



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

Mírian Feliciano da Costa

Docking, Dinâmica Molecular e Redes Neurais Artificiais na busca
por flavonoides anti-RSV

São José do Rio Preto
2018

Mírian Feliciano da Costa

Docking, Dinâmica Molecular e Redes Neurais Artificiais na busca
por flavonoides anti-RSV

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.

Financiadora: CAPES

Orientador: Prof^a. Dra. Karina Alves de Toledo

Co-orientador: Prof^a. Dra. Fátima Pereira de Souza

São José do Rio Preto
2018

Costa, Mírian Feliciano da.

Docking, Dinâmica Molecular e Redes Neurais Artificiais na busca por flavonoides anti-RSV / Mírian Feliciano da Costa . – São José do Rio Preto, 2018

211 f.: il., grafs., tabs.

Orientador: Karina Alves de Toledo

Coorientador: Fátima Pereira de Souza

Tese (doutorado) – Universidade Estadual Paulista (UNESP), Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto

1. Microbiologia. 2. Quercetina. 3. Proteínas. 4. Flavonóides.
5. Vírus Sincicial Respiratório Humano. I. Título.

CDU – 576.858

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

Mírian Feliciano da Costa

Docking, Dinâmica Molecular e Redes Neurais Artificiais na busca
por flavonoides anti-RSV

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.

Financiadora: CAPES

Comissão Examinadora

Prof^a Dra. Karina Alves de Toledo
UNESP – Assis
Orientadora

Prof. Dr. Aripuanã Watanabe
UFJF – Juiz de Fora

Prof. Dr. Luís Octávio Regasini
UNESP – São José do Rio Preto

Prof. Dr. José Celso Rocha
UNESP – Assis

Dra. Luciane Sussuchi da Silva
Hospital do Câncer de Barretos

Assis
18 de dezembro 2018

À minha mãe, minha inspiração.

AGRADECIMENTOS

A gratidão é uma das formas mais belas e nobres de expressarmos nosso apreço por estarmos vivos e pelas oportunidades de (con)vivência e de aprendizado. A vida é uma confluência de nossos relacionamentos, do quanto somos influenciados por eles e das escolhas que fazemos, onde quer que estejamos. Não há como acreditar em Deus sem acreditar que Ele usa pessoas e nossos relacionamentos com elas para moldar a dimensão de e o nosso relacionamento para com Ele. Poucas páginas não são suficientes para mensurar o quanto eu tenho a agradecer, por tudo que vivi e aprendi nesta etapa, com pessoas maravilhosas e também com outras nem tanto assim.

Sou grata primeiramente à Deus. Pelo dom da fé, pela vida, pela provisão e pela capacitação em mais este percurso. Por me agradecer com pessoas fantásticas e corajosas que toparam, digamos que “voluntariamente”, dividir mais esta jornada desafiadora comigo. Isso já seria o suficiente por ser dito, mas os corajosos merecem ser mencionados. Então vamos lá...

Aos meus pais – Anselmo e Edna – e às minhas irmãs – Débora e Priscila – os quais compõem minha agência de fomento emocional, psicológico e também financeiro. Eles compartilharam momentos únicos de alegria e vitórias, e também dividiram angústias, incertezas, dificuldades. Entretanto, estiveram sempre ali torcendo por mim e me animando a prosseguir. *“No abraço distante, ao seu lado...”* Não tenho palavras para mensurar minha gratidão a eles.

Ao meu querido – Terkimbi –, por, dentre tantos outros motivos, auxiliar-me a manter em perspectiva o essencial e a não me deixar abalar por circunstâncias e pessoas não-favoráveis.

Aos meus sobrinhos – Maria Eduarda e Miguel – pelo olhar terno, simples e belo que me animou a continuar.

À minha orientadora – Karina – por todas as (con)vivências e ensinamentos. Isso já diz muito. Também pelo exemplo pessoal de superação e de humildade, pelo apoio, pela amizade, por me encorajar a alcançar voos maiores e melhores e pelos conselhos de resiliência em momentos oportunos.

À professora Fátima pela co-orientação e pela amizade. E também ao Ícaro, pela sua sempre constante disponibilidade em ajudar.

Ao professor José Celso pela orientação e pelo suporte com a parte de Redes Neurais. Ao Diogo pela disposição em sempre ajudar. E à professora Juliana pelas dicas, pela amizade e pela ajuda com esse aspecto também.

Ao programa de pós-graduação (PPG) em Microbiologia pelas oportunidades de aprimoramento.

À seção de PG pela prestatividade e em especial aos funcionários Alex e Felipe pela solicitude.

À professora Eleni, por ser exigente e com isso ter me incentivado a me aperfeiçoar e a crescer profissionalmente. E na figura dela eu agradeço a todos os professores de toda essa minha longa jornada de conhecimento e aprendizado. Desde a mais tenra idade, aqueles que me abriram o mundo às letras, às palavras, à diversidade de linguagens, bem como os que ensinaram a poesia divina traduzida na matemática e nas suas proporções, na dança dos átomos orquestrada pela química e pela física e representada na biologia e na geografia, e assim também os arranjos sociais, suas flutuações e excentricidades ao longo do tempo. Muito obrigada! A maioria de vocês realmente professou o conhecimento e iluminou a '*sem luz*' aqui (essência do vocábulo "*aluno*") através de suas aulas e dos conhecimentos que me foram passados, permitindo-me galgar muitos desafios, dentre os quais mais este.

Ao projeto Erasmus-IBrasil pela bolsa e pela oportunidade de intercâmbio na Universidade De Roma – *Tor Vergata*.

Ao professor Alessandro Desideri por me possibilitar realizar os experimentos de dinâmica molecular em seu laboratório.

Aos meus colegas de laboratório de Roma – Federico, Blasco, Francesca, Daniele, Jéssica, Fábio(s), Kumar, Bini, Jeg, Venn – pelos bons momentos. Em especial ao Federico, pela paciência em me ensinar e em me orientar durante o tempo em que estive ali... "*Se hai fatto bene, funziona!*" E também a Jessiquinha, Kumar, Bini, Jeg e Kassim pelo auxílio mútuo e por todos finais de semana em que estivemos juntos. "*Good friends, good wine, good times.*"

Minha profunda gratidão à Joseane Scarpassa, ao Guilherme, ao senhor César e à dona Célia por todas acolhidas carinhosas durante as inúmeras idas e vindas para São José do Rio Preto.

Aos meus amigos de Assis Naiana, Pedro, Letícia, Júlia, Gilsomar, Fabiana, e Silvânia pelas alegrias, vivências compartilhadas e os três últimos pelo encorajamento na fé.

Aos meus companheiros de laboratório pelos (onipresentes) cafés e aprendizados, e pelas memórias pitorescas (“depois de mim sou eu”) ... Amandinha, Venon (Gabriela), Pri, Lari, Bruno... Obrigada! Em especial ao último pelo auxílio com os vídeos das dinâmicas.

À Capes pela bolsa.

“O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001”.

Finalmente, gratidão a você que dispôs de seu tempo para ler esta tese. Espero que, ao final, você possa perceber que o conhecimento é uma jornada: de busca, de expansão, de contradição, e de reticências... A qual nos confere humildade, pelo simples fato de que quanto mais nela avançamos, mais conscientes nos tornamos da veracidade contida no *“Só sei que nada sei”* de Sócrates... E que essa jornada nos abre os olhos, a mente e a sensibilidade, e nos leva a enxergar e ir além... E que isso vale à pena!

“O seu tempo é a sua vida (...) e por isso o melhor presente que você pode dar a alguém é o seu tempo...” Rick Warren

Minha inestimável gratidão a vocês e àqueles que porventura minha memória tenha ofuscado os nomes... Muito obrigada!

*“Não repreenda o zombador, caso contrário ele o odiará; repreenda o sábio, e ele o
amará.
Instrua o homem sábio, e ele será ainda mais sábio; ensine o homem justo, e ele
aumentará o seu saber.”
(Provérbios 9:8,9)*

*“O amor é a última consequência do conhecimento.”
(Mário César Rodrigues)*

*“Você, Quaresma, é um visionário...”
(Lima Barreto)*

RESUMO

Respirar é essencial à sobrevivência humana, tal que o trato respiratório é alvo frequente de patógenos. O Vírus Sincicial Respiratório (RSV) causa bronquiolite e pneumonia, especialmente em recém-nascidos e idosos, o que resulta em milhões de mortes e hospitalizações. Sua glicoproteína de superfície F – proteína de fusão – é responsável pela fusão do envelope viral à membrana celular. Diversos flavonoides demonstram atividade anti-RSV por meio do mecanismo virucida. Assim, os objetivos desse trabalho foram analisar a interação da proteína F modelada neste trabalho (RSVmF) com o flavonoide quercetina, seu derivado pentacetilado (Q1) e um ligante-referência (JNJ-2408068) por *docking* e dinâmica molecular (DM). Ademais, elaboramos uma rede neural artificial (ANN) assessorada pelo algoritmo genético para prever a capacidade anti-RSV de flavonoides. A modelagem da estrutura completa da proteína F, os *dockings*, a DM, e a ANN foram realizados, respectivamente, no servidor I-TASSER, no Auto-Dock/Vina 1.1.2., no Amber/14, e na plataforma Matlab®. Os experimentos de *docking* indicaram que os três ligantes preenchem a cavidade central de RSVmF e apresentam energia de interação em torno de -6,5 kcal/mol. Análises de DM revelaram tendência ao escape da cavidade para quercetina, enquanto Q1 e ligante-referência permanecem nela. Cálculos de MM-GBSA mostraram ΔG (kcal/mol) dos complexos RSVmF-quercetina= -4,92, RSVmF-Q1= -14,22, e RSVmF-inibidor = -22,78. Logo, a acetilação da quercetina aumentou sua capacidade virucida por favorecer a interação com a cavidade de RSVmF. Quanto à ANN, os resultados obtidos a partir de uma amostra (n=240 dados, referentes a 70 flavonoides), indicaram três possíveis modelos de predição, dentre os quais ANN-g2 obteve uma acurácia de 90% para predição de flavonoides “ativos” e 93,3% para “inativos”. O principal ganho referente ao nosso estudo é o respaldo à relevância de um planejamento racional de procedimentos experimentais, pela exclusão preliminar de flavonoides inativos e a pré-seleção de ativos.

Palavras-chave: RSV. Quercetina. Derivado de quercetina. Proteína F. Docking. Dinâmica Molecular. Redes Neurais Artificiais.

ABSTRACT

Breathing is a basic requirement for survival of human life and the respiratory tract is therefore a frequent target for pathogens. Respiratory syncytial virus (RSV) causes bronchiolitis and pneumonia, mainly in newborns and the elderly, which results in millions of deaths and hospitalizations. Its surface glycoprotein F – fusion protein – is responsible for the fusion of the viral envelope to cell membrane. Several flavonoids were shown to exert anti-RSV activity by virucidal mechanism. Thus, the aims of this study were to analyze the interaction of the F protein modeled herein (RSVmF) in complex with flavonoid quercetin, with its pentacetylated derivative (Q1) and with a reference-ligand (JNJ-2408068), by docking and molecular dynamics (MD). Moreover, we developed an artificial neural network (ANN), assisted by the genetic algorithm to predict the anti-RSV activity of flavonoids. The modeling of the complete protein F structure, dockings, DM, and ANN were performed, respectively, on I-TASSER server, Auto-Dock/Vina 1.1.2., Amber/14, and Matlab® platform. Docking experiments indicated that the three ligands fill RSVmF central cavity and present an interaction energy of about -6.5 kcal/mol. MD analyses revealed an escaping trend from the cavity to quercetin, whereas Q1 and the reference-ligand remain in it. MM-GBSA calculations showed ΔG (kcal/mol) for the complexes RSVmF-quercetin= -4.92, RSVmF-Q1= -14.22, and RSVmF-inhibitor = -22.78. Therefore, the acetylation of quercetin increased its virucidal activity by favoring the interaction with RSVmF cavity. ANN results were obtained from a sample (n = 240 data, regarding 70 flavonoids) and indicated three possible predictive models, among which ANN-g2 revealed an accuracy of 90% for prediction of “active” flavonoids and 93.3% for “inactive” ones. The major achievement of our study is that it has demonstrated the relevance of rational planning of experimental procedures, by the preliminary exclusion of inactive flavonoids and the pre-selection of active ones.

Keywords: RSV. Quercetin. Quercetin derivative. F protein. Docking. Molecular Dynamics. Artificial Neural Networks.

