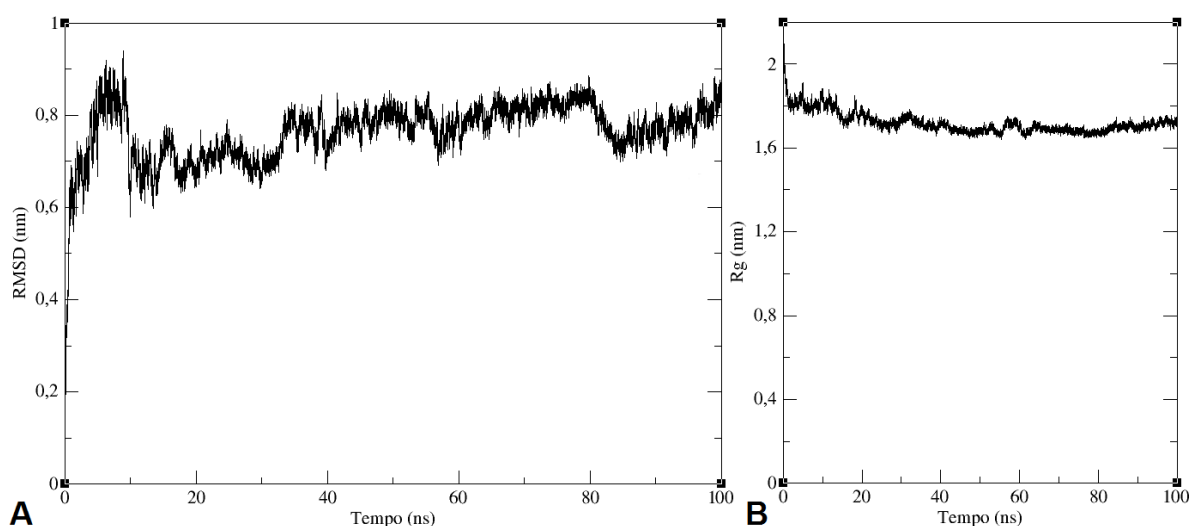


Figura 22: RMSD (A) e raio de giro (B) do *backbone* da proteína M2-1 monomérica em função do tempo à 313K.

Já o dímero da proteína M2-1 demora em torno de 50 ns para entrar em equilíbrio (Figura 23 A). Nesse caso, nesse tempo houve uma aproximação das cadeias promovendo uma leve compactação da estrutura dimérica (aproximadamente 0,2 nm) (Figura 23 B). A variação da região N-terminal foi menor, sendo observada maior variação entre os resíduos Arg12 e Asn26 para ambas as cadeias. Houve pouca movimentação dos resíduos da região de oligomerização da cadeia A, porém

houve variação estrutural dos resíduos Thr69-Ala73 do mesmo domínio da cadeia B (Figura 26 B). Ambas as cadeias apresentaram variação estrutural da alça formada pelos resíduos Asn175-Thr194. Quando se observa a cadeia A, a variação dos últimos 24 resíduos foi maior.

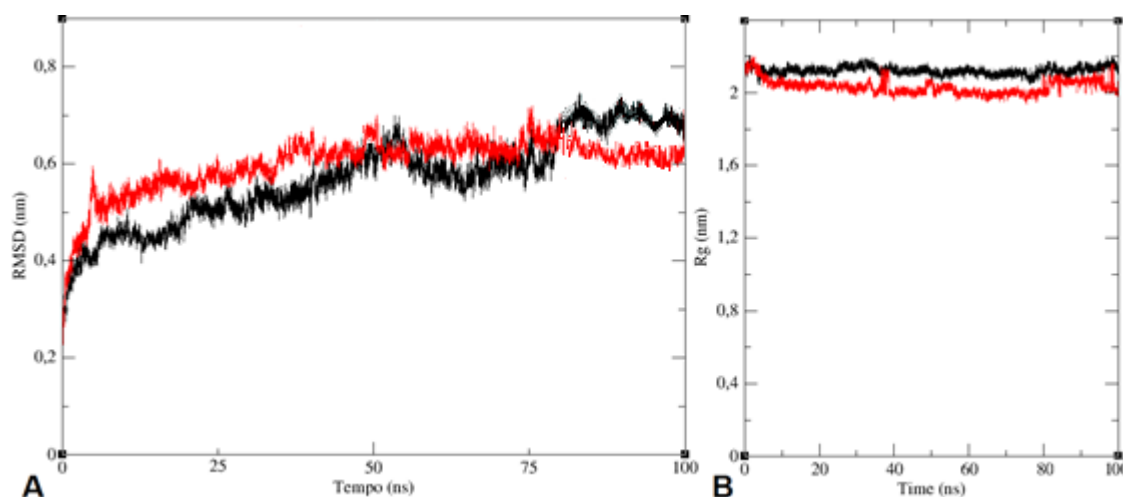


Figura 23: RMSD (A) e raio de giro (B) do *backbone* da proteína M2-1 dimérica em função do tempo à 313K. Cadeia A – preto, Cadeia B – vermelho.

A estrutura trimérica adquire uma estabilização em torno de 15 ns (Figura 24 A), sendo que as cadeias A e C foram as que tiveram maior variação em relação a estrutura inicial (aproximadamente 0,12 nm). A análise do raio de giro mostra que a cadeia A sofreu uma leve compactação (aproximadamente 0,2 nm) e a cadeia C uma leve expansão (aproximadamente 0,2 nm) (Figura 24 B). Somente a cadeia B manteve a sua estrutura inalterada. A análise do RMSF mostra que as principais variações foram nos resíduos Arg12 a Asn26 da região de oligomerização e nas alças formadas pelos resíduos Thr69-Ala73 e Asn175-Thr194 (Figura 26 C). Novamente, houve maior variação estrutural nos últimos 24 resíduos terminais.

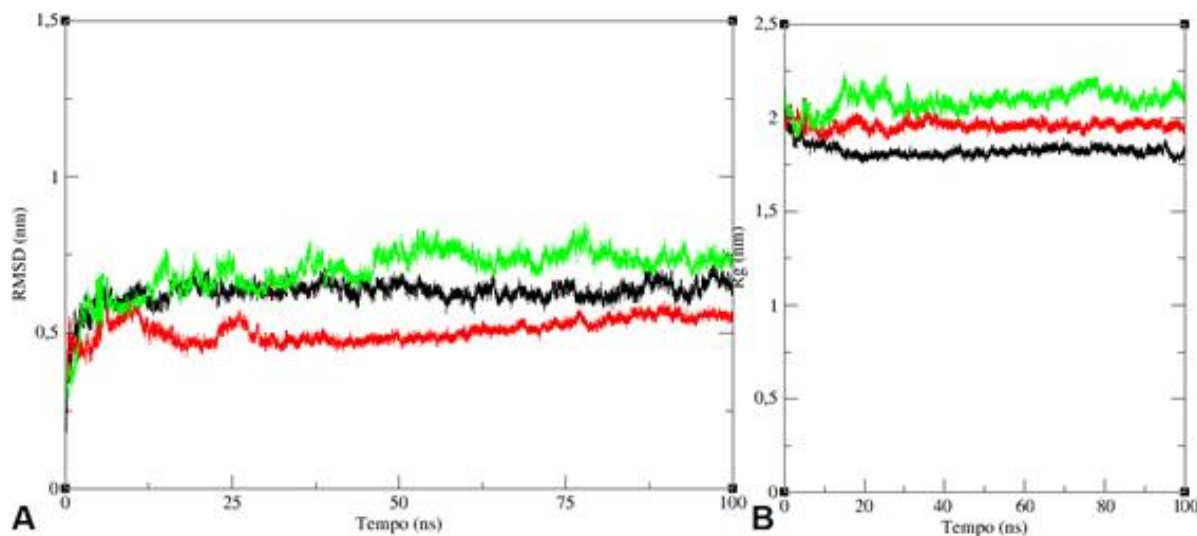


Figura 24: RMSD (A) e raio de giro (B) do *backbone* da proteína M2-1 trimérica em função do tempo à 313K. Cadeia A – preto, Cadeia B – vermelho; Cadeia C – verde.

A dinâmica molecular da proteína M2-1 tetramérica ainda não está finalizada, mas a análise prévia mostra que a estrutura começa a entrar em equilíbrio em torno de 30 ns (Figura 25 A) e que as cadeias B e D sofrem uma leve compactação (aproximadamente 0,1 nm) (Figura 25 B). A análise do RMSF mostra que os resíduos do tetrâmero sofrem pequena variação estrutural, principalmente os últimos resíduos terminais, comparado ao monômero, dímero e trímero (Figura 26 D).

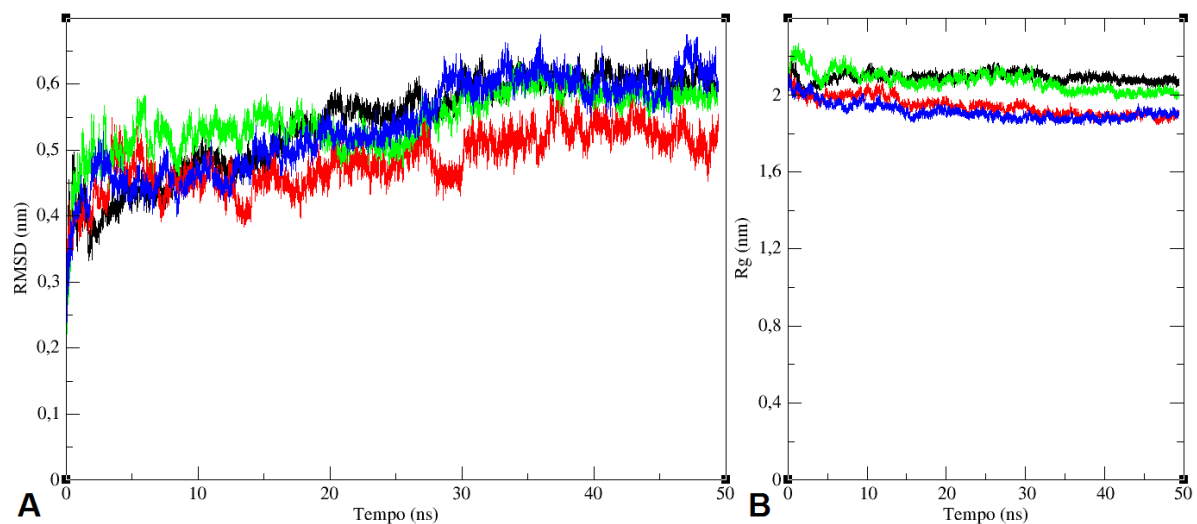


Figura 25: RMSD (A) e raio de giro (B) do *backbone* da proteína M2-1 tetramérica em função do tempo à 313K. Cadeia A – preto, Cadeia B – vermelho; Cadeia C – verde; Cadeia D – azul.

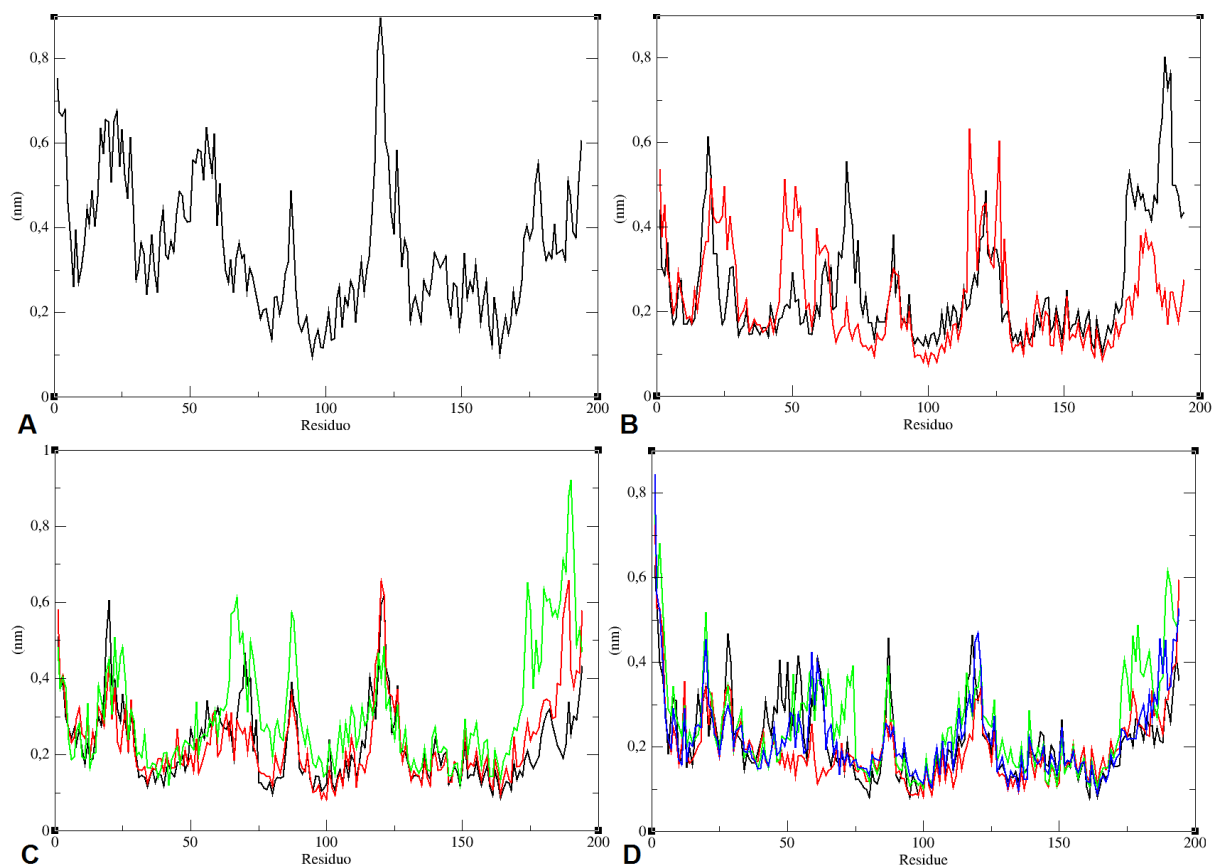


Figura 26: RMSF da proteína M2-1 monomérica (A), dimérica (B), trimérica (C) e tetramérica (D) em função do tempo à 313K. Cadeia A – preto, cadeia B - vermelho, cadeia C – verde, cadeia D – azul.

Para uma melhor visualização da variação da estrutura da proteína no decorrer do tempo, foi feito a análise de DSSP (do inglês *Database of Secondary Structure of the Protein*) que mostra a estrutura secundária de cada resíduo no decorrer do tempo (Figura 27 a 30).

Pela análise da estrutura secundária das cadeias da proteína M2-1, observou-se que o domínio de ligação à proteína P e ao RNA, formado pelos resíduos 70-175, apresenta uma estruturação bem estável. Os resíduos N-terminais são menos estruturados na forma monomérica e começam a adquirir estabilidade estrutural conforme migra para a forma tetramérica. Já os resíduos C-terminais apresentam um padrão de *coil*, independente da oligomerização (Figura 27).

Posteriormente, com a publicação da estrutura tetramérica da proteína M2-1 determinada por raio-X (Tanner et al., 2014) foi possível comparar se o modelo predito se assemelha ao modelo cristalino. A sobreposição das duas estruturas mostra que o RMSD é de 11,262 Å. Ao observarmos as estruturas sobrepostas (Figura 28), notamos que as cadeias se sobrepõem, formando o poro central e os

domínios são bem formados. Devemos ressaltar que embora a convergência não seja idêntica, a estrutura obtida da dinâmica molecular estava em movimento em um sistema aquoso, diferente da estrutura cristalina que não apresenta movimentos e, devido às condições da cristalização pode ter sua estrutura um pouco alterada.