LISTA DE ILUSTRAÇÕES

Figura 1 – O sistema respiratório	22
Figura 2 – Tipos celulares e epitélios do sistema respiratório – traqueia/brônquios, bronquíolos e alvéolos	23
Figura 3 – Fatores intrínsecos e extrínsecos que interferem na resposta inflamatória nas vias aéreas	24
Figura 4 – Formação de sincícios pelo RSV	28
Figura 5 – Formação de sincícios in vitro	28
Figura 6 – Pleomorfismo do RSV	30
Figura 7 – Representação esquemática do RSV, de suas proteínas e de seu genoma	31
Figura 8 – Estrutura resolvida da Nucleoproteína – Proteína N – do RSV	33
Figura 9 – Estrutura resolvida da Proteína de Matriz – Proteína M – do RSV	34
Figura 10 – Proteína F: sua sequência e modificações.....	37
Figura 11 – Estruturas resolvidas da Proteína de Fusão – Proteína F – em seus estados de pré e pós fusão	38
Figura 12 – Estrutura resolvida da Proteína M2-1 do RSV	39
Figura 13 – O processo de fusão entre o envelope do RSV e a membrana celular.....	44
Figura 14 – Síntese do ciclo replicativo do RSV	46
Figura 15 – O RSV e seus mecanismos de evasão do sistema imune	48
Figura 16 – O ciclo viral do RSV e algumas estratégias de inibição	51
Figura 17 – A engenharia dos anticorpos.....	52
Figura 18 – Palivizumabe (MEDI-493)	52
Figura 19 – Os sítios antigênicos da proteína F	53
Figura 20 – A Ribavirina.....	54
Figura 21 – RSV-F modelada neste trabalho e o bolsão trissimétrico dentro de sua cavidade central	56
Figura 22 – Estrutura básica dos flavonoides e alguns de seus grupos	60
Figura 23 – Acetilação da quercetina	63
Figura 24 – Do DNA à Proteína - fluxograma simplificado	68

Figura 25 – Resumo das principais técnicas de modelagem de proteínas	71
Figura 26 – Representação básica de um Docking Molecular	72
Figura 27 – Cálculos envolvidos no campo de forças	76
Figura 28 – Fluxograma simplificado de Simulação de Dinâmica Molecular	77
Figura 29 – Redes Neurais	79
Figura 30 – Algoritmo genético	90
Figura 31 – Resumo metodologia	91
Figura 32 – RSV-F pré-fusão modelada (RSVmF).....	93
Figura 33 – Avaliação de RSVmF por resíduo	96
Figura 34 – Desvio Quadrático Médio (RMSD) de RSVmF.....	99
Figura 35– Flutuação quadrática média (RMSF) de RSVmF	101
Figura 36 – Estrutura secundária RSVmF.....	103
Figura 37– Eigenvalues da matriz de covariância de RSVmF	105
Figura 38 – Porção final de RSVmF e seus aminoácidos principais	108
Figura 39 – Docking molecular dos complexos RSVmF-quercetina, RSVmF-Q1 e RSVmF-JNJ	111
Figura 40 – Simulação de DM: RSVmF-QCT, RSVmF-Q1, RSVmF-JNJ	112
Figura 41 – Variação da distância RSVmF-ligantes.....	114
Figura 42 – Quercetina, Q1 e inibidor JNJ-2408068	119
Figura 43 – Acetilação da quercetina	121
Figura 44 – Primeira simulação ANN	127
Figura 45 – Agrupamento simulação 2.....	133
Figura 46 – Matrizes de confusão da simulação 2	134
Figura 47 – Curva ROC ANN	137
Figura 48 – Desempenho dinâmico ANN-g2.....	138

LISTA DE GRÁFICOS

Gráfico 1 – MM-GBSA: Quercetina, Q1 e JNJ-2408068	117
---	-----

LISTA DE TABELAS

Tabela 1 – Análise de Ramachandran para RSVmF e 4JHW	94
Tabela 2 – Persistência das ligações de hidrogênio entre RSVmF e quercetina/ligante-referência para uma simulação de 20 ns	116
Tabela 3 – Resultados binários da simulação 1	129
Tabela 4 – Hierarquia Ativos e Inativos.....	131
Tabela 5 – Parâmetros das melhores arquiteturas ANN	135

LISTA DE ABREVIATURAS E SIGLAS

ANN *Artificial Neural Network*¹

CPU *Central Processing Unit*

CX3CR *CX3C chemokine receptor*

DM *Dinâmica Molecular*

DNA *DeoxyriboNucleic Acid*

2D *2 dimensões (bidimensional)*

FIP *Family Interacting Protein*

GCRR *G protein Cysteine-Rich Region*

GPU *Graphics Processing Unit*

HIV *Human Immunodeficiency Virus*

ICAM *Intercellular Adhesion Molecule*

IFN *Interferon*

Ig *Immunoglobulin*

IL *Interleucin*

IP *Interferon- γ -Inducible Protein*

MAVs *Mitochondrial AntiViral Signalling Protein*

MIP *Macrophage Inflammatory Protein*

mRNA *messenger RNA*

pDC *Plasmacytoid Dendritic Cell*

QSAR *Quantitative Structure-Activity Relationship*

Rab *Ras Related Protein*

RANTES *Regulated on Activation, Normal T Expressed and Secreted*

RNA *RiboNucleic Acid*

RSVmF *Proteína F modelada neste trabalho*

SAR *Structure-activity Relationship*

STAT *Signal Transducer and Activator of Transcription*

3D *3 dimensões (tridimensional)*

TLR *Toll-Like Receptor*

TNF *Tumor Necrosis Factor*

¹ A fim de evitar possíveis confusões com as siglas “RNA”, as quais podem ser referentes ao “ácido ribonucleico”, bem como às “redes neurais artificiais”, nós adotamos exclusivamente a

terminologia “ANN” para se referir às redes neurais artificiais.

SUMÁRIO

INTRODUÇÃO	18
1. INTRODUÇÃO GERAL	18
1.1 Infecções do trato respiratório	21
1.1.2 O vírus sincicial respiratório (RSV)	27
1.1.2.1 Ciclo do RSV	42
1.1.2.1.2 Estratégias Antivirais contra RSV	50
1.2 Flavonoides	58
1.3 Estratégias Computacionais	65
2. OBJETIVOS.....	83
3. METODOLOGIA	84
<i>3.1 Modelagem Molecular da Glicoproteína F do RSV</i>	<i>84</i>
<i>3.2 Preparação do modelo para simulação de Dinâmica Molecular (MD)</i>	<i>85</i>
<i>3.3 Protocolo de dinâmica molecular (MD)</i>	<i>85</i>
<i>3.4 Análises da simulação de MD</i>	<i>86</i>
<i>3.5 Procedimento de Docking Molecular</i>	<i>86</i>
<i>3.6 Elaboração da base de dados de flavonoides para Rede Neural Artificial</i>	<i>87</i>
3.7 ANN	88
3.8 Algoritmo genético (AG).....	89
4. RESULTADOS E DISCUSSÃO	92
4.1 Modelagem de RSVmF, dinâmica molecular e análises	92
4.2 Docking e Dinâmica Molecular RSVmF-quercetina, RSVmF-Q1 e análises.....	110
4.3 ANN: elaboração e implementação	125
5. CONCLUSÃO	142

6. PERSPECTIVAS.....	144
REFERÊNCIAS	147
BIBLIOGRAFIA	178
APÊNDICE I – Manuscript draft.....	179
APÊNDICE II – Simulação de DM: RSVmF-QRC	202
APÊNDICE III – Simulação de DM: RSVmF-Q1	203
APÊNDICE IV – Simulação de DM: RSVmF-JNJ	204
APÊNDICE V – Variação da distância RSVmF-ligantes – CADEIA A	205
APÊNDICE VI – Variação da distância RSVmF-ligantes – CADEIA B	206
APÊNDICE VII – Variação da distância RSVmF-ligantes – CADEIA C	207
APÊNDICE VIII – Método MM-GBSA	208
APÊNDICE IX – Propriedades físico-químicas dos ligantes quercetina, Q1 e JNJ-2408068	209
APÊNDICE X – Primeira simulação ANN.....	210
APÊNDICE XI – Matrizes de confusão da simulação 2	211

INTRODUÇÃO

1. INTRODUÇÃO GERAL

“One is too many.” (UNICEF)

Respirar é quesito básico para sobrevivência e, por conseguinte, o sistema respiratório é também o mais comumente afetado por infecções. Dados da **Organização Mundial da Saúde (OMS)** de 2015 mostram que a pneumonia é uma das principais causas de mortalidade infantil [1]. Embora no Brasil e em outras nações em desenvolvimento, o diagnóstico da causa das infecções respiratórias seja pouco explorado, o agente viral pioneiro e desencadeador dessas infecções nesses casos pediátricos é o pneumovírus vírus sincicial respiratório (RSV – do inglês *respiratory syncytial virus*) [2-9].

O RSV infecta vias aéreas superiores e pode se estender às inferiores de crianças, adultos e idosos, sendo mais crítico no primeiro e no terceiro grupo devido ao sistema imune deficitário. O vírus pode causar desde coriza evoluindo até bronquiolite e pneumonia, o que pode levar à morte. A cada ano, estima-se que 34 milhões de crianças sejam infectadas ao redor do mundo [5]. Em 2010, houve cerca de 253 mil mortes, a maioria em países em desenvolvimento [10].

O tratamento atualmente disponibilizado adota o uso de ribavirina, um composto com efeitos potencialmente tóxicos e eficácia questionável [3,4,11-13]. A prevenção é feita pela utilização de palivizumabe, um anticorpo monoclonal (Synagis®), aplicável em casos restritos, devido aos custos intrínsecos. Sua eficácia, no entanto, também tem sido questionada [9].

Desta forma, a busca por alternativas que substituam a ribavirina, e que também reduzam os custos com a prevenção dessa infecção continua patente. Tais compostos

terapêuticos devem reunir características desejáveis, tais como baixa toxicidade, segurança na administração, eficácia e, além disso, custo reduzido.

Estudos [12,14-18] mostram que o desenvolvimento da infecção por RSV depende de sinalização celular regulada pela liberação de espécies reativas de oxigênio (EROs). Os compostos fenólicos, provenientes de metabolismo secundário vegetal, possuem capacidade neutralizante de EROS. Representam, portanto, uma alternativa para modulação do processo inflamatório resultante de infecção por RSV, e, possivelmente, para redução da carga viral.

Dentre a diversidade de compostos fenólicos se encontram os flavonoides e entre as suas várias classes, os flavonóis tipo quercetina são os mais conhecidos. A literatura relata uma miríade de atividades sobre o potencial farmacológico da quercetina, os quais incluem desde propriedade antioxidante, antitumoral, anti-inflamatória, até atividade antibacteriana [19-23] e antiviral [24-29]. No entanto, os mecanismos moleculares concernentes a tais atividades ainda não foram bem elucidados.

Visando compreender as interações químicas entre moléculas pequenas como a quercetina e proteínas-alvo, diversas ferramentas e estratégias computacionais vêm sendo propostas e utilizadas. O intuito é suplantando limitações experimentais no que concerne à observação das interações químicas a nível molecular e atômico, além de buscar de forma efetiva padrões de interesse presentes dentro desse universo químico. Assim, associadas à parte experimental essas ferramentas possibilitam um melhor entendimento do fenômeno em questão.

A proposta do presente trabalho foi investigar flavonoides inibidores da replicação do RSV através de ferramentas computacionais. Foram realizadas análises das interações moleculares da proteína de fusão (proteína F) do RSV em complexo com a quercetina, pelo uso de ferramentas da biologia estrutural – *docking* e a Dinâmica Molecular. Ao mesmo tempo, procurou-se avaliar como uma modificação estrutural na quercetina – a acetilação – poderia influenciar na capacidade de interação com essa mesma proteína. Além disso, desenvolvemos uma rede neural artificial associada ao algoritmo genético para categorização de flavonoides em “ativos” ou “inativos”, a partir

de determinados parâmetros de entrada. Buscamos otimizar a investigação por moléculas inibitórias pertencentes a essa classe por estabelecer um critério seletivo anteriormente à experimentação na bancada, o que implica em maior funcionalidade experimental, e redução de custos e tempo.

1.1 Infecções do trato respiratório

“Infância – A vida em technicolor...” (Mário Quintana)

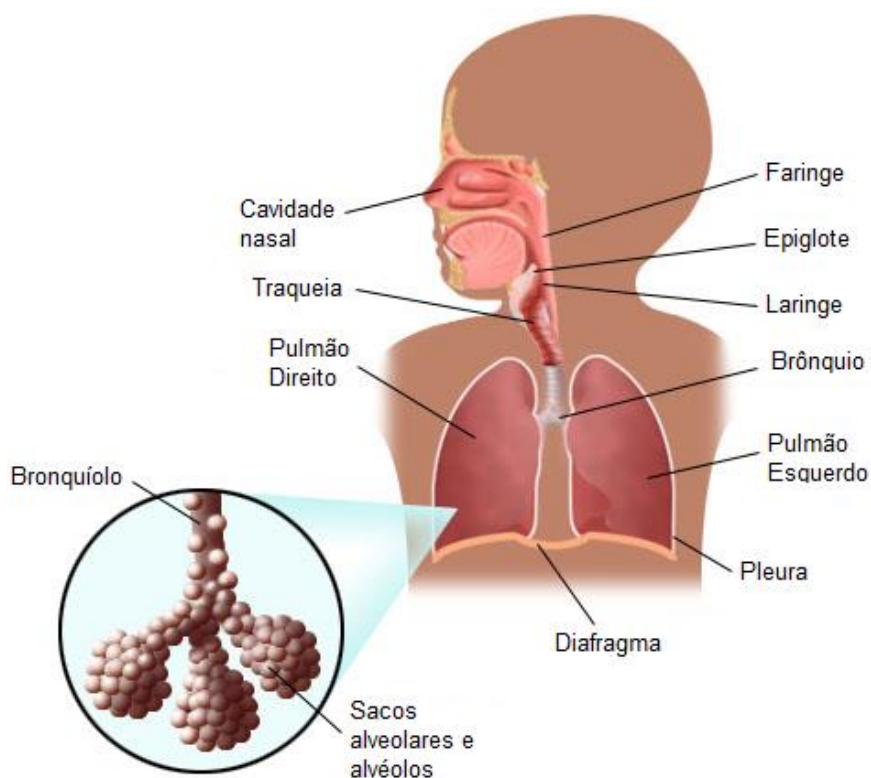
A respiração está diretamente relacionada à capacidade do nosso organismo de produzir energia e, portanto, é fundamental para nossa sobrevivência. Do latim *respirare*, composto de *re* = outra vez, e *spirare* = exalar, aspirar, inalar [30], a ação de espirar continuamente, em uma atmosfera previamente oxigenada, permitiu-nos e nos permite a permanência no planeta Terra. O radical *spir* ainda faz parte de uma gama de vocábulos presentes não só na língua portuguesa, como *espírito*, indicando, talvez, além de uma questão básica de sobrevivência, um elo transcendental da existência.