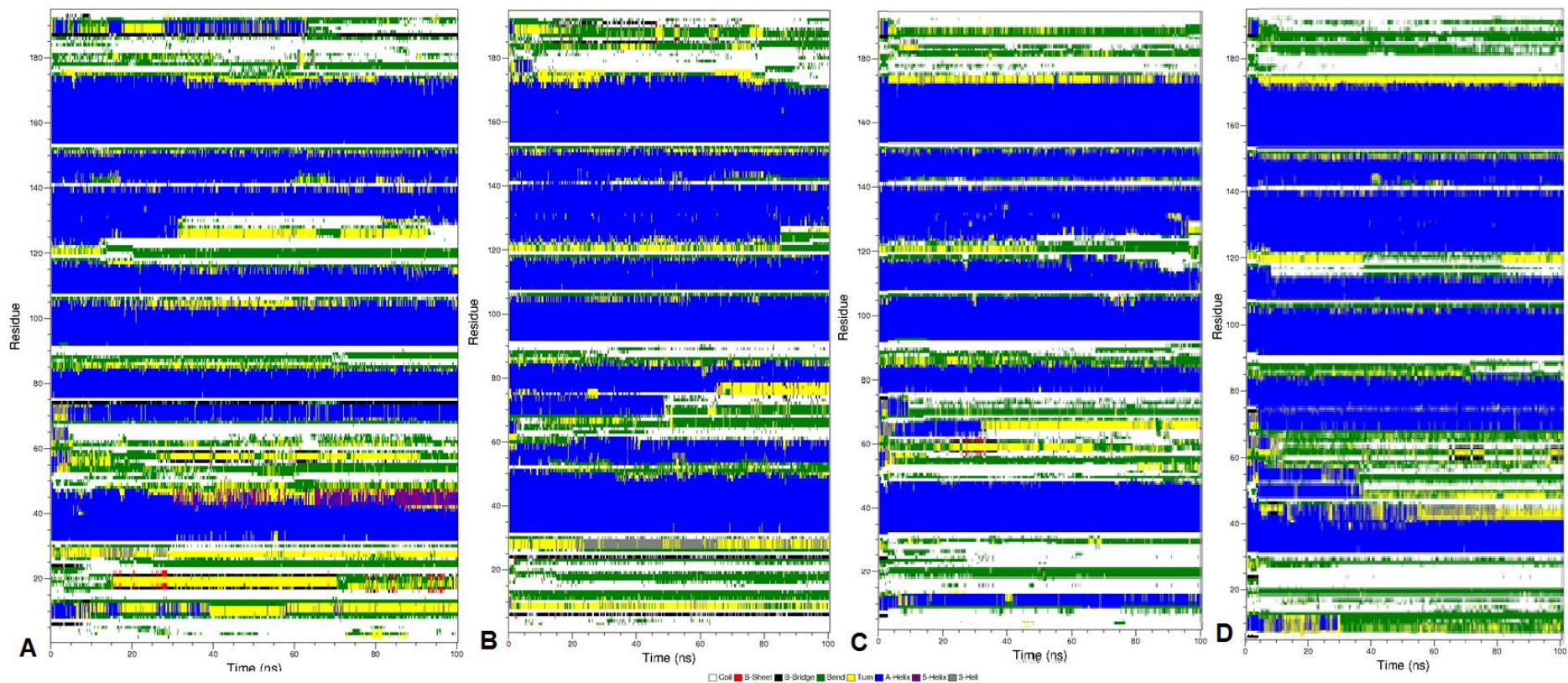


Figura 27: Gráfico padrão da estrutura secundária da proteína M2-1 monomérica (A) e, cadeias A da proteína dimérica (B), trimérica (C) e tetramérica (D).dimérica no solvente água.

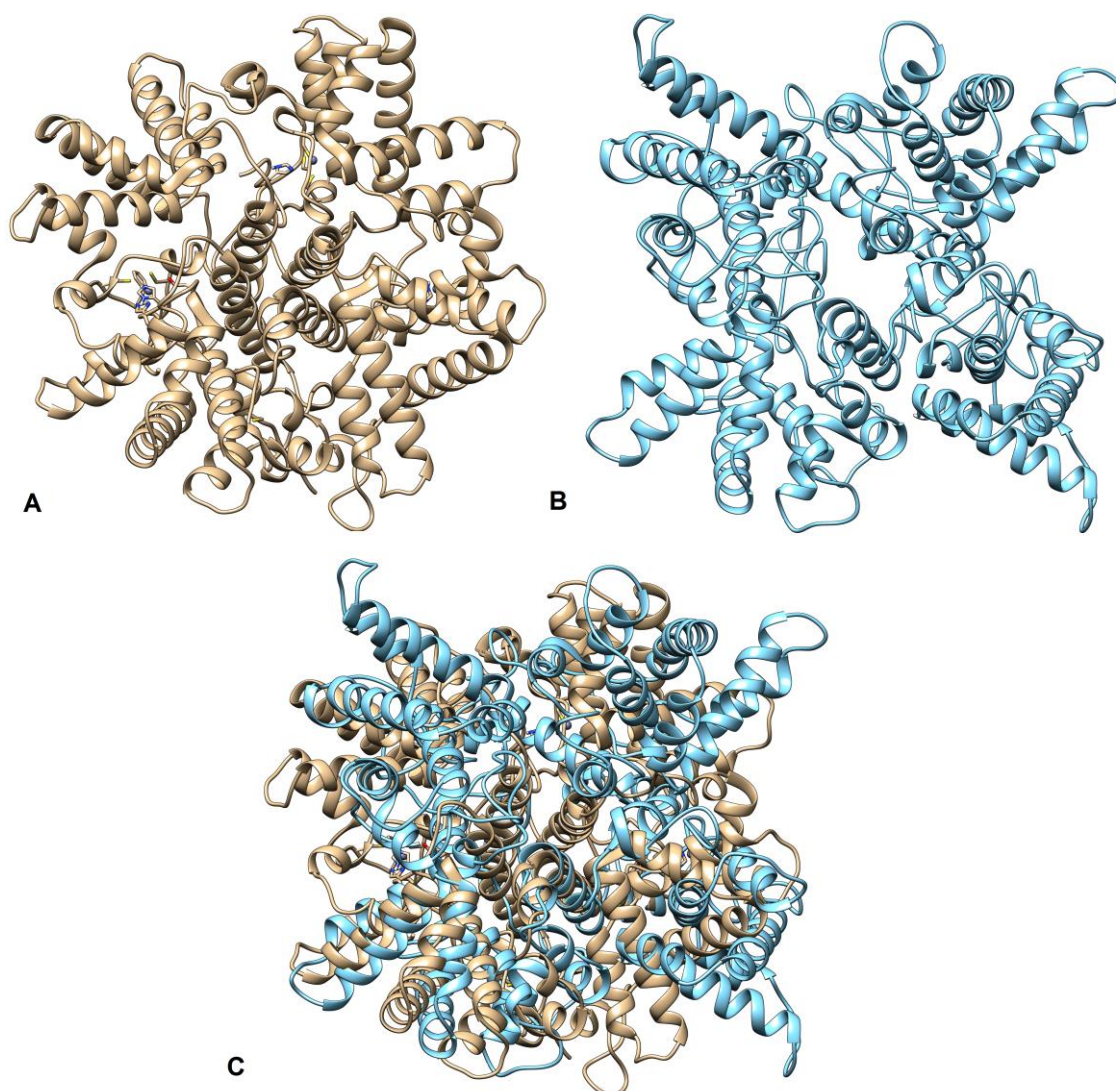


Figura 28: Sobreposição da estrutura obtida pela dinâmica molecular com a estrutura cristalina da proteína M2-1. Em ambos os modelos, há a formação do poro central e os domínios ficam definidos. M2-1 cristalina (A) (PDB = 4C3B), M2-1 obtido pela dinâmica (B) e sobreposição das estruturas (C).

4.8.3. Docking molecular.

Com a análise do *docking* molecular, podemos inferir como a Quercetina interage com a proteína M2-1. O estudo de *docking* molecular foi feito com a proteína M2-1 na forma monomérica e tetramérica (ambas as estruturas cristalinas).

O monômero foi dividido em duas regiões para fazer o estudo de *docking* (Figura 29). O sítio 1-1 (sítio 1 do monômero) é composto pelo domínio globular (domínio de ligação ao RNA e proteína P) e a porção C-terminal da proteína e, sítio 2-1 (sítio 2 do monômero) compreende o domínio “dedo de zinco” e o domínio de oligomerização.

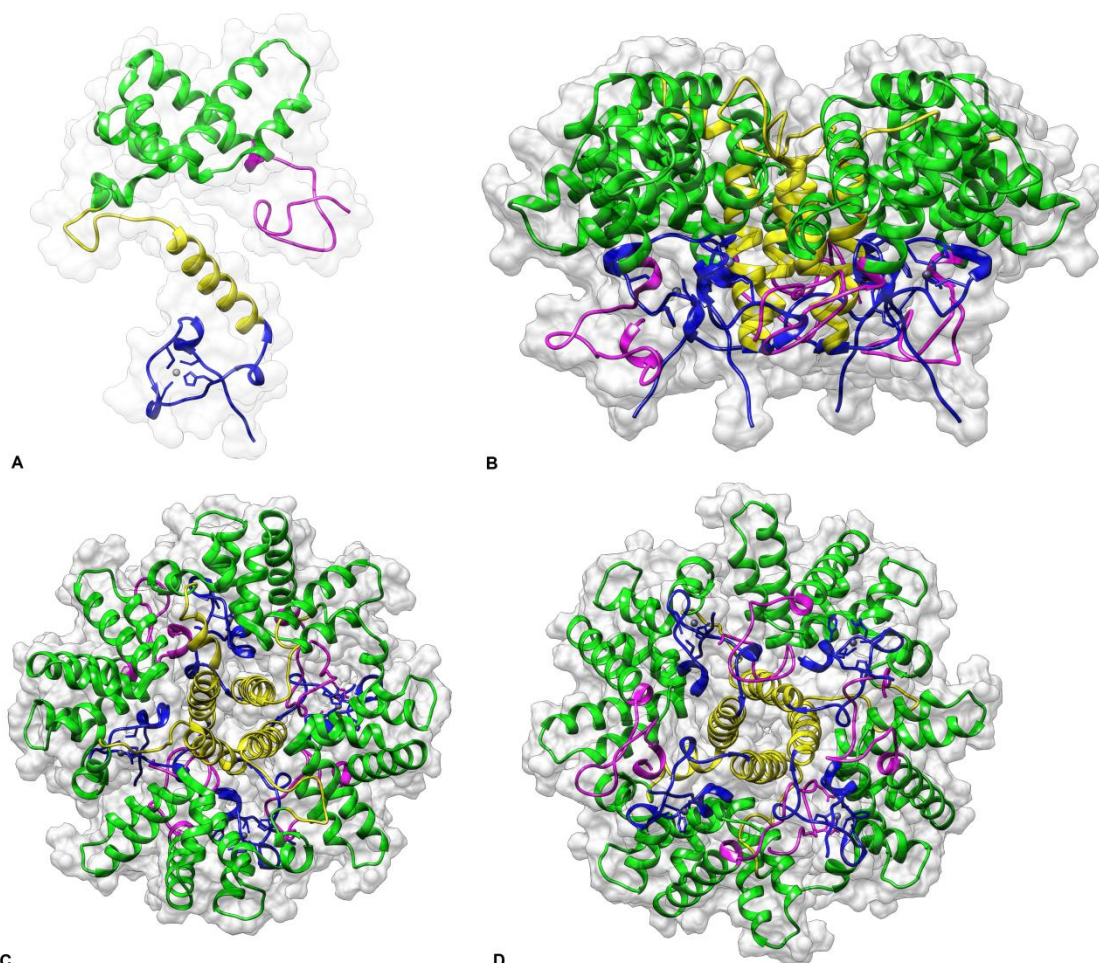


Figura 29: Domínios da proteína M2-1. Visão da proteína M2-1 monomérica (A), M2-1 tetramérica face lateral (B), face N-terminal (C) e face C-terminal (D). Os domínios foram corados em: azul – domínio dedo de zinco; amarelo – domínio de oligomerização; verde – domínio de ligação ao RNA e à proteína P; magenta – domínio C-terminal não estruturado.

No sítio 1-1, a Quercetina tem uma energia de interação de $-6,41 \text{ kcal/mol}$ (Tabela 9) e faz três ligações de hidrogênio (Met50, Asp51 e Leu103). O ligante interage em uma cavidade entre o domínio de oligomerização e o domínio globular (Figura 30 A). No sítio 2-1, a Quercetina tem uma energia de interação de $-5,21 \text{ kcal/mol}$ (Tabela 9) e faz três ligações de hidrogênio (His14 e Phe41). Neste sítio, o ligante interage com o domínio dedo de zinco, próximo ao metal de zinco (Figura 30 B).

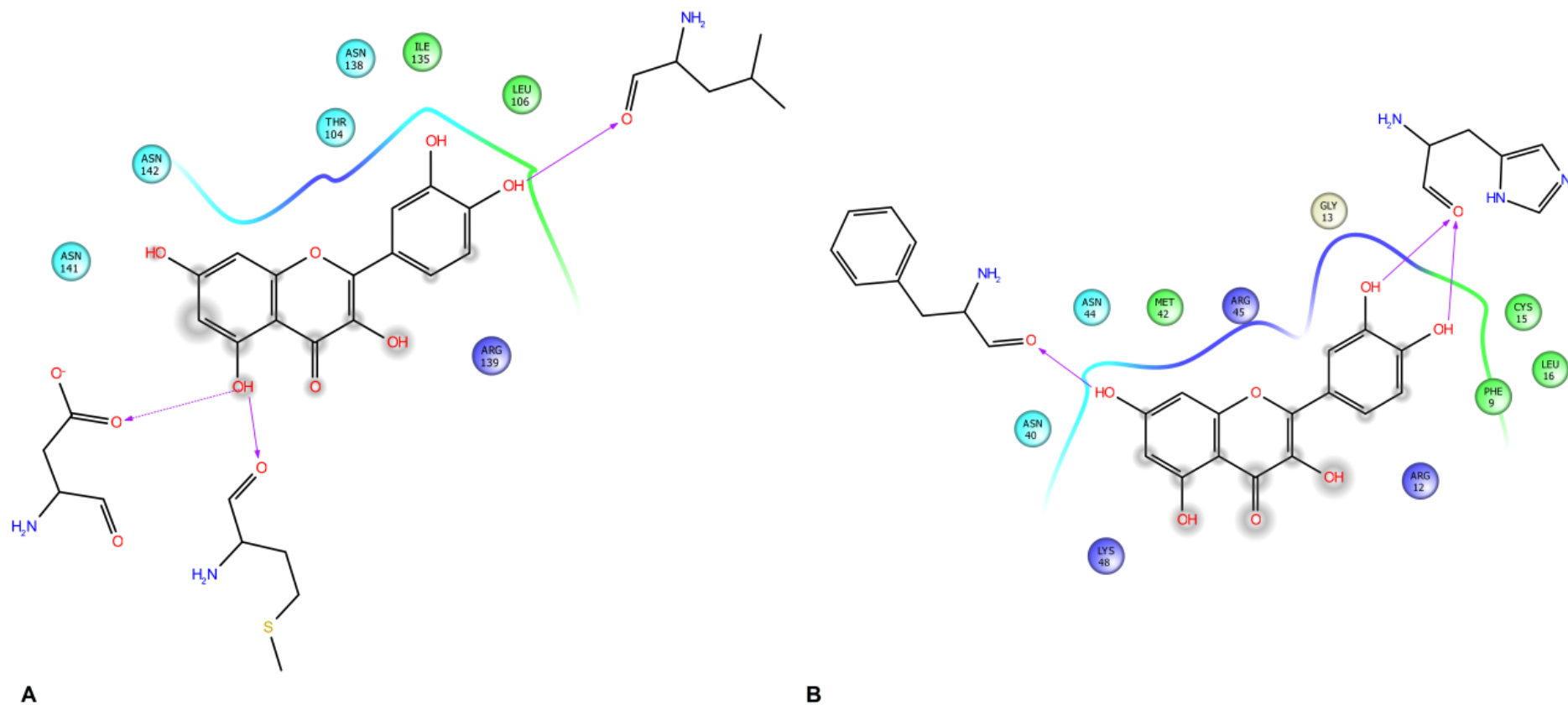


Figura 30: Docking molecular da Quercetina na proteína M2-1 monomérica. Quercetina foi docada no sítio 1 (A) (domínio de ligação ao RNA e à proteína P e resíduo N-terminal) e no sítio 2 (B) (domínio dedo de zinco e de oligomerização). Legenda dos resíduos: azul escuro – carregado positivamente; azul claro – polar; verde – hidrofóbico; branco – glicina; seta roxa – ligação de hidrogênio; sombreado cinza – exposto ao solvente.