O sistema respiratório é dividido em trato respiratório superior (vias aéreas superiores) e trato respiratório inferior (vias aéreas inferiores). [Figura 1] O primeiro é formado pela cavidade nasal, faringe e laringe. O segundo é representado pela traqueia, a qual se ramifica em brônquios e bronquíolos, e pelos pulmões.

Esse sistema é revestido por tipos diferentes de epitélio, os quais atuam como barreira contra patógenos, alérgenos e quaisquer substâncias estranhas ao organismo [Figura 2]. A presença de cílios na superfície das células do sistema respiratório, bem como a produção de muco por estruturas especializadas visam dificultar a adesão de microrganismos e os degradar, tendendo, assim, a evitar as possíveis entradas de patógenos – infecção [31]. Ademais, as células do sistema respiratório também são capazes de secretar moléculas sinalizadoras – as citocinas* – as quais assinalam para o sistema de defesa – o sistema imune – de que uma possível inflamação e/ou infecção está ocorrendo ali, e de que o reparo e/ou o combate ao agente infeccioso é necessário.

“As citocinas/citoquinas formam um grupo grande de proteínas secretadas que apresentam uma variada gama de estruturas e de funções, e que regulam e coordenam muitas atividades das células do sistema imune inato e adaptativo [...] Entre suas muitas funções, as citocinas regulam crescimento e diferenciação das células imunes, ativação das funções efetoras de linfócitos e fagócitos, bem como o direcionamento das células imunes do sangue para o tecido e dentro do tecido. As citocinas que regulam a migração celular e o movimento são chamadas *quimiocinas*.” [31, tradução nossa]

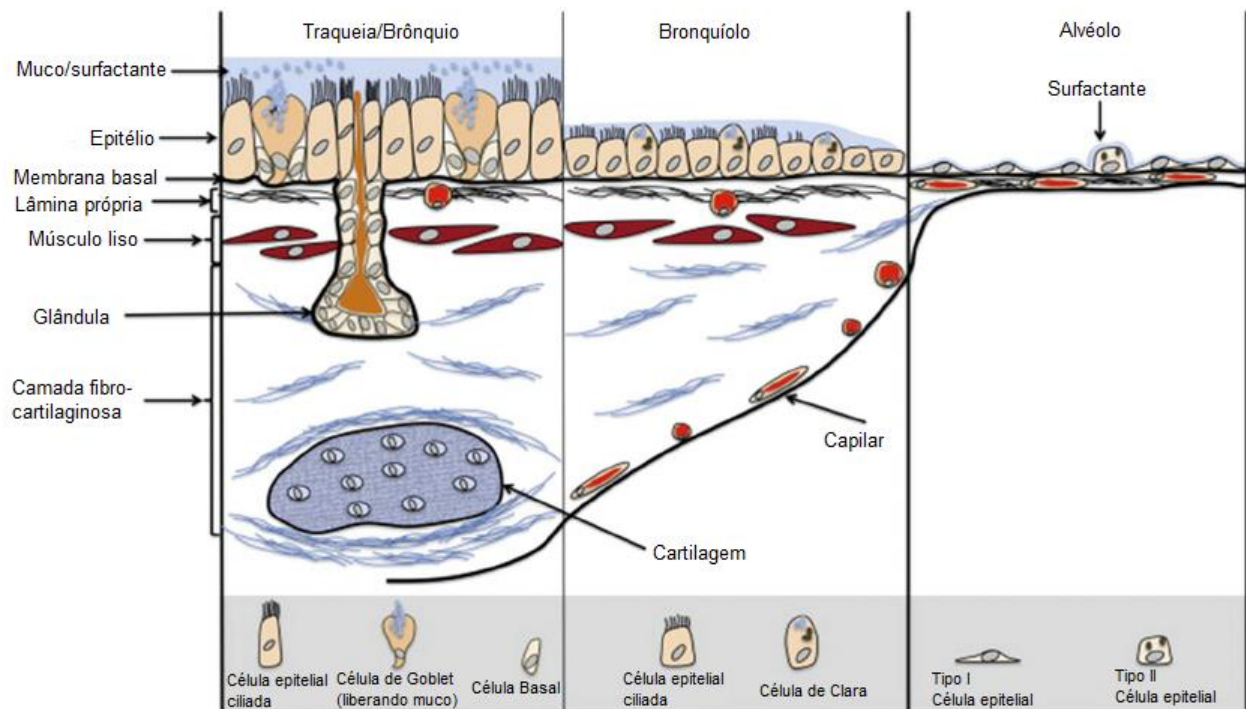
Figura 1 – O sistema respiratório



O sistema respiratório pode ser dividido em vias aéreas superiores e inferiores. As primeiras incluem a cavidade nasal, a faringe e a laringe. As segundas incluem a traqueia, ramificada em brônquios e bronquíolos, e os pulmões. Outra divisão possível é entre porção condutora (cavidade nasal, faringe, laringe, traqueia, brônquios e bronquíolos) e porção respiratória, na qual ocorrem as trocas gasosas (bronquíolos respiratórios, ductos alveolares, sacos alveolares e alvéolos).

Fonte: Adaptado [32].

Figura 2 – Tipos celulares e epitélios do sistema respiratório – traqueia/brônquios, bronquíolos e alvéolos

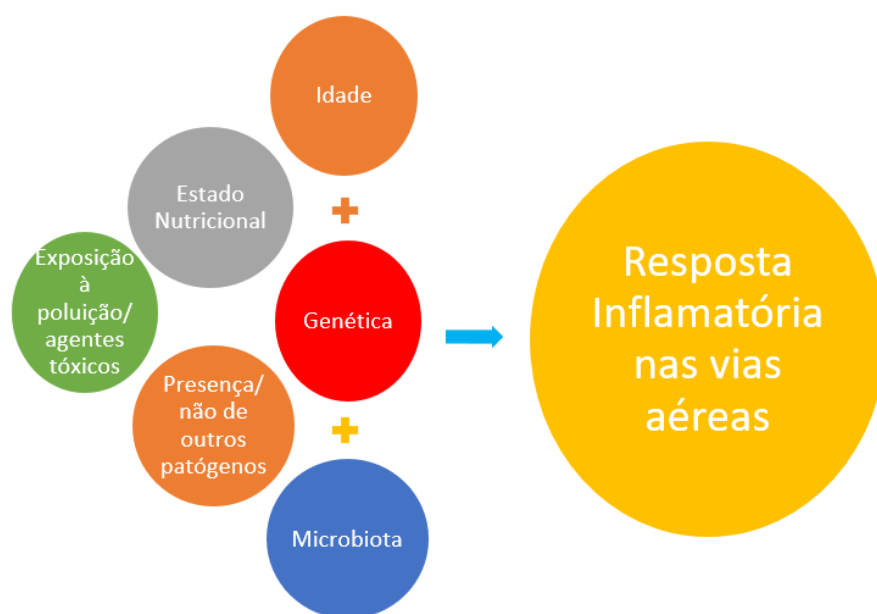


O epitélio da traqueia/brônquios é do tipo pseudo-estratificado ciliado com a presença de células de Goblet (também chamadas de células caliciformes) responsáveis pela produção de muco contendo mucinas e outras enzimas, cuja função é degradar alérgenos e alguns microrganismos. A produção de muco neste local tem papel fisiológico. Já nos bronquíolos, o epitélio é do tipo simples colunar/cúbico e ciliado com a presença de células de Clara (renomeadas células *Club* [33]), cuja função é excretar glicoproteínas, proteínas e enzimas, visando a defesa contra agentes estranhos, bem como auxiliando a evitar o colapso dos bronquíolos. Quanto aos alvéolos, há a divisão em dois tipos celulares: pneumócitos I, formando epitélio pavimentoso simples, e pneumócito II, compondo epitélio cúbico simples. A secreção de surfactantes aqui tem como função diminuir a tensão superficial e, assim, evitar o colapso do alvéolo. Fonte: Adaptado [34].

A elaboração da resposta inflamatória nas vias aéreas superiores e inferiores é dependente de diversos fatores, dentre os quais idade, estado nutricional (as condições socioeconômicas do indivíduo), genética, exposição à poluição e ao cigarro, dentre outros [2] [Figura 3]. Assim sendo, indivíduos cujo sistema imune ainda não esteja bem formado (idade), ou nos quais esse se encontre de alguma forma suprimido/deficiente (idade, estado nutricional, genética, exposição a agentes tóxicos, entre outros), apresentam

organismos potencialmente mais vulneráveis às infecções e aos danos causados por essas no sistema respiratório.

Figura 3 – Fatores intrínsecos e extrínsecos que interferem na resposta inflamatória nas vias aéreas



Diversos fatores influenciam na elaboração da resposta inflamatória nas vias aéreas. Alguns autores [revisto por [65](#),[73](#)] defendem que a idade seja o aspecto mais determinante. Em recém-nascidos e crianças até 1-2 anos, o sistema imune ainda não está maduro o suficiente, tendendo a realizar uma resposta inflamatória mais alérgica, e nos idosos, o sistema está em declínio. Os outros fatores podem atuar sinergicamente em um quadro de infecção, ora auxiliando o sistema imune (genética com menor tendência alérgica, um bom estado nutricional, a não-exposição à poluição, ausência de coinfeções e uma microbiota com menor potencial patogênico) ou o prejudicando (genética tendencialmente alérgica, má-nutrição, exposição constante a agentes tóxicos, coinfeções, microbiota potencialmente patogênica).
Fonte: o autor

As infecções respiratórias representam a forma mais comum e facilmente disseminada de infecção entre os humanos. Entre seus diversos tipos epidemiológicos, as de origem viral são notórias e frequentes. Quem não se lembra dos surtos de gripe

suína e aviária de 2009 para cá? Ou também ao longo da história, como a pandemia da gripe espanhola de 1918? Isso porque o contágio ocorre, dentre outras formas, por meio de inalação de partículas virais [35], as quais podem ser, por exemplo, provenientes de um espirro de alguém contaminado que esteja próximo. Em épocas de inverno, nas quais a tendência é a aglomeração em ambientes fechados, potencializa-se a transmissão desses vírus. Além disso, em períodos secos, como é o caso do inverno na maior parte do Brasil, a falta de umidade favorece o acúmulo de partículas e poluentes no ar além de também contribuir para o ressecamento das mucosas do trato respiratório, o que propicia a irritação nas mesmas. A transmissão facilitada pelo agrupamento em ambientes fechados, associada à irritação das mucosas devido às condições climáticas e ainda a um sistema imune não completamente formado ou deficitário, são fatores que intensificam a propagação das infecções virais, além de as agravarem.

Dentre os vírus responsáveis pelas infecções respiratórias em humanos, sobressaem adenovírus, enterovírus, **coronavírus** (CoV), **metapneumovírus** (MPV), **rinovírus** (RV), influenza, parainfluenza e o vírus sincicial respiratório (RSV) [36]. Entre os mencionados, destacam-se influenza, RV e RSV. Os três ocasionam infecções respiratórias agudas (IRAs). O primeiro é a causa mais frequente de mortes relacionadas à pneumonia nas nações desenvolvidas [36]. O segundo é o responsável pelos resfriados comuns, e também por boa parte da exacerbação dos sintomas concernentes à asma [36]. E o último, é o maior motivo de hospitalização de crianças [36], predominantemente recém-nascidos, acarretando uma alta taxa de morbimortalidade dentro desse grupo.

Apesar de todos os avanços da ciência e dos esforços atingidos no setor da saúde, principalmente com o desenvolvimento da Epidemiologia [37] e, por conseguinte, do foco no desenvolvimento de estratégias de prevenção e de cuidados em um determinado tempo e contexto histórico, as IRAs em crianças e recém-nascidos permanecem como problema de saúde pública, seja no Brasil [38-44] ou mundialmente [3-5,45-47]. Estima-se que as doenças pulmonares obstrutivas crônicas e as infecções respiratórias estarão entre as seis maiores causas de morte no ano de 2030 [48,49]. A alta taxa de morbimortalidade infantil relacionada às IRAs em países em desenvolvimento decorre,

dentre outros fatores, de agravamentos consequentes a elas, tais como infecções bacterianas secundárias, além de infecções do trato respiratório inferior (ITRI), nomeadamente pneumonia clínica (bronquiolite e pneumonia). O Brasil figura entre os 15 países com números expressivos de casos de pneumonia clínica em crianças menores de 5 anos (1,8 milhão), atingindo uma taxa estimada em 0,11 episódio/criança-ano [41]. E o estado de São Paulo aparece como líder nas taxas de óbitos infantis por IRAs em crianças menores de 5 anos, segundo dados do Ministério da Saúde [50], com uma taxa de 454 óbitos para o ano de 2011.

É na infância que se configuram grande parte das potencialidades de um indivíduo. Assim, os distúrbios que acontecem nessa fase suscitam graves consequências não só pessoais como também sociais, de forma que a saúde da criança representa uma prioridade [51], justamente por caracterizar o capital humano de uma nação. O coeficiente de mortalidade infantil além de ser um indicativo de saúde, ainda mostra a eficácia das políticas públicas de atendimento às mães e aos recém-nascidos [51,52]. A redução da mortalidade infantil, é um dos objetivos da **Organização Mundial da Saúde (OMS)** [53], por meio da implementação de políticas de saúde. No Brasil, o **Estatuto da Criança e do Adolescente (ECA)** versa sobre o compromisso de bem-estar desse grupo, pelo qual a família, o estado e a sociedade têm responsabilidades [54]. Dados da OMS de 2015 mostram que, dentre as causas mais proeminentes de mortalidade de crianças, destaca-se a pneumonia [1]. De acordo com o relatório do Fundo das Crianças das Nações Unidas (UNICEF – *United Nations Children's Fund*), uma a cada seis mortes de crianças menores de 5 anos em 2015 ocorreu devido à pneumonia [55]. Isso representa uma morte a cada 35 segundos [55]. Embora as ocorrências sejam subnotificadas, especialmente em nações em desenvolvimento, como o Brasil, o agente viral mais comumente encontrado nesses casos pediátricos é o RSV [2-9].

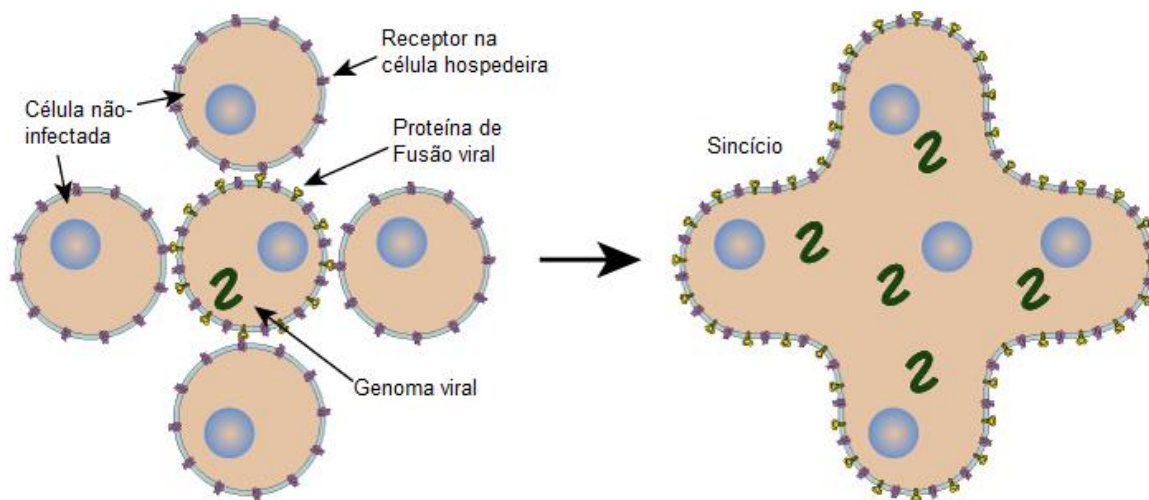
1.1.2 O vírus sincicial respiratório (RSV)

“... Porque pelo fruto se conhece a árvore...” (Mateus 12:33)

O RSV é um pneumovírus, isto é, infecta as vias respiratórias e foi recentemente reclassificado pelo Comitê Internacional de Taxonomia de Vírus (*International Committee on Taxonomy of Viruses* – ICTV) [56]. Pertence à ordem *Mononegavirale*, família *Pneumoviridae*, gênero *Orthopneumovirus* e sua renomeação para *Orthopneumovirus humano* foi proposta.

Foi descoberto em 1955 e isolado de chimpanzés de laboratório [57]. Recebeu o nome de agente da coriza do chimpanzé (*Chimpanzee Coryza Agent* – CCA) por ter sido isolado de chimpanzés que apresentavam coriza. Posteriormente, em 1956, foi identificado em um dos pesquisadores desse grupo, o qual trabalhava diretamente com os animais em questão, além de em adolescentes e em adultos, cuja sintomatologia era semelhante [57]. No entanto, a primeira suposta descrição do vírus data de 1826 [58,59], embora se acredite que ele seja bem mais antigo, e apenas não tenha sido descrito anteriormente devido às limitações técnicas concernentes à diferenciação dos agentes causadores de infecções respiratórias [58]. A patologia foi relatada em 1939 [60] e a primeira exposição clínica da doença foi feita nos anos 40 [61,62]. A presença de RSV no organismo gera características peculiares, tais como a fusão das células, produzindo células gigantes multinucleadas – o sincício [63] – e, por isso, foi denominado vírus sincicial respiratório [Figura 4 e Figura 5].

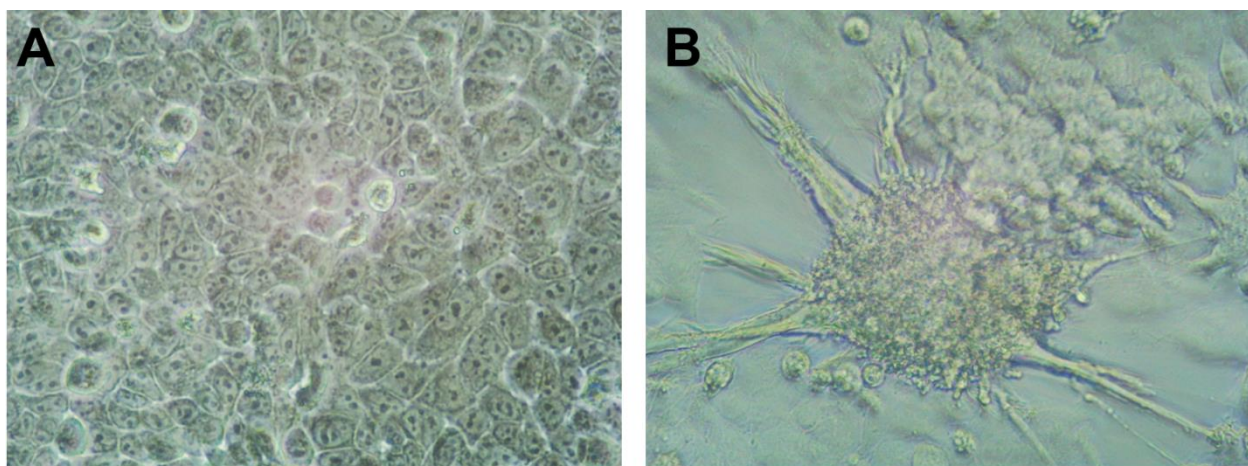
Figura 4 – Formação de sincícios pelo RSV



A fusão de células infectadas com células vizinhas, através da expressão de proteínas de fusão (proteína F) do vírus na superfície da célula hospedeira e da interação dessa proteína com receptores celulares, leva à formação de aglomerados multicelulares, os sincícios.

Fonte: Adaptado [64].

Figura 5 – Formação de sincícios *in vitro*



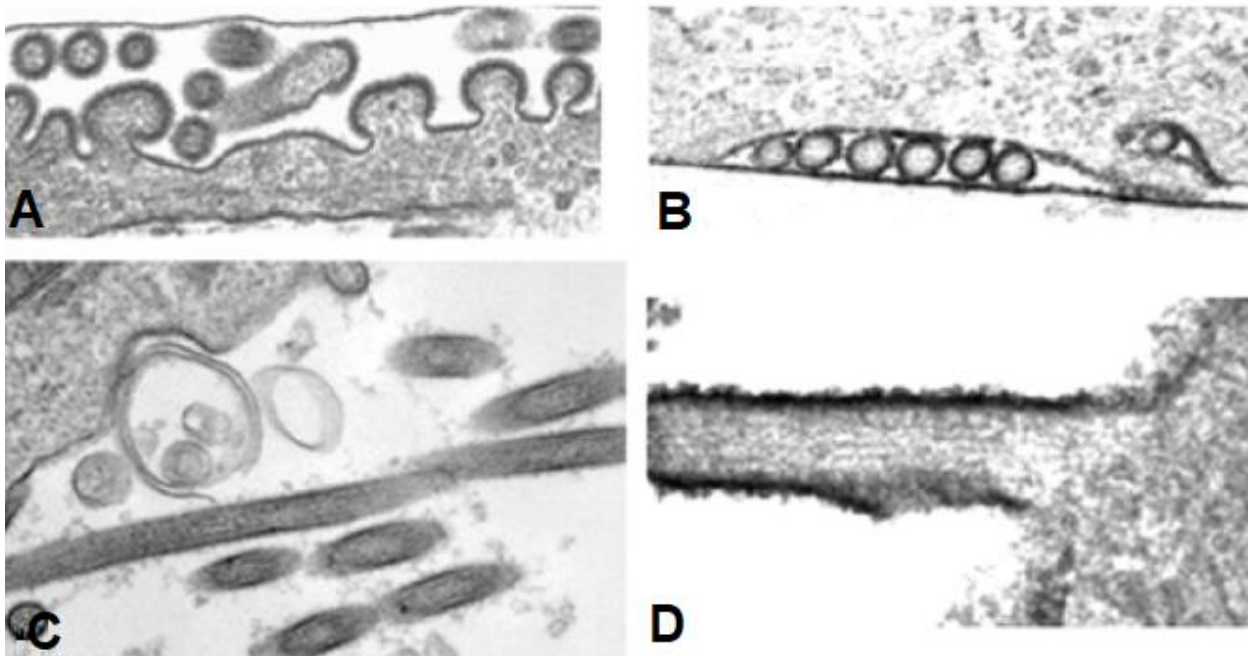
Células epiteliais humanas (5×10^4) do tipo 2 (HEP-2), derivadas de carcinoma de laringe, cultivadas em meio mínimo essencial de Eagle (MEM) 10% soro fetal bovino, anteriores à infecção por RSV (A) e posteriormente à infecção (B) com multiplicidade de infecção (MOI) 0.1 (isto é, uma proporção de uma partícula viral para cada 10 células). Aumento de 20x.

Fonte: Fotos cedidas por MSc. Larissa Figueiredo Alves Diniz (2017).

Ele infecta as células epiteliais das vias aéreas superiores, podendo se estender às inferiores de crianças, adultos e idosos, sendo mais crítico entre crianças e idosos devido ao sistema imune deficitário [revisto por [65](#)]. Além disso, existem os grupos de risco, isto é, indivíduos com maior predisposição a desenvolverem uma infecção mais exarcebada, devido à prematuridade e baixo peso, a má-formações pulmonares e cardíacas, a doenças de etiologia genética e cromossômica, como portadores de síndrome de Down, ou também à imunossupressão que afeta indivíduos soropositivos [[66-70](#)] e transplantados [revisto por [71](#)]. Embora não sejam muito comuns, existem registros de infecção ou de material genético viral em outros locais que não somente no aparelho respiratório, tais como fluido cerebral, músculo cardíaco e fígado [[2,72](#)].

É um vírus envelopado, com um diâmetro de aproximadamente 150 nm [[73](#)]. Suas partículas infectantes, os vírions, podem ser encontrados tanto na forma semi-esférica, como também na forma de filamentos, sendo, portanto, um vírus pleomórfico [[Figura 6](#)]. Ambas as formas são infectantes [[74](#)].

Figura 6 – Pleomorfismo do RSV



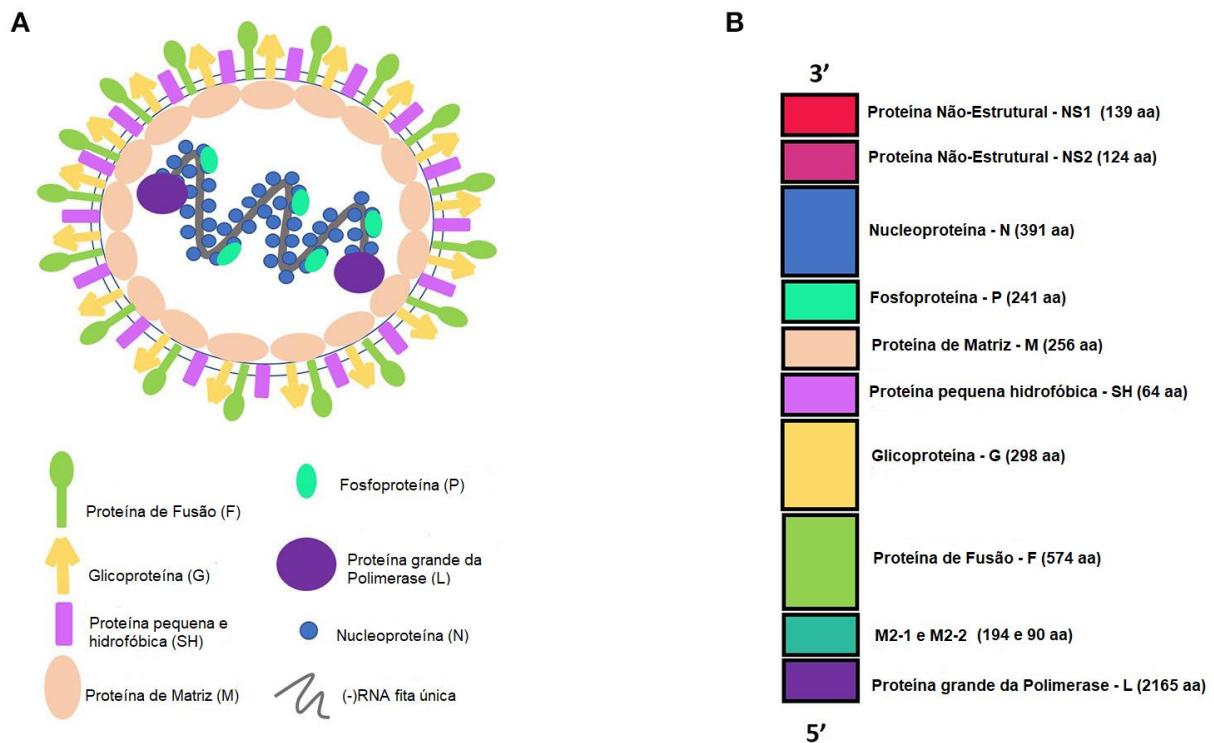
As partículas infectantes do RSV, os vírions, são pleomórficos e suas formas podem variar desde semi-esféricas (A e B) até a forma mais comumente encontrada de filamentos (C e D).

Fonte: Adaptado [74].