Embora o *docking* mostre que a energia de interação da Quercetina no monômero no sítio 1 (domínio globular) é maior que no sítio 2 (domínio “dedo de zinco” e de oligomerização), o mesmo não pode ser afirmado sobre a constante de afinidade. De acordo com a equação da energia livre de Gibbs e a relação logarítmica entre energia livre de ligação e afinidade, uma diferença de 10 vezes na atividade representa apenas uma diferença de 1,4 kcal/mol. Consequentemente, é raramente possível prever acuradamente diferenças na afinidade de interação menores que 10 vezes porque o erro na função de pontuação em análises computacionais rigorosas, como a perturbação de energia livre, está na faixa de 1,4 – 2,0 kcal/mol (Stroud et al., 2008).

O *docking* na M2-1 tetramérica foi realizado em três sítios que englobam toda a proteína (Figura 29). O ligante Quercetina foi docado na cavidade central na face N-terminal (sítio 1-4 – sítio 1 da proteína tetramérica), na face C-terminal (sítio 2-4 – sítio 2 do tetrâmero) e na interface entre duas cadeias (sítio 3-4 – sítio 3 do tetrâmero) (O termo “faces” foi retirado de Tanner et al., 2014).

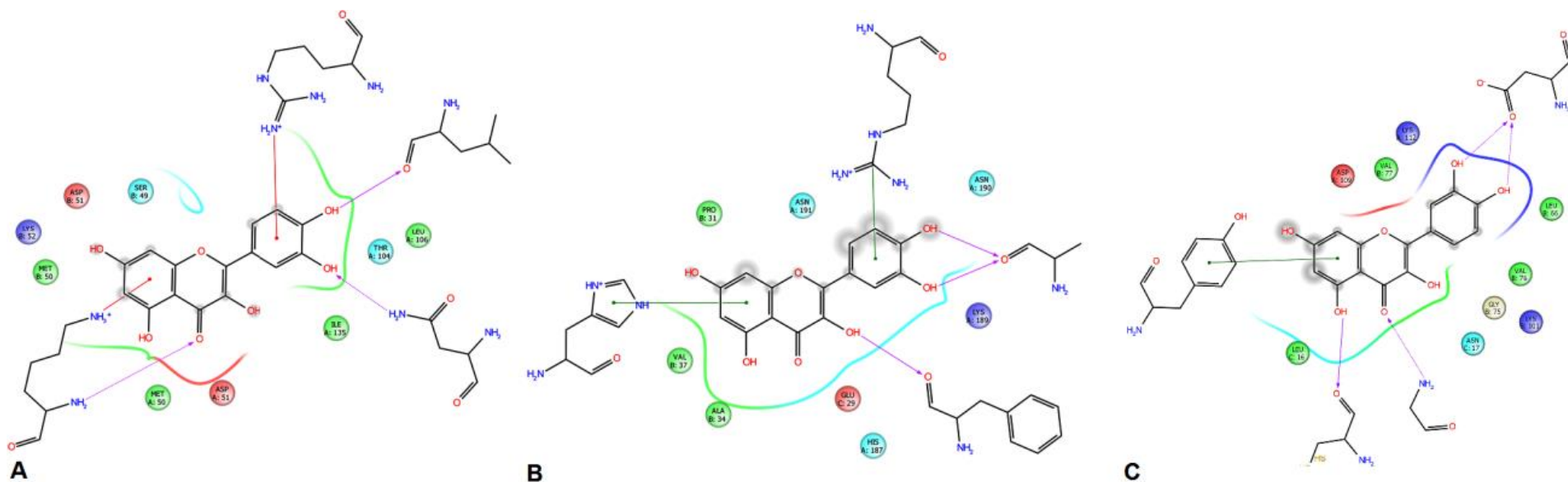


Figura 31: *Docking* molecular da Quercetina na proteína M2-1 tetramérica. A Quercetina foi docada no sítio 1 (A) (face N-terminal), sítio 2 (B) (face C-terminal) e no sítio 3 (C) (interface entre duas cadeias) Legenda dos resíduos: azul escuro – carregado positivamente; vermelho – carregado negativamente; azul claro – polar; verde – hidrofóbico; branco – glicina; seta roxa – ligação de hidrogênio; linha verde – interação pi-pi/pi-cátion; sombreado cinza – exposto ao solvente.

No sítio 1-4 a energia de interação foi de -7,36 kcal/mol (Tabela 9). O ligante interagiu numa região entre duas cadeias (Figura 31 A) fazendo três interações de hidrogênio (Lys52, Leu103 e Asn138) e duas interações do tipo pi-cátion (Lys52 e Arg139). No sítio 2-4, a energia de interação foi de -5,24 kcal/mol (Tabela 9) e a Quercetina faz três interações de hidrogênio (Ph28 e Ala188) e duas interações do tipo pi-pi (Arg3 e His33) (Figure 31 B). A molécula interage em uma região entre o domínio de “dedo de zinco” e o domínio C-terminal não estruturado próximo ao domínio de oligomerização. E para o sítio 3-4, a energia de interação foi de -7,36 kcal/mol (Table 9) e faz quatro interações de hidrogênio (Cys15, Gly18 e Asp67). O flavonol Quercetina interage em uma cavidade entre o domínio de interação com RNA e proteína P e o domínio de oligomerização, próximo ao metal zinco (Figure 31 C).

Tabela 9: Valores da energia de interação (kcal/mol) da Quercetina nos sítios 1, 2 e 3 do monômero e do tetrâmero.

	Monômero	Tetrâmero
	Energia de ligação	
Sítio 1	-6.41	-7.36
Sítio 2	-5.21	-5.24
Sítio 3	-	-7.36

O *docking* da Quercetina na proteína tetramérica mostra que as características dos sítios 1 e 2 são quase semelhantes. No sítio 1 a Quercetina é estabilizada por três ligações de hidrogênio e dois emparelhamentos pi-cátion e há cinco resíduos hidrofóbicos em torno da molécula e oito resíduos polares. No sítio 2 a Quercetina é estabilizada por três ligações de hidrogênio, dois emparelhamentos pi-pi/pi-cátion e há cinco resíduos hidrofóbicos e seis polares. O sítio 3 tem uma característica mais hidrofóbica que os outros dois. Nesse sítio, a Quercetina faz quatro ligações de hidrogênio, um emparelhamento pi-pi e tem oito resíduos hidrofóbicos e 5 polares.

Pelo *docking* podemos postular que a Quercetina interage preferencialmente no sítio 1 que no sítio 2, próximo à hélice $\alpha 2$ (na qual M2-1 interage com RNA) (Blondot et al., 2012). Também, essa hipótese é reforçada pelo fato que a Quercetina tem algumas similaridades com adenina, na qual já

se observou que os resíduos Ser58-Lys177 tem alta afinidade pelos nucleotídeos uracila e adenina (Tanner et al., 2014).

A característica da interação no sítio com maior energia de interação (sítios 1 e 2) e o no sítio com menor energia de interação (sítio 3) correlaciona com os dados experimentais, já que a interação no sítio 1 e 2 no *docking* tem um padrão de interação iônica e, a interação no sítio 3 tem um padrão de interação hidrofóbico.

5. Conclusões

A proteína M2-1 é uma proteína chave no processo de replicação do vírus sincicial respiratório humano (hRSV) por é responsável pela completa transcrição dos genes a jusante, promove a interação com a proteína P que está envolvida na formação do complexo polimerase e, na montagem da partícula viral, ela auxilia a migração da proteína M na migração para a membrana e liberação do vírus. Assim, moléculas que interajam com essa proteína e bloqueiem a sua atividade é de grande utilidade para tratar infecções que já estão em curso. Os flavonoides são compostos aromáticos encontrados em plantas e frutas cítricas que já mostraram ter atividade antimicrobiana, antiviral e antifúngica. O estudo da interação do flavonoide Quercetina com a proteína M2-1 mostra resultados promissores para esse composto. Com esse trabalho podemos concluir:

- i. A proteína M2-1 é composta por sua maioria de α -hélice, sendo uma proteína bem estruturada com temperatura de *melting* de aproximadamente 55°C.
- ii. Os resultados mostram dois tipos de sítios de ligação independentes, sendo um com constante de associação forte ($\sim 1,5 \times 10^6 \text{ M}^{-1}$) e outro tipo de sítio com constante de associação fraco ($1,1 \times 10^5 \text{ M}^{-1}$).
- iii. Por se tratar de um tetrâmero, segundo dados de cristalografia, nossos resultados sugerem a existência de um sítio de ligação forte formado pelas quatro unidades do tetrâmero e que é ocupado por um ligante (0,25 ligante/proteína=1 ligante/tetrâmero), e o segundo sítio de ligação é ocupado por um ligante (1 ligante/proteína=4 ligantes/tetrâmero);
- iv. A interação da Quercetina é não cooperativa o que significa que a ligação de uma Quercetina é independente da ligação das demais, não há aumento de afinidade após a primeira Quercetina se ligar, ou seja, a interação nos diferentes sítios é independente;
- v. Esses resultados mostram que os valores de ΔH e ΔS referentes à interação da Quercetina com os dois tipos de sítio, indicam que o sítio forte

(sítio 1) estabelece uma interação do tipo iônica ($\Delta H < 0$ e $\Delta S > 0$) e o sítio fraco (sítio 2) estabelece uma interação do tipo hidrofóbica entre o ligante e a proteína ($\Delta H > 0$ e $\Delta S > 0$).

- vi. Sabe-se que a proteína M2-1 exerce a sua função quando na forma tetramérica. Pela análise da dinâmica molecular do monômero, dímero, trímero e tetrâmero, observamos que independente da forma monomérica ou oligomérica o domínio de ligação ao RNA e à proteína P é estável e sua estrutura secundária e terciária não varia com o tempo. Já o domínio de oligomerização tende a se empacotar quando na forma monomérica, porém quando a proteína forma oligômeros, esse domínio mantém a sua forma. Também, a região N-terminal apresenta alta variação estrutural e conforme a proteína migra para a forma de tetrâmero, a mobilidade dessa região diminui. Embora não se saiba a função exata e o mecanismo de ação da região N-terminal (onde está localizado o “dedo de zinco”), a estruturação dessa região é vital para a eficiência da proteína.
- vii. O *docking* molecular da Quercetina na proteína monomérica mostra que a interação ocorre em cavidades superficiais entre a porção C-terminal da proteína e o domínio de interação ao RNA e à proteína P. Essa interação está em uma cavidade polar estabilizada por ligações de hidrogênio. No entanto, o *docking* com a forma tetramérica ocorre em dois pontos, na face N-terminal próximo à hélice $\alpha 2$ e entre cadeias na interface de oligomerização e o domínio de interação à proteína P e ao RNA. No primeiro caso, a interação é em uma cavidade com perfil iônico estabilizada por ligações de hidrogênio e emparelhamento do tipo pi-pi/pi-cátion. No segundo caso, a Quercetina interage em uma cavidade mais hidrofóbica estabilizada por ligação de hidrogênio e emparelhamento pi-pi. Então, a Quercetina pode interagir com a proteína em formação evitando sua tetramerização e interagir com a M2-1 tetramérica, assim, evitando a replicação e liberação do vírus.
- viii. Pela análise do docking propomos que um ligante com um núcleo formado por dois anéis e grupos polares são essenciais para a estabilização hidrofóbica e iônica, respectivamente.

Referências

- AGUIAR, C. L.; ALENCAR, S. M.; TSAI, S. M.; PARK, Y. K. Transformações Enzimáticas de Flavonóides. B.CCEPPA, 25, 2007, 61-76.
- AHERNE W, BIRD T, COURT SD, et al., Pathological changes in virus infections of the lower respiratory tract in children. J Clin Pathol., 1970, 23, 7–18.
- ALCALA, J. R., GRATTON, E., & PRENDERGAST, F. G. Fluorescence lifetime distributions in proteins. Biophysical Journal, 51, 1987a, 597–604.
- ALCALA, J. R., GRATTON, E., & PRENDERGAST, F. G. Interpretation of fluorescence decays in proteins using continuous lifetime distributions. Biophysical Journal, 51, 1987b, 925–936.
- ALMEIDA, J. R. G. S., LIMA, J. T., NUNES, X. P., CERQUEIRA, G. S., FIGUEIREDO, C. S. Flavonóides da dieta e sua importância para a saúde humana. Revista de Ciências da Saúde Nova Esperança. , v.2, 2004, 65 - 76.
- American Academy of Pediatrics Subcommittee on Diagnosis and Management of Bronchiolitis: Diagnosis and management of bronchiolitis. Pediatrics 118, 2006, 1774-1793. A current statement on the management of bronchiolitis: what works and what is not recommended.
- American Academy of Pediatrics: respiratory syncytial virus. In Red Book: 2009 Report of the Committee on Infectious Diseases. Edited by Pickering LK, Baker CJ, Kimberlin DW, Long SS. Elk Grove Village, IL. American Academy of Pediatrics; 2009; 560–568.
- ANDRIES, K., MOEREMANS, M., GEVERS, T., et al., Substituted enzimidazoles with nanomolar activity against respiratory syncytial virus. Antiviral Res 60, 2003, 209-219.
- ANSEJO, A., RODRÍGUEZ, L., VILLANUEVA, N. Determination of phosphorylated residues from human respiratory syncytial virus P protein that are dynamically dephosphorylated by cellular phosphatases: a possible role for serine 54, Journal of General Virology, 86, 2005, 1109–1120.
- ASENJO, A., CALVO, E., VILLANUEVA, N. Phosphorylation of human respiratory syncytial virus P protein at threonine 108 controls its interaction with the M2-1 protein in the viral RNA polymerase complex, J. Gen. Virol., 87, 2006, 3637–3642.
- BARIK, S., MCLEAN, T., DUPY, L.C. Phosphorylation of Ser232 Directly Regulates the Transcriptional Activity of the P Protein of Human Respiratory Syncytial Virus: Phosphorylation of Ser237 May Play an Accessory Role., Virol., 213, 1995, 405–412.
- BEHLING, E.B., SENDÃO, M.C., FRANCESCATO, H.D.C., ANTUNES, L.M.G., BIANCHI, M.L.P. FLAVONÓIDE QUERCETINA: ASPECTOS GERAIS E AÇÕES BIOLÓGICAS. Alim. Nutr., 5, 2004, 285-292.
- BERENDSEN, H.J.C., VAN DER SPOEL, D., VAN DRUNEN, R. GROMACS: A message-passing parallel molecular dynamics implementation. Computer Physics Communications, 91, 1995, 43-56.

BERMINGHAM, A., COLLINS, P.L. The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription Proc. Natl. Acad. Sci. USA Vol. 96, 1999, 11259–11264.

BLONDOT, M.L., DUBOSCLARD, V., FIX, J., LASSOUED, S., AUMONT-NICAISE, M., BONTEMS, F., SIZUN, C. Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein. PLoS Pathogens, 8, 2012, 1002734.

BLOUNT, R.E., MORRIS, J.A., SAVAGE, R.E. Recovery of cytopathogenic agent from chimpanzees with coryza. Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.), 92, 1956, 544–549.

BLUMBERG, B.M., GIORGI, C., KOLAKOSKY, D. N protein of vesicular stomatitis virus selectively encapsidates leader RNA in vitro. Cell, 32, 1983, 559–567.

BONFANTI, J.F., MEYER, C., DOUBLET, F., FORTIN, J., MULLER, P., QUEGUINER, L., GEVERS, T., JANSSENS, P., SZEL, H., WILLEBRORDS, R., TIMMERMAN, P., WUYTS, K., VAN REMOORTELE, P., JANSSENS, F., WIGERINCK, P., ANDRIES, K. Selection of a respiratory syncytial virus fusion inhibitor clinical candidate. 2. Discovery of a morpholinopropylaminobenzimidazole derivative (TMC353121). J Med Chem. 51, 2008, 875–896.

BOSSERT, B., CONZELMANN, K.K. Respiratory syncytial virus (RSV) nonstructural (NS) proteins as host range determinants: a chimeric bovine RSV with NS genes from human RSV is attenuated in interferon-competent bovine cells., J. Virol., 76, 2002, 4287–4293.

BOSSERT, B., MAROZIN, S., CONZELMANN, K.K. Nonstructural proteins NS1 and NS2 of bovine respiratory syncytial virus block activation of interferon regulatory factor 3., J. Virol., 77 (2003) 8661–8668.

BUTLER, M.S., Natural products to drugs: natural product derived compounds in clinical trials, Nat. Prod. Rep., 22, 2005, 162–195.

CARTEE, T.L., WERTZ, G.W. Respiratory Syncytial Virus M2-1 Protein Requires Phosphorylation for Efficient Function and Binds Viral RNA during Infection, J. Virol., 75, 2001, 12188–12197.

CARTER, S.D., DENT, K.C., ATKINS, E., et al., Direct visualization of the small hydrophobic protein of human respiratory syncytial virus reveals the structural basis for membrane permeability, FEBS Letters, 584, 2010, 2786–2790.

CASTAGNE, N., BARBIER, A., BERNARD, J., et al., Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain., J. Gen. Virol., 85, 2004, 1643–1653.

CAZAROLLI, L.H., ZANATTA, L., ALBERTON, E.H., et al., Flavonoids: prospective drug candidates. Med. Chem., 8, 2008, 1429–1440.

CHALLA, S., SCOTT, A.D., YUZHAKOV, O. ZHOU, Y., YIONG-YIP, C.L., GAO, N., THRESHER, J., YU, Q. Mechanism of Action for Respiratory Syncytial Virus Inhibitor RSV604. *Antimicrobial Agents and Chemotherapy*, 59, 2015, 1080–1087.

CHANOCK, R., FINBERG, L. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). II. Epidemiologic aspects of infection in infants and young children., *Am. J. Hyg.*, 66, 1957b, 291–300.

CHANOCK, R., ROIZMAN, B., MYERS, R. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. Isolation, properties and characterization. *Am. J. Hyg.*, 66, 1957, 281–290.

CHAPMAN, J., ABBOTT, E., ALBER, D.G., et al., RSV604, a novel inhibitor of respiratory syncytial virus replication. *Antimicrob Agents Chemother*, 51, 2007, 3346-3353.

CHENG, X., PARK, H., ZHOU, H., et al., Overexpression of the M2-2 Protein of Respiratory Syncytial Virus Inhibits Viral Replication, *J. Virol.*, 79, 2005, 13943–13952.

CIANCI, C., MEANWELL, N., KRYSTAL, M. Antiviral activity and molecular mechanism of an orally active respiratory syncytial virus fusion inhibitor. *J Antimicrob Chemother.* 55, 2005, 289–292.

COLLINS, P. L., MELERO, J. A. Progress in understanding and controlling respiratory syncytial virus : Still crazy after all these years. *Virus Research*, 162, 2011, 80–99.

COLLINS, P.L., HILL, M.G., CRISTINA, J., et al., Transcription elongation factor of respiratory syncytial virus, a nonsegmented negativestrand RNA virus. *PNAS* 93, 1996, 81–85.

COLLINS, P.L., HILL, M.G., CRISTINA, J., GROSFELD, H. Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proc Natl Acad Sci U S A* 93, 1996, 81–85.

CORNELI, H.M., ZORC, J.J., MAHAJAN, P., et al., A multicenter, randomized, controlled trial of dexamethasone for bronchiolitis. *N Engl J Med*, 357, 2007, 331-339.

CROTTY, S., ANDINO, R., Implications of high RNA virus mutation rates: lethal mutagenesis and the antiviral drug ribavirin. *Microbes Infect.* 4, 2002a, 1301-1307.

CROTTY, S., CAMERON, C., ANDINO, R. Ribavirin's antiviral mechanism of action: lethal mutagenesis? *J Mol Med* 80, 2002b, 86-95.

CURRAN, J., BOECK, R., LIN-MARQ, N., et al., Paramyxovirus Phosphoproteins Form Homotrimers as Determined by an Epitope Dilution Assay, via Predicted Coiled Coils, *Virology*, 214, 1995, 139-149.

DOBSON, J.V., STEPHENS-GROFF, S.M., MCMAHON, S.R., et al., The use of albuterol in hospitalized infants with bronchiolitis. *Pediatrics* 101,1998, 1998, 361-368.

DONGSHENG L, JANS DA, BARDIN PG, et al., Association of Respiratory Syncytial Virus M Protein with Viral Nucleocapsids Is Mediated by the M2-1 Protein. *J Virol*, 82, 2008, 8863–8870.

DOUGLAS, J.L. In search of a small-molecule inhibitor for respiratory syncytial virus. *Expert Rev Anti Infect Ther* 2, 2004, 625-639.

DOUGLAS, J.L., PANIS, M.L., HO, E., LIN, K.Y., KRAWCZYK, S.H., GRANT, D.M., CAI, R., SWAMINATHAN, S., CHEN, X., CIHLAR, T. Small molecules VP-14637 and JNJ-2408068 inhibit respiratory syncytial virus fusion by similar mechanisms. *Antimicrob Agents Chemother.* 49, 2005, 2460–2466.

DUBOSCLARD, V., BLONDOT, M.L., ELÉOUËT, J.F., et al., ¹H, ¹³C, and ¹⁵N resonance assignment of the central domain of hRSV transcription antitermination factor M2-1, *Biomol NMR Assign*, 5, 2011, 237–239.

ESPERANTE, S.A, CHEMES, L.B., SÁNCHEZ, I.E., DE PRAT-GAY, G. The respiratory syncytial virus transcription antiterminator M(2-1) is a highly stable, zinc binding tetramer with strong pH-dependent dissociation and a monomeric unfolding intermediate. *Biochemistry*, 50, 2011, 8529–39.

ESPERANTE, S.A, NOVAL, M.G., ALTIERI, T.A, DE OLIVEIRA, G.A.P., SILVA, J.L., DE PRAT-GAY, G. Fine modulation of the respiratory syncytial virus m2-1 protein quaternary structure by reversible zinc removal from its cys3-his1 motif. *Biochemistry*, 52, 2013, 6779–89.

FEARNS, R., COLLINS, P.L. Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J Virol* 73, 1999, 5852–5864.

FEHER, M., SCHMIDT, J.M. Fuzzy clustering as a means of selecting representative conformers and molecular alignments, *J. Chem. Inf. Comput. Sci.*, 43, 2003, 218–227.

FENTON, C., SCOTT, L.J., PLOSKER, G.L. Palivizumab: a review of its use as prophylaxis for serious respiratory syncytial virus infection., *Paediatr Drugs*, 6, 2004, 177–197.

FLORES, G., HORWITZ, R.I. Efficacy of beta2-agonists in bronchiolitis: a reappraisal and meta-analysis., *Pediatrics*, 100, 1997, 233–239.

GAGLIARDI, T.B., PAULA, F.E., IWAMOTO, M.A., PROENÇA-MODENA, J.L., SANTOS, A.E., CAMARA, A.A., CERVI, M.C., CINTRA, O.A.L, ARRUDA, E. Concurrent Detection of Other Respiratory Viruses in Children Shedding Viable Human Respiratory Syncytial Virus, *Journal of Medical Virology*, 2013, 85, 1852–1859.

GASTEIGER E., HOOGLAND C., GATTIKER A., DUVAUD S., WILKINS M.R., APPEL R.D., BAIROCH A.; Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press, 2005, 571–607

GHILDYAL R, LI D, PEROULIS I, et al., Interaction between the respiratory syncytial virus G glycoprotein cytoplasmic domain and the matrix protein., *J. Gen. Virol.*, 86, 2005, 1879–1884.

GHILDYAL R, MILLS J, MURRAY M, et al., Respiratory syncytial virus matrix protein associates with nucleocapsids in infected cells., *J. Gen. Virol.*, 83, 2002, 753–757.

GLEZEN WP, TABER LH, FRANK AL, et al., Risk of primary infection and reinfection with respiratory syncytial virus., *Am. J. Dis. Child.*, 140, 1986, 543–546.

GUERRA, M.E.L. Prospecção de Novos Peptídeos com Ação Leishmanicida por Dinâmica Molecular, 2014, 93 f. Dissertação (Mestrado em Biofísica Molecular), IBILCE/UNESP.

HALL CB, DOUGLAS RG JR., GEIMAN JM, Possible transmission by fomites of respiratory syncytial virus., *J. Infect. Dis.* 141, 1980, 98–102.

HALL CB, DOUGLAS RG JR., GEIMAN JM, Quantitative shedding patterns of respiratory syncytial virus in infants., *J. Infect. Dis.* 132, 1975, 151–156.

HALL CB, DOUGLAS RG JR., SCHNABEL KC, et al., Infectivity of respiratory syncytial virus by various routes of inoculation. *Infect. Immune.*, 33, 1981b, 779–783.

HALL CB, LONG CE, SCHNABEL KC, Respiratory syncytial virus infections in previously healthy working adults. *Clin. Infect. Dis.* 33, 2001, 792–796.

HALL CB, WALSH EE, LONG CE, et al., Immunity to and frequency of reinfection with respiratory syncytial virus., *J. Infect. Dis.*, 163, 1991, 693–698.

HALL T.A., BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 1999, 95-98.

HALLCB, DOUGLAS RG JR., Modes of transmission of respiratory syncytial virus. *J Pediatr* 99, 1981, 100–103.

HANSSON T, OOSTENBRINK C, VAN GUNSTEREN W. Molecular dynamics simulations. *Curr Opin Struct Biol* 12, 2002,190–196.

HARDY RW, WERTZ GW, The product of the respiratory syncytial virus M2 gene ORF1 enhances readthrough of intergenic junctions during viral transcription. *J. Virol.* 72, 1998, 520–526.

HARDY RW, WERTZ GW. The Cys(3)-His(1) motif of the respiratory syncytial virus M2-1 protein is essential for protein function. *J Virol* 74, 2000, 5880–5885.

HARRISON, M. S., SAKAGUCHI, T., & SCHMITT, A. P. Paramyxovirus assembly and budding: building particles that transmit infections. *The International Journal of Biochemistry & Cell Biology*, 42, 2010, 1416–29.

HARVEY A.L.Natural products in drug discovery *Drug Discovery Today* 13, 2008, 894–901.

HENDERSON EA, ALBER DG, BAXTER RC, et al., 1,4-benzodiazepines as inhibitors of respiratory syncytial virus. The identification of a clinical candidate., *J. Med. Chem.* 50, 2007, 1685-1692.

HENDERSON FW, COLLIER AM, CLYDE WA JR., et al., Respiratory-syncytial-virus infections, reinfections and immunity. A prospective, longitudinal study in young children., *N. Engl. J. Med.*, 300,1979, 530–534.

JIN H, ZHOU H, CHENG X, et al., Recombinant respiratory syncytial viruses with deletions in the NS1, NS2, SH, and M2-2 genes are attenuated in vitro and in vivo., *Virol.*, 273, 2000, 210–218.

KAUL, T. N.; MIDDLETON JR, E.; OGRA, P. L. Antiviral Effect of Flavonoids on Human Viruses. *Journal of Medical Virology*, 15, 1985, 71-79.

KIMURA K, MORI S, TOMITA K, et al., Antiviral activity of NMSO3 against respiratory syncytial virus infection in vitro and in vivo., *Antiviral Res.*, 47, 2000, 41-51.

KOCH MA, SCHUFFENHAUER A, SCHECK M, et al., Charting biologically relevant chemical space: A structural classification of natural products (SCONP), *PNAS*, 102, 2005, 17272–17277.

KRISSINEL E, HENRICK K, Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372, 2007, 774–797.

KUO L, GROSFELD H, CRISTINA J, HILL MG, COLLINS PL. Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. *J Virol* 70, 1996, 6892–6901

KUO, L., FEARN, R., COLLINS, P. Analysis of the gene start and gene end signals of human respiratory syncytial virus: quasi-template initiation at position 1 of the encoded mRNA. *J Virol*, 71, 1997, 4944-4953.

LAKOWICZ, J. R. Principles of Fluorescence Spectroscopy. New York: Kluwer Academic Publishers/Plenum Press, 2nd ed., 1999.

LAMB RA, PARKS GD, Paramyxoviridae: the viruses and their replication. In: Knipe DM, Howley PM, editors. *Fields virology*. Philadelphia: Lippincott, Williams and Wilkins; 2006. p. 1449–96.

LEACH A.R. Molecular Modelling: Principles and Applications. [S.I.]: Longman, 1996.

LI D, JANS DA, BARDIN PG, et al., Association of Respiratory Syncytial Virus M Protein with Viral Nucleocapsids Is Mediated by the M2-1 Protein *J Virol.*, 82, 2008, 8863–8870.

LINDAHL, E., HESS, B., VAN DER SPOEL, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Molecular modeling annual*. 7, 2001, 306-317.

LO MS, BRAZAS RM, HOLTZMAN MJ. Respiratory syncytial virus nonstructural proteins NS1 and NS2 mediate inhibition of Stat2 expression and alpha/beta interferon responsiveness., *J. Virol.*, 79, 2005, 9315–9319.

LUTTICK A, LIN B, MORTON C, et al., Characterization of a new class of polycyclic RSV inhibitors. *Antiviral Res.*, 74, 2007, A66.