Seu material genético consiste de uma fita de RNA genômico não-segmentada de sentido negativo, isto é, sentido 3'-5', com aproximadamente 15.2 quilo-bases (kb – *kilo-base pair*, logo, 15.222 nucleotídeos), o qual codifica onze proteínas [Figura 7], nomeadamente NS1, NS2, N, P, M, SH, G, F, M2 (subdividida em M2-1 e M2-2) e L. Suas proteínas se dividem em não-estruturais (aquelas que não são parte da partícula viral) – NS1 e NS2 – e estruturais – todas as demais, as quais compõem a partícula viral: N, P, M, SH, G, F, M2-1, M2-2 e L. As cepas virais são classificadas em dois grandes grupos antígenicos A e B, os quais diferem principalmente quanto à variabilidade genética da proteína G [75-77]. Por ser um vírus cujo genoma é não-segmentado, e, portanto, não realiza mudança antígenica (*antigenic shift*, em inglês) como o vírus influenza, sua taxa

de variabilidade é devida, principalmente, ao erro intrínseco à leitura de sua RNA polimerase, a qual não conta com mecanismos de correção e edição [77].

Figura 7 – Representação esquemática do RSV, de suas proteínas e de seu genoma



A - O RSV possui três glicoproteínas de superfície responsáveis pela mediação do processo de reconhecimento celular com o hospedeiro e posterior adesão e fusão (respectivamente, proteínas G e F, auxiliadas pela proteína SH). A proteína M age no processo de empacotamento de novas partículas virais, enquanto que as proteínas L, N, P compõem o complexo ribonucleoprotéico (RNP), e juntamente com as proteínas M2-1 e M2-2 atuam no processo de replicação e transcrição do material genético viral. As proteínas não-estruturais NS1 e NS2 desempenham papel importante nos mecanismos de evasão do sistema imune. **B** - O material genético do RSV consiste de uma fita de RNA simples não-segmentado e com polaridade negativa. Esse genoma é codificado em 11 proteínas, sendo que a proteína M2 se subdivide em M2-1 e M2-2.

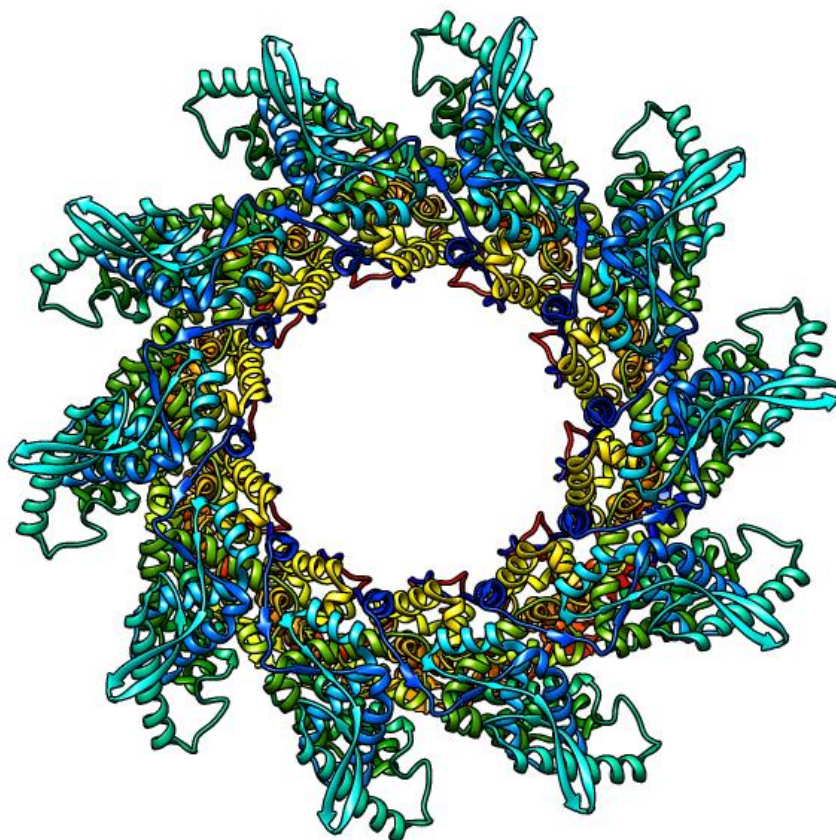
Fonte: Adaptado [78].

Pela ordem de leitura do RNA (3'–5') [Figura 7], os primeiros genes a serem codificados são os das proteínas não-estruturais (NS – *nonstructural*) NS1 e NS2, as

quais são proteínas pequenas com 139 e 124 aminoácidos, respectivamente. Elas podem formar complexos homoméricos (NS1-NS1, NS2-NS2), como também podem se associar em um complexo heteromérico (NS1-NS2). Podem atuar sinergisticamente [79,80], com foco no bloqueio da resposta imune inata da célula epitelial através da inibição da secreção de interferon dos tipos I e III (α e β , λ). [79,80]

Na sequência, vem a nucleoproteína N, a qual é composta de 391 aminoácidos e possui uma estrutura helicoidal [Figura 8]. É responsável por proteger (encapsular) o material genético viral resultante dos processos de replicação (o antigenoma, ou seja, o genoma de polaridade positiva – (+)RNA) e transcrição (síntese de RNA mensageiro – mRNA – a partir do (+)RNA), a fim de evitar sua degradação por nucleases [81,82].

Figura 8 – Estrutura resolvida da Nucleoproteína – Proteína N – do RSV



A nucleoproteína do RSV apresenta estrutura helicoidal e é responsável por encapsular o material genético resultante dos processos de replicação e transcrição. Observa-se a visão superior da proteína. A estrutura foi escolhida por critérios de completude, e não de resolução.

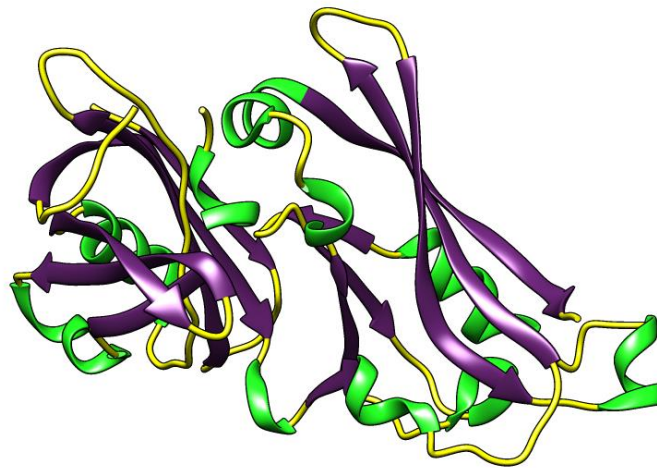
Fonte: Protein Data Bank [83] (Código de Identificação PDB: 2WJ8). Edições de imagem feitas no programa Chimera [84].

Consequente à nucleoproteína, vem a fosfoproteína P, cuja estrutura contém 241 aminoácidos. Ela é fosforilada em diversas serinas e treoninas, sendo que tal processo é realizado em diferentes taxas entre as diferentes cepas, com papel importante na regulação da replicação e do brotamento do vírus [revisto por 85]. Interage tanto com a proteína N, quanto com a proteína L, formando o complexo ribonucleoprotéico (RNP), e

também interage com a proteína M. Ela atua como cofator e estabilizador do complexo da polimerase [81].

A próxima proteína é proteína de matriz M, que é uma proteína não-glicosilada [Figura 9]. Forma uma camada protéica embaixo do envelope viral, envolvendo o nucleocapsídeo e interagindo com a membrana viral [78]. Desenvolve papel fundamental no empacotamento do vírus bem como em sua liberação. Sua ausência prejudica a formação dos vírions [86]. Possui 256 aminoácidos e sua dimerização é fundamental para exercer sua função [87].

Figura 9 – Estrutura resolvida da Proteína de Matriz – Proteína M – do RSV



A proteína de Matriz é fundamental para o empacotamento e liberação do vírus. Sua forma ativa é um dímero e é uma proteína não-glicosilada. Observa-se a visão lateral da proteína. A estrutura foi escolhida por critérios de completude, e não de resolução.

Fonte: Protein Data Bank [83] (Código de Identificação PDB: 2VQP). Edições de imagem feitas no programa Chimera [84].

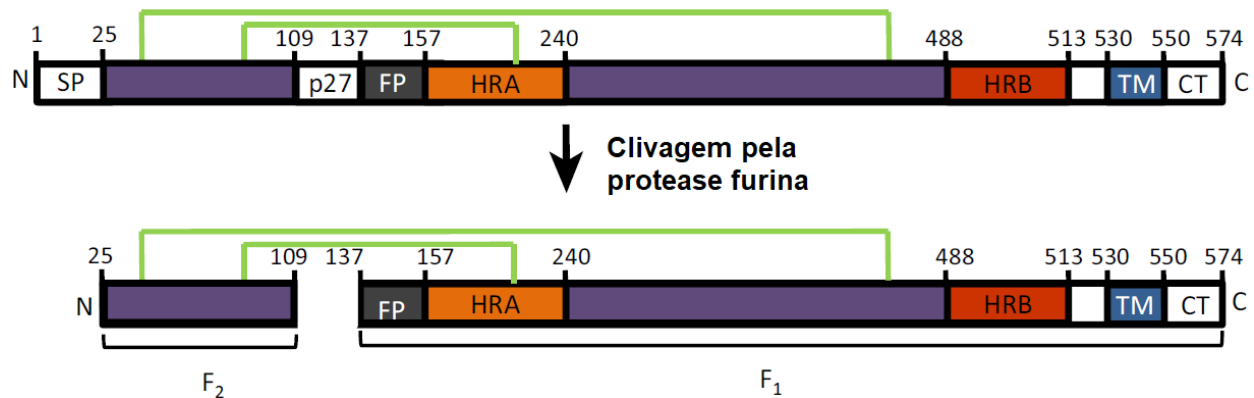
As proteínas seguintes são as glicoproteínas transmembrânicas SH, G e F. Elas desenvolvem papel primordial na adesão da partícula viral e na fusão de sua membrana à membrana da célula que será infectada. A proteína SH (SH – *Small Hydrophobic*, proteína “hidrofóbica pequena”) apresenta 64 aminoácidos, no subtipo A, e 65

aminoácidos, no subtipo B [88], e é uma proteína transmembrânica do tipo II. Isto significa, basicamente, que seu domínio C-terminal se localiza externamente à membrana celular viral, enquanto que seu domínio N-terminal é citoplasmático [89]. Sua função parece estar relacionada à formação de poros, com características de canais iônicos, nas células infectadas, as quais, provavelmente, sofrem com alteração da permeabilidade da membrana [90,91].

A glicoproteína de membrana G é responsável por mediar o processo de adesão do RSV à célula hospedeira. É uma proteína transmembrânica do tipo II, assim como a proteína SH. Contém em sua estrutura de 289 a 299 aminoácidos, dependendo da cepa [88], sendo extremamente glicosilada [revisto por 88,92]. Aproximadamente 60% do seu peso molecular é carboidrato [88]. A proteína G é expressa em duas formas: uma forma transmembrânica (mG) e uma forma secretada (sG), as quais diferem quanto à presença/ausência dos domínios N-terminal e transmembrânico [93], sendo que a segunda não os contém. A forma sG corresponde a quase 80% do total de proteína G liberada durante os estágios iniciais da infecção por RSV [94], e auxilia a bloquear a opsonização e neutralização do vírus por anticorpos anti-G específicos produzidos pelo hospedeiro [95]. A parte externa da proteína G (seu ectodomínio) contém dois domínios característicos de mucinas, ou seja, domínios altamente glicosilados. Entre eles, existe uma região altamente conservada de 13 aminoácidos, a qual contém tanto um CX3C *motif* (sequência conservada de aminoácidos, a qual interage com o receptor de quimiocina – CX3CR1 – na célula epitelial) e uma região rica em cisteína (GCRR), a última formada por quatro resíduos de cisteínas ligados por duas pontes dissulfeto [96,97]. Tanto o *motif* CX3C quanto a GCRR são importantes para modulação do sistema imune do hospedeiro, sendo CX3C *motif* um dos fatores essenciais para adesão da proteína G nas células epiteliais [98-102]. A proteína G também interage com os proteoglicanos de heparan-sulfato (HSPG – **H**eparan-**S**ulfate **P**roteo**G**lycans) comumente encontrados nas superfícies celulares, embora isso venha sendo contestado como artefato de cultura *in-vitro* [73, 103].

A glicoproteína de membrana F – proteína de fusão – é responsável pela fusão da membrana viral à célula epitelial do hospedeiro. Ela é uma glicoproteína do tipo I, isto é, seu domínio N-terminal se localiza externamente à membrana celular viral, enquanto que seu domínio C-terminal é citoplasmático [89]. Possui 574 aminoácidos [104]. A proteína F existe em dois estados de conformação principais: pré e pós-fusão. Na superfície viral, as duas possibilidades conformacionais são possíveis de serem encontradas [105]. Essencialmente, considerando o ciclo celular no hospedeiro, essa proteína é sintetizada como um precursor F0 na forma monomérica. Posteriormente à passagem pelo aparelho Golgiense, os monômeros são clivados após os aminoácidos 109 e 136 [106,107], por uma protease celular do hospedeiro chamada furina [108,109], resultando em três polipeptídeos. Os produtos finais da clivagem, ou seja, as porções N e C (-terminais) dão origem às subunidades F2 e F1, as quais são interligadas por duas pontes dissulfeto [110,111] [Figura 10, linhas em verde]. O trímero formado é exposto na superfície do vírion. No entanto, ele é meta-estável, ou seja, ele não está na conformação termodinâmica mais equilibrada, sendo que nessa configuração é denominado RSV-F pré-fusão [Figura 11A]. Por um estímulo ainda desconhecido [112], a proteína sofre uma série de alterações conformacionais que culminam em seu estado de pós-fusão [Figura 11B]. A interação de RSV-F com alguns receptores celulares tem sido sugerida, tais como TLR4, ICAM-1 e nucleolinas [113-117]. A proteína F é o principal foco de desenvolvimento de estratégias farmacêuticas de prevenção e de tratamento contra infecções por RSV, por ser essencial à infecção viral [118] e, além disso, por ser a principal proteína responsável pela formação de sincícios, característicos dessa infecção [Figuras 4 e 5].

Figura 10 – Proteína F: sua sequência e modificações

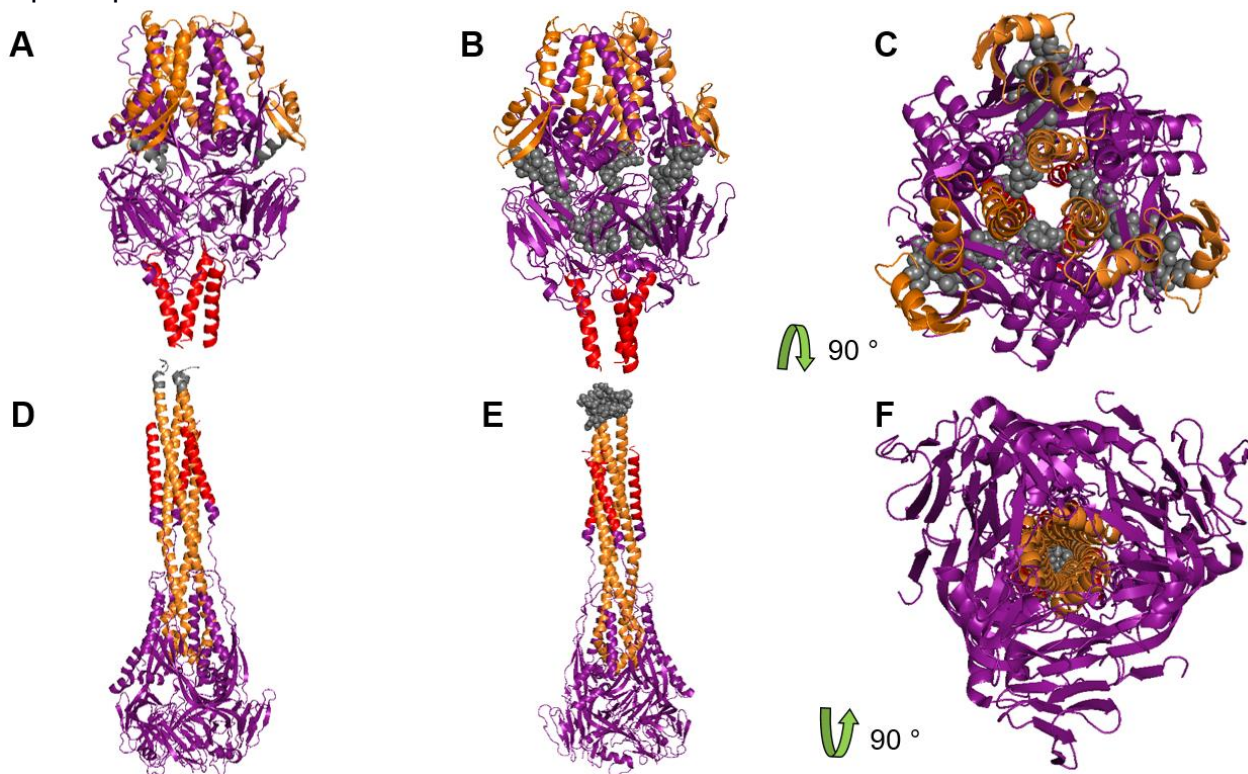


A proteína de Fusão do RSV. Representação esquemática da estrutura primária de um monômero de RSV F, antes (A) e depois da clivagem pela protease furina (B).

Legenda: HR (Região de repetição de padrão *Heptad*, isto é, sequência de 7 aminoácidos, nomeados genericamente a b c d e f g, em que “a” e “d” são geralmente hidrofóbicos e “b” “c” e “f” tendem a apresentar cargas. A estrutura formada, nesse caso, é um *coiled coil* [119]) A e B são indicadas em laranja e vermelho, respectivamente. As pontes dissulfeto são representadas pelas linhas verdes. SP - peptídeo sinal; p27 - peptídeo 27; FP - peptídeo de fusão; HRA e HRB - regiões de repetição *heptad* A e B; TM - domínio transmembrana; CT - cauda citoplasmática.

Fonte: Adaptado [112]

Figura 11 – Estruturas resolvidas da Proteína de Fusão – Proteína F – em seus estados de pré e pós fusão



A proteína F do RSV possui dois estados principais de conformação: pré-fusão, um estado meta-estável (A), (B) e (C) e pós-fusão, um estado estável (D), (E) e (F).

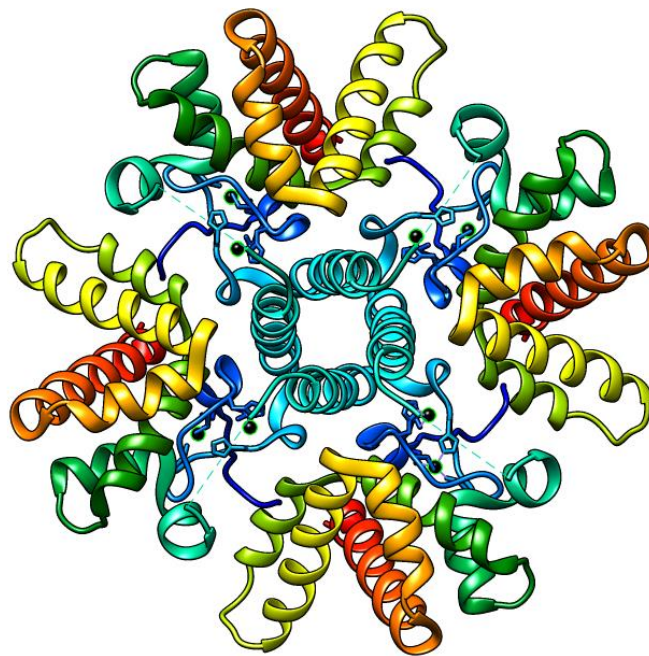
Em laranja, a região HRA, em vermelho a região HRB, e em cinza, realçado no formato de bolas o peptídeo de fusão. Observam-se a visão lateral da proteína em (A), (B), (D), e (E) e superior em (C) e (F). As estruturas foram escolhidas por critérios de completude, e não de resolução.

Fonte: Protein Data Bank [83] (Códigos de Identificação PDB: 4JHW (A) e 3RRR (B)). Edições das imagens feitas no programa Chimera [84].

As últimas proteínas a serem lidas são as proteínas M2-1, M2-2 e a proteína L. As proteínas M2-1 e M2-2 participam e regulam os processos de replicação e transcrição do RSV. A M2-1, uma fosfoproteína de 190 aminoácidos [Figura 12], regula a transcrição dos mRNA em conjunto com a proteína P, atuando como um fator de processividade e antiparada, e contribuindo, desta forma, para o trabalho de leitura da proteína L [120]. Para exercer sua função antiparada, a M2-1 necessita estar em sua forma tetramérica

[121]. Outra importante função é que ela auxilia na interação da proteína M com o nucleocapsídeo (formado pelas proteínas L, N e P) para a liberação da partícula viral [121]. Já a M2-2, uma proteína de 94 aminoácidos, medeia a transição da transcrição para replicação, regulando negativamente a síntese de RNA (replicação) [122].

Figura 12 – Estrutura resolvida da Proteína M2-1 do RSV



A fosfoproteína M2-1 atua na transcrição do vírus como um fator de processividade e antiparada. Para exercer sua função antiparada, a M2-1 deve estar em sua forma tetramérica. Observa-se a visão superior da proteína. A estrutura foi escolhida por critérios de completude, e não de resolução.

Fonte: Protein Data Bank [83] (Código de Identidade PDB: 4C3D). Edições de imagem feitas no programa Chimera [84].

A última e a maior proteína do RSV a ser lida é a proteína L (L – *Large*, “grande”). Ela é composta de 2165 aminoácidos, e atua como a maior subunidade catalítica do

complexo RNA-polimerase-RNA-dependente (RpRd). Sua principal função é na replicação e transcrição do genoma viral, reguladas pelo complexo RNP [120].

O conhecimento das estruturas protéicas do RSV auxilia na elaboração de novas estratégias de controle e prevenção da patologia, o que é urgente. Estima-se que a taxa de episódios de infecção por RSV em 2005 em todo mundo e em crianças menores de 5 anos foi de quase 34 milhões, sendo que dessas, 3,4 milhões requeriram hospitalizações [5]. Considerando-se todas as idades, a taxa sobe para 64 milhões. Em 2010, houve cerca de 253 mil mortes, a maioria em países em desenvolvimento [10]. Nos Estados Unidos, de 11 a 17 mil adultos morrem devido à infecção por RSV anualmente [123]. Além da problemática de saúde pública a nível global, o RSV também representa um peso em termos de gastos com as taxas de admissões hospitalares e internações, e de perda de produtividade e estresse por parte dos pais e cuidadores. Os gastos anuais com internações relacionadas ao RSV em crianças menores de 5 anos foram estimados em £54 milhões no Reino Unido [124]. Um estudo publicado por pesquisadores do Canadá, estimou em \$ 736,69 (dólares canadenses) os custos por internação, sendo que o tempo de permanência no hospital foi, em média, de 7 dias [125]. Esses gastos tendem a ser maiores quando considerados os grupos de risco.

Para além da infecção por RSV estritamente, existe ainda o problema das consequências da infecção *per si*. As IRAs virais no trato respiratório inferior de neonatos e crianças tendem a ocasionar sequelas graves, devido, principalmente, ao remodelamento das frágeis vias aéreas em resposta à presença do patógeno. Uma das consequências, que vem sendo relacionada à infecção por RSV na tenra idade é o desenvolvimento de asma [126-129], a qual também é uma problemática de saúde pública a nível global [revisto por 130]. No Brasil, no período de 2003 a 2013, ela foi responsável por 38% das internações hospitalares [131].

Portanto, prevenir infecções por RSV ou elaborar estratégias mais eficientes de tratamento, significa não só fazer jus ao compromisso de bem-estar das crianças, como também reduzir gastos públicos e particulares concernentes ao tratamento do estado patológico causado pelo vírus. Além disso, diminuir a possibilidade de desenvolvimento

de doenças respiratórias crônicas como a asma, desonera o sistema público de saúde. A elaboração de novas estratégias perspassa o entendimento não só da estrutura do vírus, como também da interação com o hospedeiro por meio de seu ciclo replicativo, ou seja, de seu ciclo de multiplicação.

1.1.2.1 Ciclo do RSV

“To be or not to be, that is the question.” (William Shakespeare)

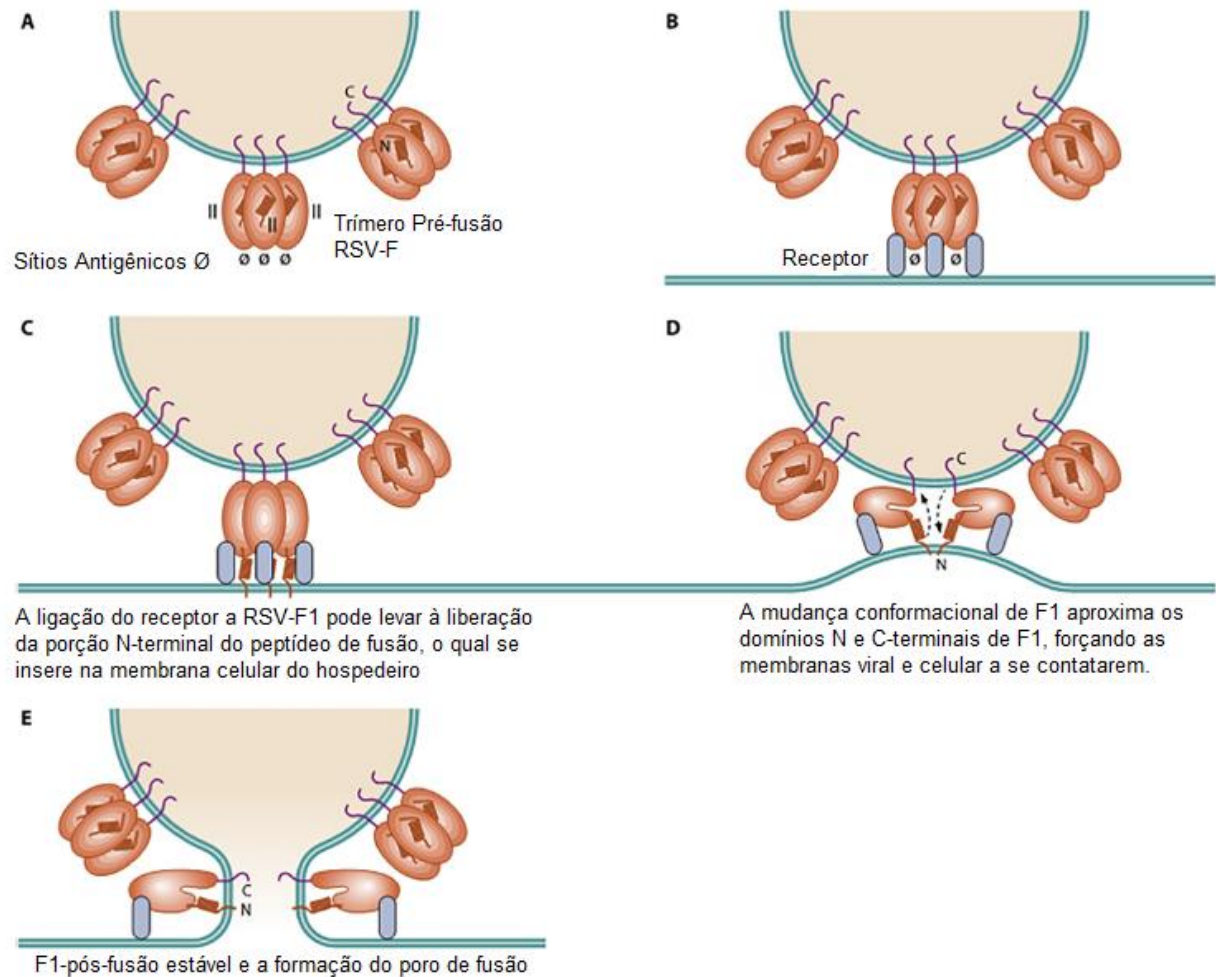
Vírus são parasitas intracelulares obrigatórios. Se por um lado eles existem como entidades biológicas por apresentarem uma estrutura organizada, embora acelular, e um material genético que codifica para proteína, por outro, no entanto, eles dependem de uma maquinaria celular para executar todos os seus processos metabólicos. Ou seja, eles só “são” dentro de uma célula. A dicotomia do “ser ou não ser” abrange do micro ao macro. E o RSV não foge a ela.