MARQUARDT ED, Cost of ribavirin therapy for respiratory syncytial virus infection. *J Pediatr.*, 126, 1995, 847.

MARTY, F.M, CHEMALY, R.F., LIAKOPOULOU, E., DENT, J.C., POWELL, K. A double-blind, randomised, placebo-controlled study to evaluate the safety and efficacy of RSV604 in adults with respiratory syncytial virus (RSV) infection following stem cell transplantation. IX International Symposium on Respiratory Viral Infections, Hong Kong, China, 2007.

MASON SW, ABERG E, LAWETZ C, et al., Interaction between Human Respiratory Syncytial Virus (RSV) M2-1 and P Proteins Is Required for Reconstitution of M2-1-Dependent RSV Minigenome Activity, *J. Virol.*, 77, 2003, 10670–10676

MASON, S. W., ABERG, E., LAWETZ, C., DELONG, R., WHITEHEAD, P., & LIUZZI, M. Interaction between Human Respiratory Syncytial Virus (RSV) M2-1 and P Proteins Is Required for Reconstitution of M2-1-Dependent RSV Minigenome Activity. *Society*, 77, 2003, 10670–10676.

MERIC C, SPEHNER D, MAZARIN V, Respiratory syncytial virus nucleocapsid protein (N) expressed in insect cells forms nucleocapsid-like structures., *Virus. Res.*, 31, 1994, 187-201.

MORRIS, G. M., HUEY, R., LINDSTROM, W., SANNER, M. F., BELEW, R. K., GOODSSELL, D. S. AND OLSON, A. J. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Computational Chemistry*, 16, 2009, 2785-91.

MOURA, F. E. A, PERDIGÃO, A. C. B., RIBEIRO, J. F., FLORÊNCIO, C. M. G. D., OLIVEIRA, F. M. S., PEREIRA, S. A R., ... DURIGON, E. L. Respiratory syncytial virus epidemic periods in an equatorial city of Brazil. *Influenza and Other Respiratory Viruses*, 7, 2013, 1128–35.

Othumpangat S, Gibson L, Samsell L, et al., NGF is an essential survival factor for bronchial epithelial cells during respiratory syncytial virus infection., *PLoS ONE*, 4, 2009, e6444.

PATEL, J. M. A Review of Potential Health Benefits of Flavonoids. *Lethbridge Undergraduate Research Journal*.3, 2008.

PELLETIER AJ, MANSBACH JM, CAMARGO CA Jr., Direct medical costs of bronchiolitis hospitalizations in the United States., *Pediatrics*, 118, 2006, 2418-2423.

PETERS TR, CROWE JE JR., Respiratory syncytial virus. In: Long SS, editor. *Principles and practice of pediatric infectious diseases*, 3rd edition. Philadelphia, PA: Churchill Livingstone; 2008. 1112–1117.

PETTERSEN EF, GODDARD TD, HUANG CC, COUCH GS, GREENBLATT DM, MENG EC, FERRIN TE. UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem*. 25, 2004, 1605-12.

RAMASWAMY M, SHI L, MONICK MM, et al., Specific inhibition of type I interferon signal transduction by respiratory syncytial virus., *Am. J. Respir. Cell Mol. Biol.*, 30, 2004, 893–900.

RAMASWAMY, SHI L, VARGA SM, et al., Respiratory syncytial virus nonstructural protein 2 specifically inhibits type I interferon signal transduction. *Viol.* 344, 2005 328-339.

ROY, A., KUCUKURAL, A., ZHANG, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5, 2010, 725-738.

SANNER, M.F., *Python: A Programming Language for Software Integration and Development*. *J. Mol. Graphics Mod.*, 17, 1999, 57-61.

SANTOS NSO, ROMANOS MTRV, WIGG MD, *Introdução à Virologia Humana*, 2ª edição, Rio de Janeiro, Guanabara Koogan, 2008.

SCHLENDER J, BOSSERT B, BUCHHOLZ U, et al., Bovine respiratory syncytial virus nonstructural proteins NS1 and NS2 cooperatively antagonize alpha/beta interferon-induced antiviral response. *J Virol.*, 74, 2000, 8234–8242.

SELWAY JTW, *Antiviral Activity of Flavones and Flavans*, V. Cody, E. Middleton, J.B. Harborne, Editors, *Plant Flavonoids in Biology and Medicine: Biochemical, Pharmacological and Structure-Activity Relationships*, Alan R. Liss, Inc., New York, London , 1986, 521–536.

SHAY DK, HOLMAN RC, NEWMAN RD, et al., Bronchiolitis-associated hospitalizations among US children, 1980-1996. *JAMA.*, 282, 1999, 1440-1446.

SIDWELL RW, BARNARD DL, Respiratory syncytial virus infections: recent prospects for control. *Antiviral Res*, 71, 2006, 379-390.

SIMÕES CMO, SCHENKEL EP, GOSMANN G, et al., *Farmacognosia - da Planta ao Medicamento*, 5ª ed., Editora da UFSC: Santa Catarina, 2004.

SINNOKROT, M.O., SHERRILL, D. Unexpected Substituent effects in face-to-face π -stacking interactions. *J of Physical Chemistry A*. 107, 2003, 8377-8379.

SPANN KM, SCHENKEL EP, GOSMANN G, et al., Suppression of the induction of alpha, beta, and lambda interferons by the NS1 and NS2 proteins of human respiratory syncytial virus in human epithelial cells and macrophages., *J. Virol.* 78, 2004, 4363–4369.

SPANN KM, TRAN KC, COLLINS PL, Effects of Nonstructural Proteins NS1 and NS2 of Human Respiratory Syncytial Virus on Interferon Regulatory Factor 3, NF-kappaB, and Proinflammatory Cytokines, *J. Virol.*, 79, 2005, 5353–5362.

STOKES HL, EASTON AJ, MARRIOTT AC, Chimeric pneumovirus nucleocapsid (N) proteins allow identification of amino acids essential for the function of the respiratory syncytial virus N protein., *J. Gen. Virol.*, 84, 2003, 2679–2683.

TABOADA, P. BARBOSA, S., CASTRO, E. GUTIÉRREZ-PICHEL, M., MOSQUERA, V. Effect of solvation on the structure of human serum albumin in aqueous-alcohol mixed solvents. *Chemical Physics*, 340, 2007, 59-68.

TANG RS, NGUYEN N, CHENG X, JIN H. Requirement of cysteines and length of the human respiratory syncytial virus M2-1 protein for protein function and virus viability. *J Virol* 75, 2001, 11328–11335.

TANNER, S. J., ARIZA, A., RICHARD, C.-A., KYLE, H. F., DODS, R. L., BLONDOT, M.-L., ... BARR, J.N. Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 2014, 1580–5.

TRAN T-L, CASTAGNE N, BHELLA D, et al., The nine C-terminal amino acids of the respiratory syncytial virus protein P are necessary and sufficient for binding to ribonucleoprotein complexes in which six ribonucleotides are contacted per N protein protomer., *J. Gen. Virol.*, 88, 2007, 196–206.

TRAN T-L, CASTAGNÉ N, DUBOSCLARD V, et al., The Respiratory Syncytial Virus M2-1 Protein Forms Tetramers and Interacts with RNA and P in a Competitive Manner, *J. Virol.*, 83, 2009, 6363–6374.

TRAN, T.-L., CASTAGNÉ, N., DUBOSCLARD, V., NOINVILLE, S., KOCH, E., MOUDJOU, M., ... ELÉOUËT, J.-F. The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. *Journal of Virology*, 83, 2009, 6363–74.

VISON JA, Flavonoids in foods as in vitro and in vivo antioxidants. *Flavonoids in the Living System*. Mantley and Buslig Plenum Press, New York, 1998.

WALSH EE, PETERSON DR, FALSEY AR, Risk factors for severe respiratory syncytial virus infection in elderly persons. *J Infect Dis*, 189, 2004, 233–238.

WETZEL S, SCHUFFENHAUER A, ROGGO S, et al., *Chimia* 61, 2007, 355–360.

WHITEHEAD SS, BUKREYEV A, TENG MN, et al., Recombinant respiratory syncytial virus bearing a deletion of either the NS2 or SH gene is attenuated in chimpanzees., *J. Virol.*, 73, 1999, 3438–3442.

WU H, PFARR DS, TANG Y, et al., Ultra-potent antibodies against respiratory syncytial virus: effects of binding kinetics and binding valence on viral neutralization. *J Mol Biol.*, 350, 2005, 126-144.

WYDE PR, CHETTY SN, TIMMERMAN P, et al., Short duration aerosols of JNJ 2408068 (R1170591) administered prophylactically or therapeutically protect cotton rats from experimental respiratory syncytial virus infection. *Antiviral Res*, 60, 2003, 221-231.

YANG J., YAN R., ROY, A., XU, D., POISSON, J., ZHANG, Y. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12, 2015, 7-8.

ZHANG, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9, 2008, 40.

ZHOU H, CHENG X, JIN H. Identification of amino acids that are critical to the processivity function of respiratory syncytial virus M2-1 protein. *J Virol* 77, 2003, 5046–5053

ZSILA, F., BIKÁDI, Z., SIMONYI, M., et al. Probing the binding of the flavonoid quercetin to human serum albumin by circular dichroism electronic absorption spectroscopy and molecular modeling methods. *Biochemical pharmacology*, 65, 2003, 447-56.

GILLET LC, NAVARRO P, TATE S, RÖST H, SELEVSEK N, REITER L, BONNER R, AEBERSOLD R. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* 11, 2012, O111.016717–O111.016717.

YANG J, YAN R, ROY A, XU D, POISSON J, ZHANG Y. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 2015, 12, 7-8.

SCHNEIDMAN-DUHOVNY D, INBAR Y, NUSSINOV R, WOLFSON HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *NAR*, 2005, 33, W363-W367.

Anexo I

Clonagem de produto de PCR usando o protocolo TOPO® XL Gel Purification Kit (Invitrogen®).

Purificação de produto de PCR

Gel

As instruções são para preparo de gel de agarose 0,8%.

- Misturar 0,4g de agarose em 50mL de tampão TAE 1X
 - Dissolver a agarose no microondas;
 - Após diminuir a temperatura da solução, acrescentar 20-40µL de solução de Cristal Violeta (2mg/mL). O meio deve ficar com uma coloração roxa;
 - Colocar a solução na forma de gel com o pente e deixar solidificar;
 - Colocar o gel na cuba de eletroforese e cobrir com tampão TAE 1X sem corante.
-
- Adicionar 8 µL de Cristal Violeta 6X pra cada 40 µL do produto da PCR e aplicar no gel;
 - Correr o gel à 80 volts e deixar correr até que o material da PCR tenha entrado no gel (um pequena banda azulada);
 - Usando uma lâmina, remova a banda de DNA do gel e transferir para um tubo tipo eppendorf de 2mL;
 - Estimar o volume do pedaço de agarose (1mg ~1µL);
 - Adicionar 2,5 vezes o volume de iodeto de sódio 6,6M e agitar bem o tubo;
 - Incubar à 42°C até que o gel tenha dissolvido;
 - Adicionar 1,5 vezes o volume de Tampão de Ligação (Binding Buffer) e homogeneizar;
 - Montar a coluna de purificação e adicionar a solução com a amostra na coluna;
 - Centrifugar de 2.000 a 3.000 xg em uma microcentífuga por 30 segundos;
 - Colocar o líquido filtrado novamente na coluna e repetir a centrifugação.
 - Repetir o paço anterior mais uma vez;
 - Descartar o líquido filtrado;
 - Adicionar 400 µL de lavagem final 1X (1X Final Wash) e centrifugar de 2.000 a 3.000 xg em uma microcentífuga por 30 segundos;
 - Repetir o paço anterior e descartar o líquido filtrado (800 µL);
 - Centrifugar novamente a coluna em velocidade máxima (>10.000 xg) por pelo menos 1 minuto para secar a coluna;
 - Transferir a coluna para um novo tubo tipo eppendorf de 1,5mL;

- Adicionar 20 µL tampão TE diretamente na coluna e incubar por 1 minuto à temperatura ambiente;
- Centrifugar novamente a coluna em velocidade máxima ($>10.000 \times g$) por pelo menos 1 minuto para eluir o DNA no tubo.

Clonando no TOPO-XL-PCR

- Adicionar 4 µL do PCR purificado no gel em 1 µL do vetor pCR-XL-TOPO;
- Homogeneizar gentilmente e incubar por 5 minutos em temperatura ambiente ($\sim 25^{\circ}\text{C}$);
- Adicionar 1 µL de solução de parada (6X TOPO cloning Stop Solution) e misturar a solução;

Transformação

- Adicionar 2 µL da reação de clonagem em um tubo com células quimicamente competente One Shot® e homogeneizar gentilmente. Não pipetar e evitar a formação de bolhas;
- Incubar por 30 minutos no gelo;
- Dar um choque térmico por 30 segundos à 42°C ;
- Colocar novamente no gelo por 2 minutos;
- Adicionar 250 µL de meio SOC;
- Colocar no shaker à 37°C por 1 hora;
- Aplicar 50-150 µL de célula transformante em meio de cultura sólido acrescido de canamicina (50 µL/mL).
- Incubar overnight à 37°C

Anexo II

Protocolo de Reação de Minipreparação (QIAprep® Spin Miniprep Kit (Qiagen®))

- Pré-incubação

1. Repicar uma colônia bacteriana em um tubo tipo Falcon com 1 a 5ml de meio LB líquido com antibiótico apropriado.
2. Deixar as colônias crescerem por 16 horas no shaker a 37°C.
3. Centrifugar por 20 minutos a 4°C a 35.000rpm.

- Reação de Miniprep usando a centrífuga (protocolo Qiagen®)

1. Suspender o pellet de células bacterianas em 250µl de Tampão P1 e transferir para um eppendorf.
2. Adicionar 250µl de Tampão P2 e agitar por inversão de 4 a 6 vezes.
3. Adicionar 350µl de Tampão N3 e misturar imediatamente por inversão de 4 a 6 vezes.
4. Centrifugar por 10 minutos a 13.000 rpm (~17.900g) em uma microcentrífuga.
5. Pipetar o sobrenadante e aplicar na coluna QIAprep.
6. Centrifugar por 1 minuto.
7. Descartar o líquido filtrado.
8. Lavar a coluna com 750µl de Tampão PE e centrifugar por 1 minuto.
9. Descartar o líquido filtrado.
10. Centrifugar por 1 minuto para remover tampão residual.
11. Transferir a coluna QIAprep para um eppendorf de 1,5ml.
12. Adicionar 30µl de Tampão EB ao centro da coluna QIAprep e deixar em repouso por 1 minuto.
13. Centrifugar por 1 minuto.

Anexo III

Protocolo para extração de DNA de gel (QIAquick®)

1. Excisar o fragmento de DNA do gel de agarose.
2. Pesar o pedaço de gel em um tubo eppendorf.
3. Para 1 volume de gel, adicionar 3X o volume do Tampão QG (100mg~100µl) (Para gel de agarose maior que 2%, adicionar 6X de QG).
4. Incubar a 50°C por 10 minutos (ou até o gel dissolver).
5. Após o gel dissolver, observar se a cor da mistura está amarela (similar ao tampão QG). (Se a cor da mistura estiver laranja ou violeta, adicionar 10µl de acetato de sódio 3M, pH 5 e misturar. A cor da mistura ficará amarela)
6. Adicionar 1 volume do gel de isopropanol à amostra e misturar.
7. Encaixar uma coluna QIAquick em um tubo coletor de 2ml.
8. Aplicar a amostra na coluna QIAquick e centrifugar por 1 minuto a 13.000rpm.
9. Descartar o filtrado.
10. Adicionar 750µl de Tampão PE à coluna QIAquick e centrifugar por 1 minuto.
11. Descartar o filtrado.
12. Centrifugar mais uma vez a coluna QIAquick por 1 minuto a 13.000rpm para remover tampão residual.
13. Transferir a coluna QIAquick para um eppendorf de 1,5ml.
14. Adicionar 30µl de Tampão EB ao centro da coluna QIAquick e deixar incubado por 4 minutos.
15. Centrifugar por 1 minuto.
16. Para a análise em gel de agarose, usar 1 volume de tampão de corrida para 5 volumes de DNA purificado.