O primeiro passo em direção à infecção de uma célula é superar as barreiras químicas e físicas que se interpõem no caminho. Não é fácil. No caso do RSV, as células a serem “conquistadas” (ou talvez “invadidas”, a depender do ponto de vista) são as células epiteliais do aparelho respiratório superior e inferior. O RSV tem preferência por infectar a superfície apical dessas células [132], ou seja, o topo delas. Para obter sucesso nessa infecção, ele precisa primeiramente resistir à presença de substâncias tais como muco, surfactantes e peptídeos antimicrobianos, os quais são produzidos por células especializadas no tecido epitelial como mecanismo de lubrificação e de defesa [33,39, [Figura 2](#)], não só contra o RSV, mas também contra quaisquer outros patógenos. Além dessas substâncias, existe ainda o movimento dos cílios, o qual tende a dificultar a adesão do vírus às células, como também a favorecer o transporte de microrganismos, eventualmente presos no muco, para o ambiente ácido do estômago.

Se superadas as barreiras químicas e físicas, o RSV deve aderir à célula, para então poder adentrá-la e iniciar seu ciclo replicativo. Portanto, é necessário haver uma interação entre suas proteínas de superfície e os receptores celulares. Ademais, é fundamental que a célula seja susceptível e permissiva à infecção, isto é, seja infectada e ofereça as condições para que o vírus se multiplique. Isso é chamado de tropismo [133]. As proteínas do RSV responsáveis pelo processo de adesão e fusão à célula epitelial são

as glicoproteínas de membrana G e F. A proteína G medeia o processo de adesão, pela interação com heparan-sulfatos e o receptor de quimiocina CXCR3 [98-102]. E a proteína F realiza o processo de fusão da membrana do envelope viral com a membrana celular, por meio de alterações conformacionais na sua própria estrutura, como também por estímulos desencadeados pela interação com receptores celulares tais como TLR-4, nucleolina e ICAM-1 [113-117]. Um modelo de fusão bem elucidativo, a partir da interação da proteína F com o receptor nucleolina foi proposto por Griffiths e colegas [73] [Figura 13]:

Figura 13 – O processo de fusão entre o envelope do RSV e a membrana celular



O envelope do RSV contém várias RSV-F proeminentes, ancoradas via seus domínios transmembranas. (A) RSV-F no estado de pré-fusão, com a maior parte dos epítomos de neutralização do peptídeo de fusão localizados na região N-terminal. O maior sítio antigênico Ø existe apenas na conformação de pré-fusão e é perdido após a fusão. A interação entre o trímero de RSV-F e o receptor pode causar uma mudança dramática na conformação de RSV-F pré-fusão (B) o que leva à inserção do peptídeo de fusão na membrana celular da célula do hospedeiro (C) forçando o contato entre o envelope viral e a membrana celular (D). Apesar de apenas dois monômeros de RSV-F serem simplismente mostrados aqui, é necessária a força combinada de múltiplas mudanças de conformação de RSV-F para superar a barreira termodinâmica em mesclar as membranas, para assim estabelecer um poro de fusão que possibilite a liberação do nucleocapsídeo viral no citoplasma da célula hospedeira.

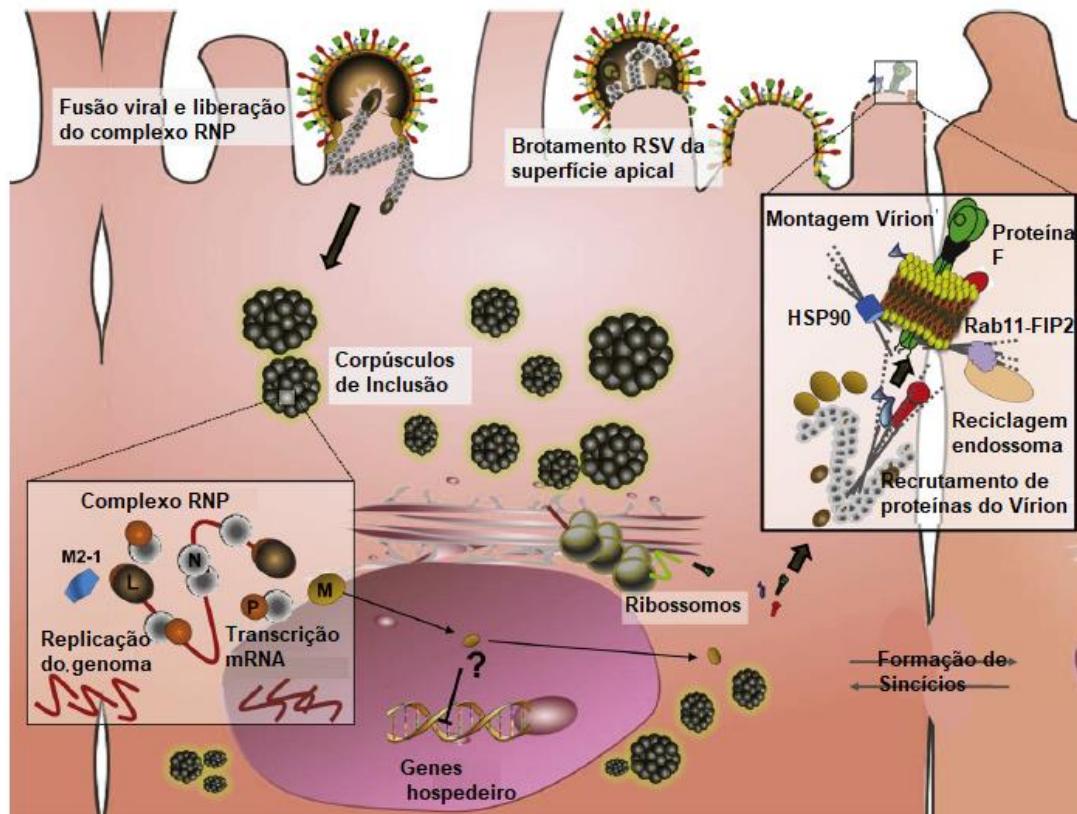
Fonte: Adaptado [73].

Além da fusão, o RSV também é capaz de adentrar a célula por outro mecanismo chamado macropinocitose, embora ainda não seja bem estabelecido o papel dos receptores e de quais deles atuem nesse sistema [134].

Após o processo de fusão, ou de macropinocitose, o nucleocapsídeo viral é liberado no citoplasma celular e a replicação deve ser iniciada. No caso do RSV, os sítios de replicação viral são conhecidos com corpúsculos de inclusão (CI) [73]. Embora o papel dos corpúsculos de inclusão não seja completamente elucidado, eles têm participação na síntese de RNA viral e na modulação das respostas imunes do hospedeiro [73]. A formação de CI em células infectadas se inicia quase que ao mesmo tempo em que se começa a síntese de proteínas virais (ao redor de 6 horas após a infecção), e seu tamanho aumenta gradativamente conforme o processo de replicação continua [135].

O RNA viral liberado no citoplasma é então coberto pela proteína N, de forma a evitar seu reconhecimento por nucleases e, assim, impedir sua degradação [81,82]. O posicionamento da proteína N sobre o RNA viral possibilita o trabalho do complexo RpRd (proteína L acessoriada pelas proteínas P, N, M2-1 e M2-2), o que constitui o centro do sistema de replicação do RSV [120]. A relação complexa de interação entre as proteínas do complexo RpRd resultará na alternância entre os processos de replicação do material genético viral ((-)RNA->(+)RNA) e transcrição do mesmo, produzindo assim suas proteínas e possibilitando o empacotamento e a geração de novos vírions [Figura 14]. Apenas as proteínas N, P, M e F são estritamente essenciais para a formação de novas partículas virais [136].

Figura 14 – Síntese do ciclo replicativo do RSV



O modelo de replicação do RSV em células epiteliais das vias aéreas (AEC – *airway epithelial cells*). 1- O RSV se liga e entra na célula epitelial, como descrito na figura 13. 2- O complexo RNP se dissocia do restante da estrutura de revestimento viral e é liberado no citoplasma. 3- Uma vez dentro do citoplasma, o complexo RNP inicia o processo de replicação-transcrição do genoma viral, o qual ocorre principalmente dentro dos corpúsculos de inclusão. As principais proteínas atuantes no complexo são as proteínas L, N e P, acessoradas pelas proteínas M2-1 e M. 4- Nos estágios iniciais da infecção, a proteína M migra para o núcleo, possivelmente para bloquear a transcrição de genes do hospedeiro. Nas etapas finais da infecção, a proteína M se localiza novamente no citoplasma através de um sinal de exportação núcleo dependente de Crm-1. 5- A proteína F se desloca aos lipídios de superfície para mediar a formação de filamentos para os vírions que sairão na superfície da célula hospedeira. Ali, a proteína F recruta outras proteínas virais, tais como as proteínas G, SH, N, P e M para formar um complexo na superfície das células infectadas e unir as proteínas do RSV. No vírion, a proteína M se liga a proteína M2-1 no complexo RNP, recrutando essa estrutura para a membrana celular a fim de possibilitar a montagem do vírion. Nesse estágio, a ação da actina facilita a montagem e liberação do vírion. Além disso, a proteína de choque térmico (HSP90 – *heat shock protein*) liga a actina às proteínas virais. Finalmente, o brotamento do vírion ocorre na superfície apical da célula AEC através do endossoma de reciclagem, o que envolve a participação da proteína Rab11-FIP-2.

Fonte: Adaptado [137].

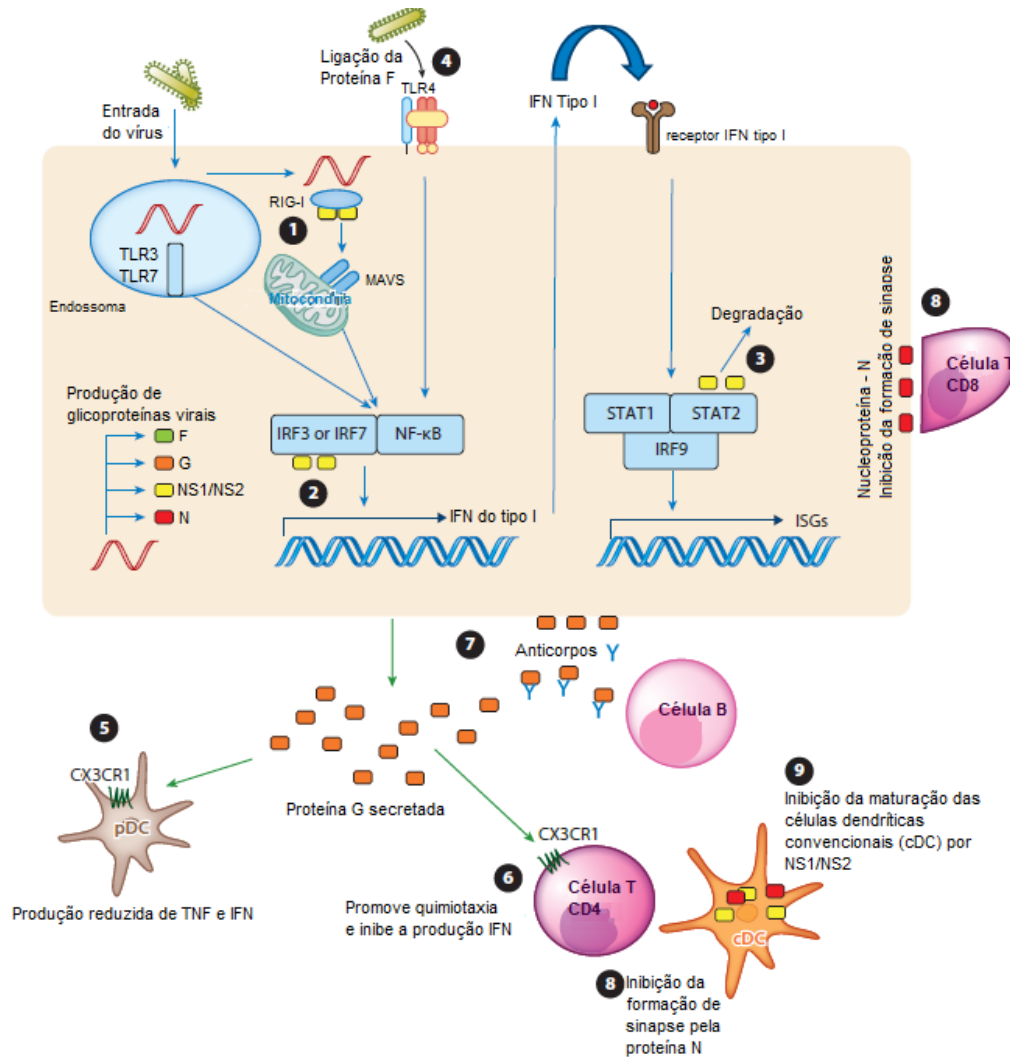
Para ter sucesso na infecção, o RSV necessita driblar os mais variados mecanismos de defesa do organismo. Assim que ocorre a interação das proteínas G e F com a superfície celular, seja em células epiteliais, como também em células dendríticas e macrófagos residentes no tecido epitelial, as partículas virais são reconhecidas por receptores específicos. Esses receptores específicos de sinalização incluem receptores de reconhecimento padrão (PRRs), receptores do tipo Toll (TLRs), receptores de gene induzíveis pelo ácido retinóico (RIG-1) e receptores do tipo Nod (NLRs) [138]. Eles promovem, então, o desencadeamento de uma cascata de sinalização na célula infectada, a qual impulsionará um estado antiviral nas células vizinhas ainda não-infectadas, e recrutará células de defesa para combaterem o vírus.