Anexo IV

Parâmetros da dinâmica molecular.

```
title           = M21 MD
; Run parameters
integrator      = md           ; leap-frog integrator
nsteps         = 50000000      ; 2 * 50000000 = 100000 ps, 100 ns
dt             = 0.002         ; 2 fs
; Output control
nstxout        = 1000         ; save coordinates every 2 ps
nstvout        = 1000         ; save velocities every 2 ps
nstxtcout      = 1000         ; xtc compressed trajectory output every 2 ps
nstenergy      = 1000         ; save energies every 2 ps
nstlog         = 1000         ; update log file every 2 ps
; Bond parameters
continuation    = yes         ; Restarting after NPT
constraint_algorithm = lincs   ; holonomic constraints
constraints     = all-bonds    ; all bonds (even heavy atom-H bonds) constrained
lincs_iter      = 1           ; accuracy of LINCS
lincs_order     = 2           ; also related to accuracy
; Neighborsearching
ns_type         = grid        ; search neighboring grid cells
nstlist        = 5           ; 10 fs
rlist          = 1.4         ; short-range neighborlist cutoff (in nm)
rcoulomb       = 1.4         ; short-range electrostatic cutoff (in nm)
rvdw           = 1.4         ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype     = PME         ; Particle Mesh Ewald for long-range electrostatics
pme_order       = 4           ; cubic interpolation
fourierspacing = 0.16        ; grid spacing for FFT
; Temperature coupling is on
tcoupl         = V-rescale    ; modified Berendsen thermostat
tc-grps        = Protein Non-Protein; two coupling groups - more accurate
tau_t          = 0.1 0.1     ; time constant, in ps
ref_t          = 310 310     ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl         = Berendsen    ; Pressure coupling on in NPT
pcoupltype     = isotropic    ; uniform scaling of box vectors
tau_p          = 2.0         ; time constant, in ps
ref_p          = 1.0         ; reference pressure, in bar
compressibility = 4.5e-5      ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions
pbc            = xyz          ; 3-D PBC
; Dispersion correction
DispCorr       = EnerPres    ; account for cut-off vdW scheme
; Velocity generation
gen_vel        = no          ; Velocity generation is off
```

APENDICE

Biophysical characterization of the interaction between M2-1 protein of Human Respiratory Syncytial Virus (hRSV) and quercetin.

Thiago Salem Pançonato Teixeira^{1,3} (thiago_spt@yahoo.com.br), Marcelo Andrés Fossey^{2,3} (mafossey@gmail.com), Luis Octávio Regasini⁴ (luisregasini@gmail.com), Bruno Rafael Pereira Lopes⁵, Karina Alves de Toledo⁵ (karinafcla@gmail.com), Fátima Pereira de Souza^{2,3} (fatyssouza@gmail.com) *.

¹ – Instituto de Biociências, Letras e Ciências Exatas, UNESP Univ Estadual Paulista, Campus São José do Rio Preto, Departamento de Biologia. ² - Instituto de Biociências, Letras e Ciências Exatas, UNESP Univ Estadual Paulista, Campus São José do Rio Preto, Departamento de Física. ³ - Instituto de Biociências, Letras e Ciências Exatas, UNESP Univ Estadual Paulista, Campus São José do Rio Preto, Centro Multiusuário de Inovação Biomolecular, Laboratório de Biologia Molecular. ⁴ - Instituto de Biociências, Letras e Ciências Exatas, UNESP Univ Estadual Paulista, Campus São José do Rio Preto, Departamento de Química e Ciências Ambientais; ⁵ – Faculdade de Ciências e Letras, UNESP Univ Estadual Paulista, Campus Assis, Departamento de Ciências Biológicas, Laboratório de Imunologia Celular e Molecular.

* Rua Cristóvão Colombo, 2265, Jardim Nazareth, São José do Rio Preto – SP, 15054-000, Brazil, tel: +55 18 3221-2463.

Abstract

Human Respiratory Syncytial Virus (hRSV) is the major causative agent of acute respiratory infections. Among its eleven proteins, M2-1 is a transcription antiterminator, making it an interesting target for antivirals. Quercetin is a flavonol that has been shown to inhibit some virus infectivity and replication. The M2-1 gene was cloned, expressed and the protein was purified with affinity and molecular exclusion chromatography. Thermal stability and secondary structure were analyzed by circular dichroism and, the interaction with Quercetin was analyzed by fluorescence spectroscopy. To understand the interaction mechanism was performed molecular docking. The purified protein is mainly composed by α -helix, being the melting temperature of 328.6 K. The titration of M2-1 with Quercetin showed it interacts in two sites, one with a strong constant association (site 1 = $\sim 1.5 \times 10^6 \text{ M}^{-1}$) and another with a weak one (site 2 = $1.1 \times 10^5 \text{ M}^{-1}$). This interaction is not cooperative, and the interaction of Quercetin at site 1 is through ionic interaction and at site 2 is through hydrophobic interaction. The docking of the ligand shows that it interacts on the N-terminus face in a more ionic pocket and, between the oligomerization and RNA and P protein interaction domain, in a more hydrophobic pocket, as predicted by experimental data. So, we postulated that this ligand could be acting as an inhibitor of the protein avoiding virus replication and budding.

Key-words: hRSV, M2-1, Quercetin, Biophysical Analyses, Molecular docking.

Introduction

Among the viral agents responsible for acute lower respiratory infection (ALRI) in infants and children, Human Respiratory Syncytial Virus is the leading cause [1, 2]. Each year, hRSV is responsible for over 30 million of ALRI cases with a result of 60.000-190.000 children deaths under five years old, elderly and immunocompromised [3]. It is estimated that 70% of infants under one year old are infected with the hRSV and 100% of two year old children have been infected at least once [4, 5].

RSV is classified in Mononegavirales order, *Paramyxoviridae* family, *Paramyxovirinae* subfamily, *Pneumovirus* genus. They are enveloped virus with helical capsid [6] and its genome is composed by a single strand RNA negative polarity, non-segmented of about 15 kb which encompasses 10 genes in the following order 3'-NS1-NS2-N-P-M-SH-G-F-M2-L-5', but translates eleven proteins because of an open reading frame at matrix-2 protein (M2) forming M2-1 and M2-2.

Four of the eleven proteins (N, P, L and M2-1) form the complex RNA-dependent RNA polymerase (RdRp), which is involved to genome replication and transcription. The M2-1 protein is a factor that prevents the premature transcription stop, hence, it is important to all mRNAs transcription [7]. Also, during the virus capsule formation, the N-terminus portion of M protein associates with M2-1 protein [8, 9] and with N and P protein [10] to interrupt the viral transcription and directs the virus assembly and budding.

Structurally, M2-1 protein is a tetramer [11] mainly composed by alpha helix [12, 13] and has four distinct regions: the first 32 residues in N-terminus links to zinc (zinc finger) and corroborates to transcription (is indispensable to its function, but the correct functions is unknown) [14, 15, 16]; the oligomerization domain situated between His33 and Gly62; a globular domain composed mainly by alpha helix (Ala63-Ser170) capable to link to P protein and RNA; and a C-terminus region with low organization degree [13].

The treatment for RSV infections is mainly symptomatic, including oxygen support and hydration. However, the mortality is high with annual cost to the health system [17, 18]. Despite the long term research with hRSV, there is no effective drug or available vaccine in the market. There are few drugs that lack specificity and are expensive, like ribavirin [19, 20, 21, 22, 23, 24].

Natural products are relevant source of antiviral agents [25, 26]. Flavonoid is a group of phenolic compounds widely diverse among natural products which have many activities as antitumor, antioxidant, antiviral and anti-inflammatory [27, 28]. Quercetin (3,3',4',5,7-pentahydroxyflavone) is an excellent free radical scavenger and a metal chelator and is one of the most potent bioflavonoids studied that possesses a variety of pharmacological activities including anti-inflammatory, antineoplastic, cardio protective and anticancer activities [29, 30].

Despite the description of quercetin having anti-cancer activity [31, 32, 33, 34] it is believed that this effect is due the anti-oxidant protective effect. On the other hand, quercetin has shown to have anti-viral activity. It was able to inhibit HBsAg and HBeAg secretion in HBV infected cells [35], HCV replication [36] HIV-1 reverse transcriptase [37], and transcriptase from avian myeloblastosis, Rous-associated virus-2 and Maloney murine leukemia virus [38].

Ahmed-Belkacem and co-authors (2014) [39] reported flavonoids, mainly quercetin, was able to inhibit approximately 80% of HCV RdRp activity. This effect was due to the binding of the flavonol to the entrance of RNA template tunnel. In 1985, Kaul et al. [40], observed that quercetin was able to inhibit RSV infection and replication. So probably, quercetin is binding to some proteins of the RdRp of RSV avoiding the replication or transcription.

As M2-1 is a target protein to viral replication, transcription and budding of hRSV, and because of the promising properties of quercetin, we investigated the interaction of both molecules by fluorescence spectroscopy and by bioinformatics tools.

Material and Methods.

Expression, purification and identification of M2-1 protein: A cDNA from Human Respiratory Syncytial Virus (hRSV) A2 strain M2-1 was inserted into pGEX4T-1 vector (GE healthcare) allowing expression of full-length M2-1 that was N-terminally fused to a GST affinity tag and a thrombin protease recognition sequence. Resulting plasmid was transformed into the cultures of *Escherichia coli* BL21 Gold (DE3) (Agilent Technologies), which were grown to OD₆₀₀ 0.6–0.8 at 37 °C and induced with 400 µM isopropyl β-D-1-thiogalactopyranoside (IPTG), with 50 µM zinc sulfate (ZnSO₄) to aid in the correct folding of the zinc finger, for 16 h at 28 °C. Cells were recovered by centrifugation and resuspended in 5 mL lysis buffer [50 mM Tris-Cl, pH 7.5, 1 M NaCl, 1x complete protease inhibitor mixture EDTA-free (Sigma Aldrich), 1 mM β-mercaptoethanol, 1 mM CaCl₂, 10 mM MgSO₄] per 100 ml of pellet culture. Cells were incubated on ice for 30 min, and then lysis was completed by sonication (15 cycles – 15 seconds on and 45 seconds off – potency of 36 W). Clarified supernatant containing soluble GST-M2-1 protein was applied in batch to glutathione Sepharose 4B resin (GE Healthcare) at 4°C for 4 hours with rotation. The resin was pelleted and washed with 4 volumes of lysis buffer and 5 volumes with PBS buffer. M2-1 was cleaved from GST on the resin overnight at 16°C by thrombin protease, which was expressed and purified in-house. Cleaved M2-1 protein was eluted by pelleting the resin and collecting the supernatant, which was applied over a 30 mL Sepharose 20/300 size exclusion chromatography (SEC) column (GE Healthcare) equilibrated with PBS buffer. The identity of the purified protein was confirmed by accurate mass determination by denaturing mass spectrometry. Protein was confirmed RNA-free by spectrophotometry (OD 260/280 with a ratio of 0.27). The protein concentration was determined spectrophotometrically using a molar extinction coefficient (ϵ_{280}) of 13200 M⁻¹cm⁻¹, calculated using the ExPASy ProtParam tool and in concordance with a previous report [13].

Primary protein structure sequencing: The primary sequence of the M2-1 protein was determined to see if it was expressed correctly. The protein was determined by the Shotgun method [41] by the group of Prof. Dr. Mario Sérgio Palma at Centro de Estudos de Insetos Sociais – UNESP.

Circular Dichroism (CD) Spectroscopy: GST-free M2-1 was analyzed with a Jasco J-810 spectropolarimeter at 25°C using 1 cm thick quartz cells. Protein samples were loaded into the cuvettes, and spectra between 200 and 260 nm were recorded in a velocity of 20 nm/min, with 2 seconds answer and slit opening of 1.0 nm, totalizing 10 accumulations. The baseline was determined with the same buffer without protein and subtracted from the protein spectra. The results were expressed as milli degrees. The fraction of secondary structure was calculated with the program of spectrum deconvolution CDPPro [42]. With this data was possible to see the integrity of the target protein.

Thermal denaturation: The thermal denaturation of the M2-1 protein was performed with the spectropolarimeter Jasco-815 with Peltier temperature control. The measure conditions were; temperature range 20°C to 80°C, with heating velocity of 1°C/min and readings at each 5°C. It was used a cuvette of 5 mm, wavelength of measurement 222nm with 2 seconds answer and slit opening of 1.0 nm, totalizing three accumulations.

Sample preparation: Stock solution of quercetin (MM = 338,3 Da) was prepared in 10% ultra-pure water and 90% ethanol, and M2-1 protein was prepared in a solution of PBS pH = 7.4. To determine the concentration of quercetin stock solution was used a molar extinction coefficient of $\epsilon_{375}=21.880 \text{ M}^{-1}\text{cm}^{-1}$ at 376nm [43] and, to determine the M2-1 concentration was used a molar extinction coefficient of $\epsilon_{280}=13.200 \text{ cm}^{-1}\text{M}^{-1}$ at 280nm [44].

Fluorescence spectroscopy: The fluorescence suppression was performed with the Cary Eclipse Fluorescence Spectrophotometer with Single Cell Peltier temperature control. To the titration of the protein with quercetin, first was determined the concentration of the protein to use a proportion of ligand:protein of 0.1 to 4.5. The titration was performed with temperature control. It was used a 1 cm quartz cuvette, with an emission and excitation slit opening of 5 nm. The excitation was in 280nm and emission was recorded between 295 until 500 nm, totalizing five accumulations.

Molecular docking: All molecular docking calculus were performed with AutoDock 4.0 [44]. The GRID were built centralized in the protein with a box width of 100x100x100 with spacing points of 0.5 Å. It was used Lamarckian Genetic Algorithm (LGA) search with Genetic Algorithm (GA) number of 200 steps, population size of 150 and medium of maximum evaluation number. The remaining parameters were kept as default. The results were analyzed with the MGL AutoDockTools [45] and the pictures were made with Chimera program [46] and Maestro 2015-4 [47].

Results and Discussion

The M2-1 gene was cloned in a 603bp length (Figure S1) inserted in pGEX4T-1 vector and transformed in BL21 Gold (DE3) strain. Five clones were selected to express the protein (Figure S2) and all of them expressed the protein M2-1 fused to GST (~48.47kDa).

Thus, the purified protein was sequenced by the Shotgun method. The primary sequence (Figure 1) was aligned with BLASTp Tool resulting in a 67% coverage and 46% identity with the M2-1 protein of hRSV (access code - AAX23995).

```

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYE
RDEGDKWRNKKFELGLEFPNLPPYIDGDVKLTQS
MAIIRYIADKHNMLGGCPKERAEISMLEGAVLDI
RYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDR
LCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCL
DAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQ
GWQATFGGGDHPPKSDLVPRGSMSRRNPCKFEIR
GHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRI
LKSMDKSIDTLSEISGAAELDRTEEYALGVVGVL
ESYIGSINNITKQSACVAMSKLLTELNSDDIKKL
RDNEELNSPKIRVYNTVISYIESNRKNNKQTIHL
LKRLPADVLKKTIKNTLDIHKSITINNPKELTVS
DTNDHAKNNDTT

```

Figure 1: M2-1 protein sequenced by Shotgun method. In bold is the GST sequence, in italic is M2-1 sequence, in bold and italic and underlined is thrombin cleaving site and underlined is the amino acid sequence found by Shotgun Method.

Once the protein was gotten, was necessary to observe if the protein was structured on the temperatures used in fluorescence experiments. The results of the circular dichroism show that the protein has a secondary pattern of approximated 40% of alpha helix, 10% of beta sheet, 19% of turns and 31% of random coils (50% turns and random coils) (Figure 2), showing that the protein is well structured. Our results are in accordance with the results found by Tran and colleagues [13] that shows the M2-1 protein is composed by 38% alpha helix, 12% beta sheet, 10% turns and 40% random coils.

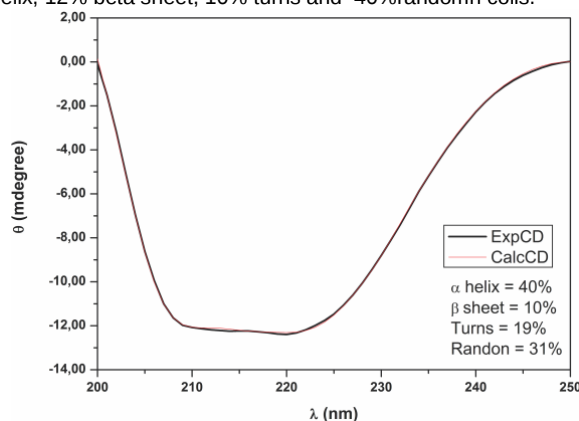


Figure 2: Far-UV CD spectra of M2-1 protein. The results are expressed as mili degrees. The dashed line represents the adjustments of the experimental data obtained from the deconvolution of the CD spectrum.

The temperature of melting obtained was $T_m = 328.6$ K showing that M2-1 has a good thermal stability considering the temperature range of many physiologic processes, whose is 310 K. By the analysis of the denaturation spectra (Figure 3) can be seen that in the biologic temperature range (until approximated 310 K), in the experimental conditions, there has been 10% denaturation, not interfering in the isothermal titration experiments performed to identify the interaction profile between M2-1 protein and Quercetin.

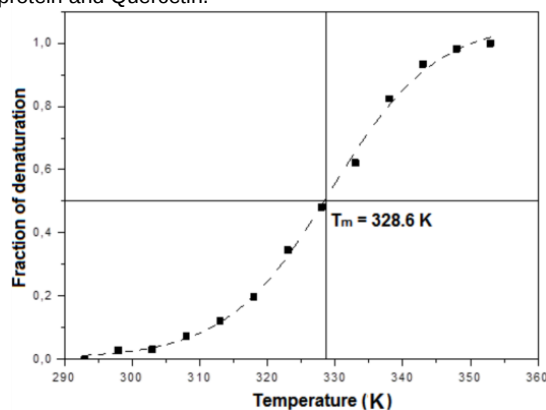


Figure 3: Thermal denaturation of the M2-1 protein. At the temperature of 328.6 K there is a balance in the concentration of the M2-1 protein in native state and denatured protein. The dashed line is a nonlinear adjustment of the point measure.

In Figure 4, are shown the fluorescence emission spectra at 298 K, as the suppression at 304 K and 310 K have a similar fluorescence characteristic.

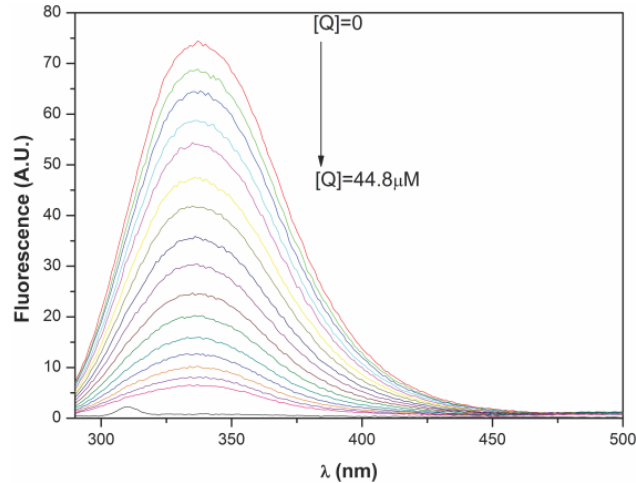


Figure 4: Fluorescence emission spectra titration of M2-1 protein with Quercetin at 25°C. The M2-1 protein was in a concentration of 10 μM and the concentration of Quercetin ([Q]) varied from 0 to 44.8 μM.

In the analysis of this result is admitted that the interaction of Quercetin with M2-1 suppresses totally the M2-1 fluorescence, it is, after each addition of Quercetin the remniscent fluorescence corresponds to the protein that has not interacted with Quercetin. So it is possible to calculate the fraction of the protein complexed with the ligand (f_b) throughout the fluorescence variation with the ligand increment:

$$f_b = \frac{F_0 - F}{F_0} = \frac{[PL]}{[P_T]}, \quad (1)$$

where F_0 is M2-1 protein fluorescence in the absence of the ligand, F is the fluorescence after the addition of the ligand, $[PL]$ is the concentration of the protein complexed with n ligands and $[P_T]$ is the total concentrations of the protein in the solution. Therefore, there is a relation between the fraction of the ligand bonded and the fraction of the protein complexed with the ligand f_b :

$$\vartheta = \frac{[L_b]}{[P_T]} = n f_b \quad (2)$$

where

$[L_b]$ = concentration of ligand bonded to the protein,
 n = number of ligands bonded to each protein.

Hence, can be constructed the Scatchard graphic and get the association constant and the number of ligand bound per protein. The result of this analysis (Figure 5) shows that there are two independent binding sites, one with a strong binding association ($\sim 1.5 \times 10^6 \text{ M}^{-1}$) being occupied by 0.25 ligand and another site with weak binding association ($\sim 1.1 \times 10^5 \text{ M}^{-1}$) being occupied by one ligand. The literature describes that the M2-1 protein forms tetramer, therefore, this result suggests the existence of one strong binding site formed by the four units of the tetramer that is occupied by one ligand (0.25 ligand/protein = 1 ligand/tetramer), and the second binding site is occupied by one ligand (1 ligand/protein = 4 ligands/tetramer) and the binding is not cooperative.

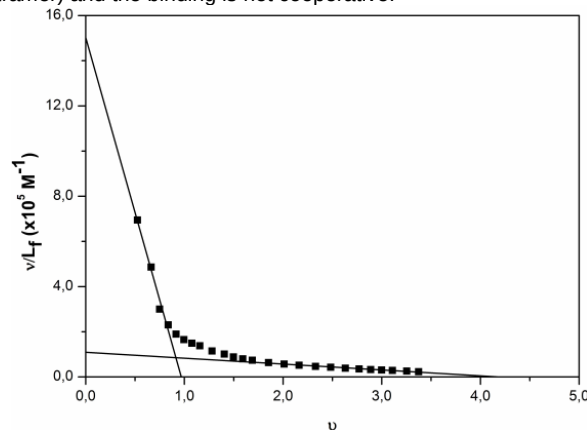


Figure 5: Scatchard graphic of the interaction between M2-1 protein and Quercetin. It is observed two different interaction sites that are not cooperative. The straight line represents the best adjustments.

With the objective of identifying the type of interaction between the M2-1 protein and Quercetin with the two binding sites, were performed titration in three different temperatures, 298 K, 304 K and 310 K.

The acquired results are shown in Table 1, the values of the two association constants K1 (strong site) and K2 (weak site) obtained in three different temperatures.

Table 1: Constant association of site 1 (K1) and site 2 (K2) in three different temperatures.

Constant	298 K	304 K	310 K
K1 ($\times 10^6 \text{ M}^{-1}$)	1.85 $\pm$ 0.04	1.60 $\pm$ 0.10	1.50 $\pm$ 0.10
K2 ($\times 10^5 \text{ M}^{-1}$)	0.82 $\pm$ 0.06	1.18 $\pm$ 0.02	1.09 $\pm$ 0.01

With the values of the constant can be constructed the Van'Hoff graphic to get the enthalpy interaction variation energy (ΔH), that is given by the declivity of the graphic $-\ln(K)$ vs $(1/T)$, where $R = 1,98 \text{ cal/Mol.K}$. In Figure 6, is shown this graphic for the two associations constant, where is verified that the strong site has a negative variation enthalpy (-3.6 Kcal/mol) (exothermic reaction), and the weak site has a positive variation enthalpy (4.4 Kcal/mol) (endothermic reaction).

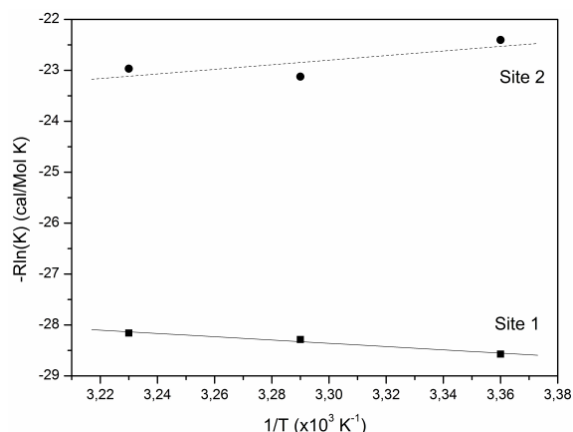


Figure 6: Van'Hoffe graph related to M2-1 protein interaction with Quercetin.

Then was calculated the Gibbs free energy with the association constant K through the equation:

$$\Delta G = -RT\ln(K) = \Delta H - T\Delta S \quad (3)$$

From this equation, can be get the values of Gibbs free energy (ΔG), the enthalpy variation (ΔH) and the entropy variation (ΔS), where is verified that the interaction between the ligand and the protein is energetic favorable ($\Delta G < 0$).

The results from Table 2 lists the thermodynamic parameters obtained in the spectrofluorometric titration experiments and in Figure 07 the variation of these in the temperature range explored. These results show that the values of ΔH and ΔS related to the two different sites, shows that the strong site (site 1) has an ionic interaction ($\Delta H < 0$ and $\Delta S > 0$) and the weak site (site 2) has a hydrophobic interaction between protein and ligand ($\Delta H > 0$ and $\Delta S > 0$).

As the N-terminus face of M2-1 is highly positive and the C-terminus face is also positive with four negative points [11], we can postulated that the negative points of Quercetin (formed by the hydroxyl and ketone group) could lead this molecule to interact at these two points. When those sites are full, the surface between the chains (which has neutral charge points), could accommodate the non-polar chain of Quercetin through hydrophobic interaction.

Table 2: ΔG , ΔH and $T\Delta S$ values of the interaction of the M2-1 protein with Quercetin to the both binding sites.

Kcal/mol	298K	304K	310K
$\Delta G1$	-8.15 ± 0.02	-8.60 ± 0.05	-8.73 ± 0.06
$\Delta G2$	-6.7 ± 0.05	-7.00 ± 0.01	-7.10 ± 0.07
$\Delta H1$	-3.2 ± 0.6	-3.2 ± 0.6	-3.2 ± 0.06
$\Delta H2$	4.4 ± 0.4	4.4 ± 0.4	4.4 ± 0.4
$T\Delta S1$	5.3 ± 0.9	5.4 ± 0.9	5.5 ± 0.9
$T\Delta S2$	11.1 ± 0.8	11.4 ± 0.2	11.5 ± 0.1

$\Delta G1$ = free energy variation on site 1, $\Delta G2$ = free energy variation on site 2, $\Delta H1$ = enthalpy variation on site 1, $\Delta H2$ = enthalpy variation on site 2, $T\Delta S1$ = entropy variation on site 1, $T\Delta S2$ = entropy variation on site 2.

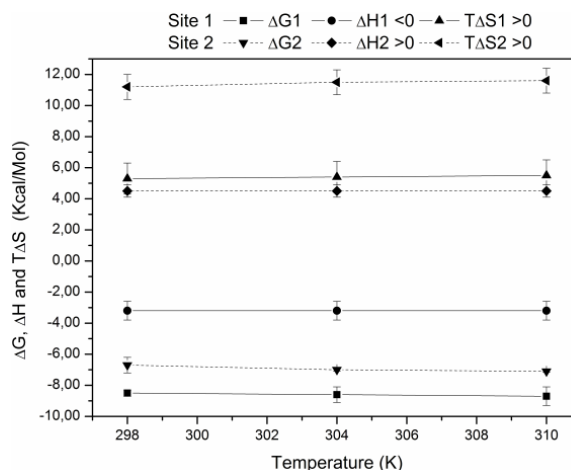


Figure 7: Dependence of ΔG , ΔH and $T\Delta S$ with temperature for the interaction of m2-1 protein with Quercetin in site 1 and 2.

Molecular docking.

The molecular docking study was performed on M2-1 monomer and tetramer. The monomer was divided in two regions (Figure S3) to perform the molecular docking. Site 1-1 (site 1 of the monomer) is encompassed by globular domain (P protein and RNA binding domain) and protein C-terminus, and site 2-1 (site 2 of the monomer) encompasses the zinc finger and oligomerization domain. On site 1-1, the Quercetin has a binding energy of -6.41 kcal/mol (Table 3) and makes three hydrogen bonds (Met50, Asp51 and Leu103). The ligand interacts in a pocket between the oligomerization domain and the

globular domain (Figure 8-A). On site 2-1, Quercetin has a binding energy of -5.21 kcal/mol (Table 3) and makes three hydrogen bonds (His14 and Phe41). In this site, the ligand interacts with the zinc finger domain, close to zinc metal (Figure 8-B).

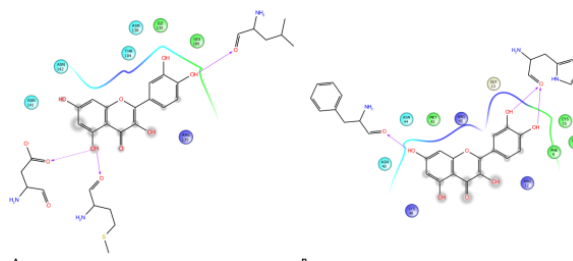


Figure 8: Molecular docking of Quercetin on M21 monomer. Quercetin was docked on site 1 (A) (RNA and P protein binding domain and N-terminus residues) and on site 2 (B) (zinc finger and oligomerization domain). Dark blue – positive charged; light blue – polar; green – hydrophobic; white – glycine; purple arrow – hydrogen bond; gray shadow – solvent exposure.

The docking on M2-1 tetramer was performed in three points (Figure S3) that encompasses the whole molecule. The ligand Quercetin was docked on the central cavity on N-terminus (site 1-4 – site 1 of the tetramer) and C-terminus (site 2-4 – site 2 of the tetramer) face and on the interface between two chains (site 3-4 – site 3 of the tetramer) (These faces names were taken from Tanner et al., 2014).

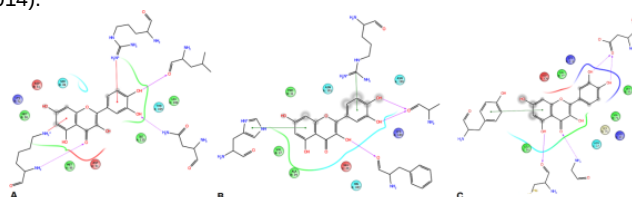


Figure 9: Molecular docking of Quercetin on M21 tetramer. Quercetin was docked on site 1 (A) (N-terminus face), site 2 (B) (C-terminus face) and on site 3 (C) (interface between two chains). Dark blue – positive charged; red – negative charged; light blue – polar; green – hydrophobic; white – glycine; purple arrow – hydrogen bond; green line – pi-pi/pi-cation stacking; gray shadow – solvent exposure.

On site 1-4 the binding energy of quercetin was -7.36 kcal/mol (Table 3). The ligand interacted in a region between 2 chains (Figure 9-A), making three hydrogen bonds (Lys52, Leu103 and Asn138) and two pi-cation interactions (Lys52 and Arg139). On site 2-4, the binding energy was -5.24 kcal/mol (Table 3) and Quercetin made three hydrogen bonds (Ph28 and Ala188) and two pi-pi stacking (Arg3 and His33) (Figure 9-B). The molecule interacts in a region between the zinc finger domain and the unstructured C-terminus domain, close to oligomerization domain. For site 3-4, the binding energy was -7.36 kcal/mol (Table 3) and made four hydrogen bonds (Cys15, Gly18 and Asp67). The flavonol Quercetin interacts in a pocket between the P protein and RNA domain and the oligomerization domain, close to zinc metal (Figure 9-C).

Table 3: Binding energy values (kcal/mol) of Quercetin on sites 1, 2 and 3 of the monomer and the tetramer

	Monomer	Tetramer
	Binding energy	Binding energy
Site 1	-6.41	-7.36
Site 2	-5.21	-5.24
Site 3	-	-7.36

By the docking we can postulated that Quercetin interacts preferably on site 1 than on 2, close to $\alpha 2$ helix (where M2-1 binds to RNA) [48]. Also, this hypothesis can be reinforced by the fact that Quercetin has some similarity to adenine which was shown that the residues Ser58-Lys177 have high affinity for uracil and adenine [11].

The characteristic of the interaction on the site with strong binding association (site 1 and 2) and the one with weak binding association (site 3) correlates between experimental data and molecular docking data. Sites 1 and 2, on docking, have a more ionic profile than on site 3 that has a more hydrophobic profile

Conclusion

In conclusion, the M2-1 protein is mainly composed by alpha helix, having a temperature of melting of about 328.6 K. From fluorescence suppression, we observed that Quercetin interacts in two sites, one with a strong constant association (site 1 = $\sim 1.5 \times 10^6 \text{ M}^{-1}$) and another with weak association constant (site 2 = $1.1 \times 10^5 \text{ M}^{-1}$). The site 1 is occupied by 0.25 ligand / protein, or 1 ligand/tetramer, and site 2 is occupied by 1 ligand/protein, or 4 ligands/tetramer. This interaction is not cooperative, meaning that the interaction of one molecule is independent of another. The thermodynamics analysis shows that the interaction of Quercetin at site 1 is through ionic interaction ($\Delta H < 0$ and $\Delta S > 0$) and at site 2 is through hydrophobic interaction ($\Delta H > 0$ and $\Delta S > 0$).

The molecular docking of Quercetin in monomeric form shows that the interaction occurs in superficial pockets between the protein C-terminus and RNA and protein P interaction domain. This interaction is on a polar pocket stabilized by hydrogen bonds. However, the docking with the tetrameric form occurs in two points, on N-terminus face close do $\alpha 2$ helix and between chains in the interface the interaction of oligomerization and RNA and P protein interaction domain. In the first case, the interaction is on a more ionic pocket stabilized by hydrogen bonds and pi-pi/pi-cation stacking and, on the second case Quercetin interacts in a more hydrophobic pocket stabilized by pi-pi stacking and hydrogen bond. So, Quercetin could interact with M2-1 in formation avoiding its tetramerization and interacts with M2-1 tetramer, which could avoid virus replication and budding.

Acknowledgments.

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP – Processo 13/24355-2). We thanks to Prof. Dr. Altair Benedito Moreira to for the access to the fluorescence spectrometer

(Processo FAPESP: 2014/17511-0), to Dr^a Nathalia Baptista Dias and Prof. Dr. Mário Sergio Palma to perform the protein sequencing.

References

- [1] Borchers AT, Chang C, Gershwin ME, Gershwin LJ. Respiratory syncytial virus—a comprehensive review. *Clin Rev Allergy Immunol*, 45, 2013, 331–379.
- [02] Collins PL, Melero JA. Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years. *Virus Res* 162, 2011, 80–99.
- [03] Nair H, Nokes DJ, Gessner BD, Dherani M, Madhi SA, Singleton RJ, O'Brien KL, Roca A, Wright PF, Bruce N, Chandran A, Theodoratou E, Sutanto A, Sedyaniingsih ER, Ngama M, Munywoki PK, Kartasasmita C, Simões EA, Rudan I, Weber MW, Campbell H. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *Lancet* 375, 2010, 1545–1555.
- [04] Baker KA, Ryan ME. RSV infection in infants and young children. What's new in diagnosis, treatment, and prevention? *Postgrad Med* 106, 1999, 97–9, 103–4, 7–8 passim.
- [05] Henderson FW, Collier AM, Clyde WA Jr, Denny FW. Respiratory syncytial virus infections, reinfections and immunity. A prospective, longitudinal study in young children. *N Engl J Med*, 300, 1979, 530–534.
- [06] Santos NSO, Romanos MTR, Wigg MD, Introdução à Virologia Humana, 2^a edição, Rio de Janeiro, Guanabara Koogan, 2008.
- [07] Collins PL, Hill MG, Grosfeld H. Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *PNAS* 93, 1996, 81–85.
- [08] Ghildyal R, Li D, Peroulis I, Shields B, Bardin PG, Teng MN, Collins PL, Meanger J, Mills J. Interaction between the respiratory syncytial virus G glycoprotein cytoplasmic domain and the matrix protein., *J. Gen. Virol.*, 86, 2005, 1879–1884.
- [09] Dongsheng Li, David A. Jans, Phillip G. Bardin, Jayesh Meanger, John Mills, Reena Ghildyal. Association of Respiratory Syncytial Virus M Protein with Viral Nucleocapsids Is Mediated by the M2-1 Protein. *J Virol*, 82, 2008, 8863–8870.
- [10] Ghildyal R, Mills J, Murray M, Vardaxis N, Meanger J. Respiratory syncytial virus matrix protein associates with nucleocapsids in infected cells., *J. Gen. Virol.*, 83, 2002, 753–757.
- [11] Tanner SJ, Ariza A, Richard CA, Kyle HF, Dods RL, Blondot ML, Wu W, Trincão J, Trinh CH, Hiscox JA, Carroll MW, Silman NJ, Eléouët JF, Edwards TA, Barr JN. Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 2014, 1580–5.
- [12] Dubosclard V, Blondot ML, Eléouët JF, Bontems F, Sizun C. 1H, 13C, and 15N resonance assignment of the central domain of hRSV transcription antitermination factor M2-1, *Biomol NMR Assign*, 5, 2011, 237–239.
- [13] Tran TL, Castagné N, Dubosclard V, Noinville S, Koch E, Moudjou M, Henry C, Bernard J, Yeo RP, Eléouët JF. The Respiratory Syncytial Virus M2-1 Protein Forms Tetramers and Interacts with RNA and P in a Competitive Manner, *J. Virol.*, 83, 2009, 6363–6374.
- [14] Hardy RW, Wertz GW. The Cys(3)-His(1) motif of the respiratory syncytial virus M2-1 protein is essential for protein function. *J Virol* 74, 2000, 5880–5885.
- [15] Tang RS, Nguyen N, Cheng X, Jin H. Requirement of cysteines and length of the human respiratory syncytial virus M2-1 protein for protein function and virus viability. *J Virol* 75, 2001, 11328–11335.
- [16] Zhou H, Cheng X, Jin H. Identification of amino acids that are critical to the processivity function of respiratory syncytial virus M2-1 protein. *J Virol* 77, 2003, 5046–5053
- [17] Pelletier AJ, Mansbach JM, Camargo CA Jr. Direct medical costs of bronchiolitis hospitalizations in the United States., *Pediatrics*, 118, 2006, 2418–2423.
- [18] Nair H, Brooks WA, Katz M, Roca A, Berkley JA, Madhi SA, Simmerman JM, Gordon A, Sato M, Howie S, Krishnan A, Ope M, Lindblade KA, Carosone-Link P, Lucero M, Ochieng W, Kamimoto L, Dueger E, Bhat N, Vong S, Theodoratou E, Chittaganpitch M, Chimah O, Balmaseda A, Buchy P, Harris E, Evans V, Katayose M, Gaur B, O'Callaghan-Gordo C, Goswami D, Arvelo W, Venter M, Briesse T, Tokarz R, Widdowson MA, Mounts AW, Breiman RF, Feikin DR, Klugman KP, Olsen SJ, Gessner BD, Wright PF, Rudan I, Broor S, Simões EA, Campbell H. Global burden of respiratory infections due to seasonal influenza in young children: a systematic review and meta-analysis. *Lancet* 2010; 378:1917–30
- [19] Sidwell RW, Barnard DL. Respiratory syncytial virus infections: recent prospects for control. *Antiviral Res*, 71, 2006, 379–390.
- [20] Crotty S, Andino R. Implications of high RNA virus mutation rates: lethal mutagenesis and the antiviral drug ribavirin. *Microbes Infect.* 4, 2002, 1301–1307.
- [21] Crotty S, Cameron C, Andino R. Ribavirin's antiviral mechanism of action: lethal mutagenesis? *J Mol Med* 80, 2002, 86–95.
- [22] Fenton C, Scott LJ, Plosker GL. Palivizumab: a review of its use as prophylaxis for serious respiratory syncytial virus infection., *Paediatr Drugs*, 6, 2004, 177–197.
- [23] Wu H, Pfarr DS, Tang Y, An LL, Patel NK, Watkins JD, Huse WD, Kiener PA, Young JF. Ultra-potent antibodies against respiratory syncytial virus: effects of binding kinetics and binding valence on viral neutralization. *J Mol Biol.*, 350, 2005, 126–144.
- [24] American Academy of Pediatrics: respiratory syncytial virus. In *Red Book: 2009 Report of the Committee on Infectious Diseases*. Edited by Pickering LK, Baker CJ, Kimberlin DW, Long SS. Elk Grove Village, IL. American Academy of Pediatrics; 2009; 560–568.
- [25] Butler MS. Natural products to drugs: natural product derived compounds in clinical trials, *Nat. Prod. Rep.*, 22, 2005, 162–195.
- [26] Harvey AL. Natural products in drug discovery *Drug Discovery Today* 13, 2008, 894–901.
- [27] Simões CMO, Schenkel EP, Gosmann G, *Farmacognosia - da Planta ao Medicamento*, 5^a ed., Editora da UFSC: Santa Catarina, 2004.
- [28] Cazarolli LH, Zanatta L, Alberton EH, Figueiredo MS, Folador P, Damazio RG, Pizzolatti MG, Silva FR. Flavonoids: prospective drug candidates. *Med. Chem.*, 8, 2008, 1429–1440.
- [29] Wiegand H, Boesch-Saadatmandi C, Regos I, Treutler D, Wolfram S, Rimbach G, Effects of quercetin and catechin on hepatic glutathione-S transferase (GST), NAD(P)H quinone oxidoreductase 1 (NQO1), and antioxidant enzyme activity levels in rats, *Nutr. Cancer* 61, 2009, 717–722.

- [30] Binsack R, Boersma BJ, Patel RP, Kirk M, White CR, Darley-Usmar V, Barnes S, Zhou F, Parks DA. Enhanced antioxidant activity after chlorination of quercetin by hypochlorous acid. *Alcohol. Clin. Exp. Res.* 25, 2001, 434–443.
- [31] Murakami A, Ashida H, Terao J. Multitargeted cancer prevention by quercetin. *Cancer letters.* 269, 2008, 315–325.
- [32] Xavier CP, Lima CF, Rohde M, Pereira-Wilson C. Quercetin enhances 5-fluorouracil-induced apoptosis in MSI colorectal cancer cells through p53 modulation. *Cancer chemotherapy and pharmacology.* 68, 2011, 1449–1457.
- [33] Dajas F. Life or death: neuroprotective and anticancer effects of quercetin. *Journal of ethnopharmacology.* 143, 2012, 383–396.
- [34] Murphy EA, Davis JM, McClellan JL, Carmichael MD. Quercetin's effects on intestinal polyp multiplicity and macrophage number in the Apc(Min/+) mouse. *Nutrition and cancer.* 63, 2011, 421–426.
- [35] Wu LL, Yang XB, Huang ZM, Liu HZ, Wu GX. *In vivo* and *in vitro* antiviral activity of hyperoside extracted from *Abelmoschus manihot* (L) medik. *Acta pharmacologica Sinica.* 2007;28:404–409.
- [36] Pisonero-Vaquero S, Garcia-Mediavilla MV, Jorquera F, Majano PL, Benet M, Jover R, Gonzalez-Gallego J, Sanchez-Campos S. Modulation of PI3K-LXRalpha-dependent lipogenesis mediated by oxidative/nitrosative stress contributes to inhibition of HCV replication by quercetin. *Laboratory investigation; a journal of technical methods and pathology.* 94, 2014, 262–274.
- [37] Yu YB, Miyashiro H, Nakamura N, Hattori M, Park JC. Effects of triterpenoids and flavonoids isolated from *Alnus firma* on HIV-1 viral enzymes. *Archives of pharmacal research.* 30, 2007, 820–826.
- [38] Spedding G, Ratty A, Middleton E. Jr. Inhibition of reverse transcriptases by flavonoids. *Antiviral research.* 12, 1989, 99–110.
- [39] Ahmed-Belkacem A, Guichou JF, Brillet R, Ahnou N, Hernandez E, Pallier C, Pawlotsky JM. Inhibition of RNA binding to hepatitis C virus RNA-dependent RNA polymerase: A new mechanism for antiviral intervention. *Nucleic Acids Res* 42, 2014, 9399–9409.
- [40] Kaul TN, Middleton Jr E, Ogra PL. Antiviral Effect of Flavonoids on Human Viruses. *Journal of Medical Virology,* 15, 1985, 71-79.
- [41] Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* 11, 2012, O111.016717–O111.016717.
- [42] Sreerama N, Woody RW. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set., *Anal Biochem.* 15, 2000, 252-60
- [43] The Merck Index, 11th Ed., Entry# 8044.
- [44] Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. **Protein Identification and Analysis Tools on the ExPASy Server;** (In) John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press (2005). pp. 571-607.
- [44] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. **J. Computational Chemistry**, 16, 2009, 2785-91.
- [45] Sanner MF. Python: A Programming Language for Software Integration and Development. *J. Mol. Graphics Mod.*, 17, 1999, 57-61.
- [46] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem.* 25, 2004, 1605-12.
- [47] Maestro, version 10.4, Schrödinger, LLC, New York, NY, 2015.
- [48] Blondot ML., Dubosclard V, Fix J, Lassoued S, Aumont-Nicaise M, Bontems F, Sizun C. Structure and functional analysis of the RNA- and viral phosphoprotein-binding domain of respiratory syncytial virus M2-1 protein. *PLoS Pathogens*, 8, 2012, 1002734.