O reconhecimento das partículas virais leva à produção de várias quimiocinas e citocinas pró-inflamatórias, tais como IP-10, MCP-1, MIP-1 α , MIP-1 β , RANTES, IL-1 β , IL-6, IL-8 e TNF- α [139,140], o que estimula o recrutamento de células do sistema imune inato tais como monócitos, macrófagos, células NK, neutrófilos e células dendríticas. Além disso, a célula infectada também passa a produzir interferons (IFN) do tipo I (α e β) e liberá-los no meio extracelular. Isso induz as células vizinhas não-infectadas ainda, a transcreverem genes estimulados por interferon (ISG – *interferon stimulated genes*), proporcionando a essas um estado antiviral característico, o que já as prepara para uma eventual infecção.

Ademais, as células do sistema inato, interagem com as células do sistema adaptativo – linfócitos T e B – de forma a tentar otimizar a defesa do organismo contra o patógeno, como também a gerar memória a partir da produção de anticorpos. Entretanto, a produção de anticorpos contra RSV (IgA e IgG) é marcadamente decrescente com o passar do tempo. Após um ano de infecção, os anticorpos IgG anti-RSV não foram detectados em 25-50% em crianças pequenas, e apresentaram uma taxa de detecção de apenas 25% em quase 75% dos adultos [141-143]. Isso evidencia que a produção de anticorpos não gera memória de longa duração contra RSV, o que favorece re-infecções.

Os principais mecanismos desenvolvidos pelo RSV para escapar do sistema imune são mostrados na [Figura 15](#).

Figura 15 – O RSV e seus mecanismos de evasão do sistema imune



O RSV prejudica a elaboração de uma resposta imune efetiva, por interagir com a imunidade do hospedeiro por meio de diversos mecanismos. 1- A proteína NS2 se liga a RIG-1 e dificulta a resposta imune inata via MAVs. 2- A proteína NS1 bloqueia a ligação de IRF3 ao promotor de IFN-β; a proteína G e a proteína N também inibem a produção de interferon do tipo I. 3- As proteínas NS1 e NS2 aumentam a taxa de degradação de STAT2. 4- A proteína F se liga a TLR4 e pode causar dessensitização da via de sinalização de TLR. 5- A proteína G secretada (sG) pode se ligar a CX3CR1 nas células dendríticas e 6- em alguns linfócitos, ocasionando alteração da quimiotaxia e redução de função. 7- sG pode atuar como uma "isca" para anticorpos de neutralização. 8- A proteína N pode desfazer a sinapse imunológica formada por linfócitos TCD4+ e TCD8+. 9- As proteínas NS1-NS2 reduzem a maturação de células dendríticas, atenuando a eficácia delas em apresentar antígenos. Abreviações: cDC – células dendríticas convencionais; ISG – gene estimulado por interferon; pDC – célula dendrítica plasmocitoide; TLR – receptor do tipo Toll.

Fonte: Adaptado [65].

Resumindo, o ciclo replicativo do RSV envolve as etapas de adesão e entrada, desnudamento e liberação do material genético, replicação e transcrição, empacotamento da partícula viral e, por fim, o brotamento e a liberação do vírus. Apesar de em alguns desses passos, uma ou duas proteínas serem protagonistas, é a atuação conjunta de todas elas que propicia o sucesso da infecção, justamente pela junção de diversos mecanismos de modulação do sistema imune. Desenvolver estratégias antivirais requer não só o entendimento desses passos e da estrutura e atuação das proteínas virais, como também a compreensão de que existe uma delicada rede de interações químicas, físicas e biológicas estabelecidas entre o vírus e a célula. As estratégias podem focar diferentes proteínas virais ou celulares em uma ou diferentes etapas, como também se fixar em uma enzima ou proteína específica de uma etapa. Vamos a elas.

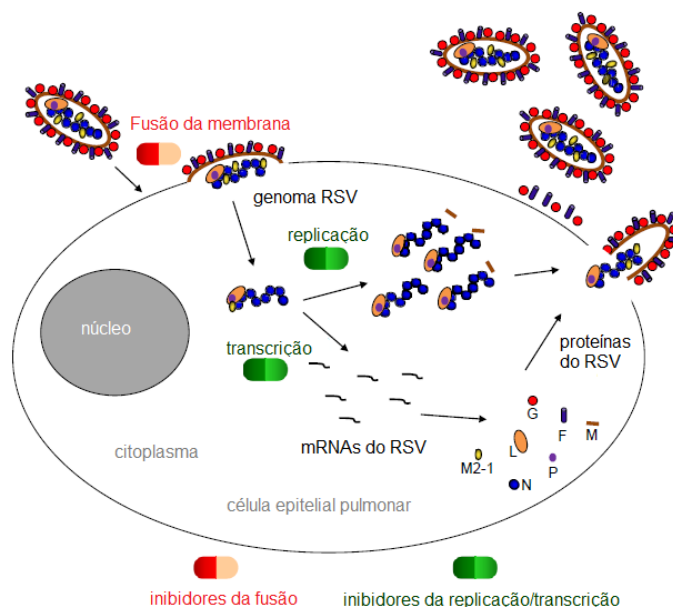
1.1.2.1.2 Estratégias Antivirais contra RSV

*“... Mas onde se achará a sabedoria?
E onde está o lugar
do entendimento?...” (Jó 28:20)*

Elaborar estratégias antivirais, assim como antibacterianas e antifúngicas requer lidar com um desafio pertinente e persistente nesses casos: o surgimento de mutantes resistentes. Devido à alta taxa de erros íntinseca ao processo de replicação-transcrição viral, no caso do RSV ligada à falta de mecanismos de reparo e edição em sua RNA-polimerase [77], a resistência é um fator crucial, o que pressupõe a necessidade constante de elaboração, aprimoramento e combinação de estratégias anti-RSV.

Diferentes etapas do ciclo reprodutivo do vírus na célula podem ser exploradas, na tentativa de inibir a replicação viral [Figura 16]. Focar na inibição da entrada do vírus, por exemplo, apresenta a vantagem de haver uma quantidade menor de vírus com a qual se deve lidar. Nas etapas posteriores, de replicação, brotamento e liberação do vírus, a aposta é na especificidade do composto em questão, de interagir especificamente com os processos e proteínas do vírus, de forma a não causar malefícios à célula hospedeira. Neste caso, cabe mencionar o antiviral acyclovir, um análogo de guanina, pela sua alta seletividade e eficiência na inibição da replicação do vírus do herpes [144].

Figura 16 – O ciclo viral do RSV e algumas estratégias de inibição



O passo-a-passo do ciclo do RSV e as estratégias mais comuns de intervenção, as quais consistem no bloqueio da fusão e/ou da síntese de RNA.

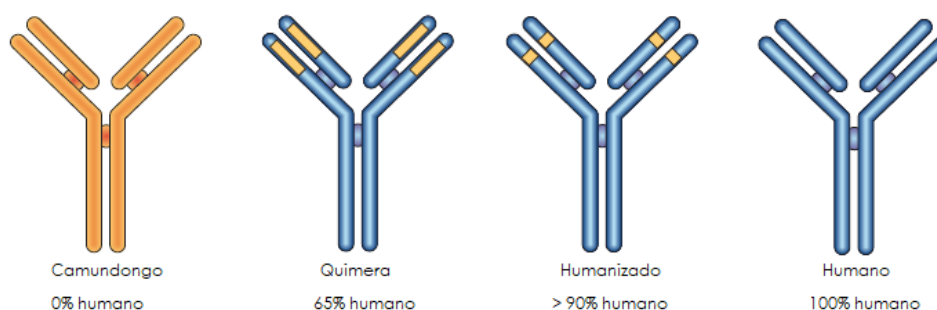
Fonte: Adaptado [3].

Os únicos antivirais disponíveis atualmente contra RSV são o palivizumabe (Synagis®) para prevenção e a ribavirina para tratamento. Além desses, medidas paliativas de cuidado também são utilizadas.

O palivizumabe (MEDI-493) é um anticorpo monoclonal humanizado [Figuras 17 e 18, 145], direcionado à proteína F nas conformações de pré e pós fusão [146]. Logo, sua ação consiste em se ligar à proteína F em seu respectivo sítio (sítio II – aminoácidos 254 a 277 [Figura 19, 146, 147]) e inibir a fusão do vírus à célula [148]. Ele deve ser aplicado em cinco doses mensais, durante a temporada de infecção por RSV, na concentração de 15 mg/kg e intramuscularmente em crianças pertencentes aos grupos de risco – aquelas com má-formação pulmonar, com doenças cardíacas, prematuras, acometidas por desnutrição, imunocomprometidas pelo vírus da imunodeficiência humana (do inglês

human immunodeficiency virus – HIV), dentre outras [149]. Sua eficácia, no entanto, tem sido questionada entre os diferentes grupos de risco [41]. Além disso, já foram identificados mutantes resistentes ao palivizumabe, tanto em modelos experimentais, como também clínicos [150-152]. Os custos de tal profilaxia dificultam sua aplicação em larga escala (1 mililitro da solução intramuscular a 100 mg/mL custa em torno de \$2.960 (dois mil, novecentos e sessenta dólares americanos) [153]).

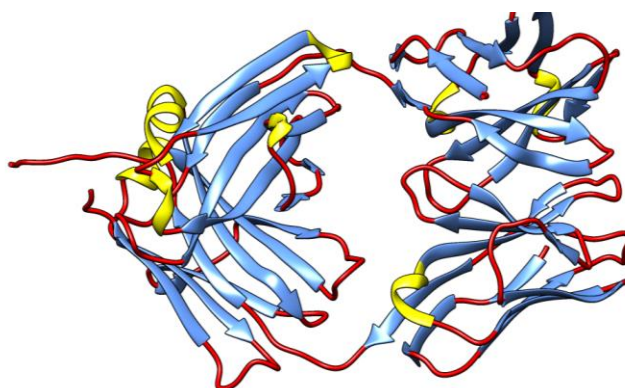
Figura 17 – A engenharia dos anticorpos



O palivizumabe é um anticorpo humanizado, isto é, resulta da combinação, por meio de tecnologia de engenharia genética, de anticorpos humanos e de camundongos.

Fonte: Adaptado [154]

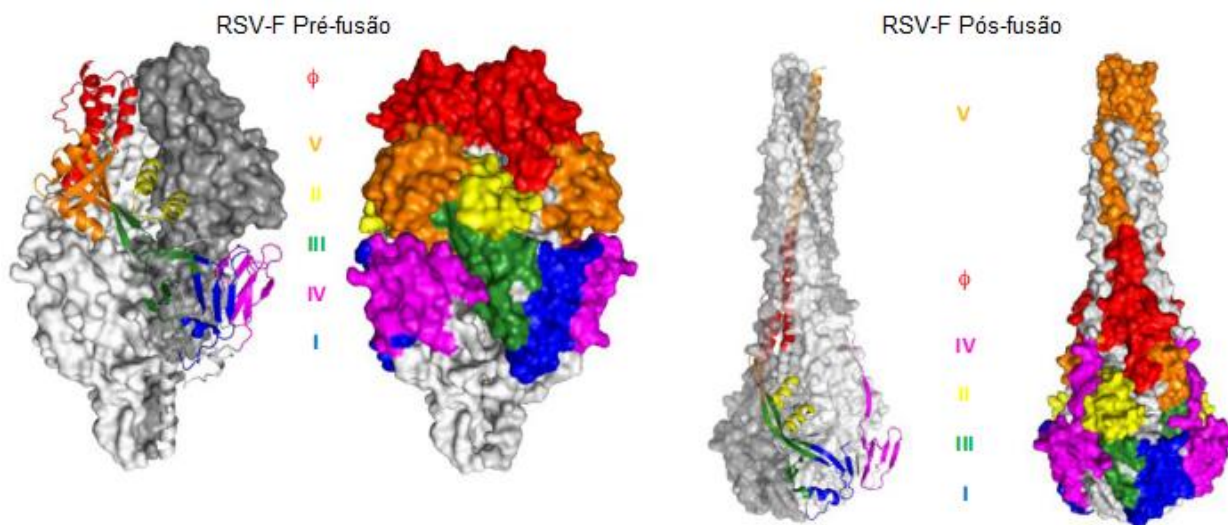
Figura 18 – Palivizumabe (MEDI-493)



A estrutura resolvida do anticorpo Palivizumabe (código de identificação PDB:2HWZ). Edição da imagem feita no programa Chimera [84].

Fonte: o autor.

Figura 19 – Os sítios antigênicos da proteína F

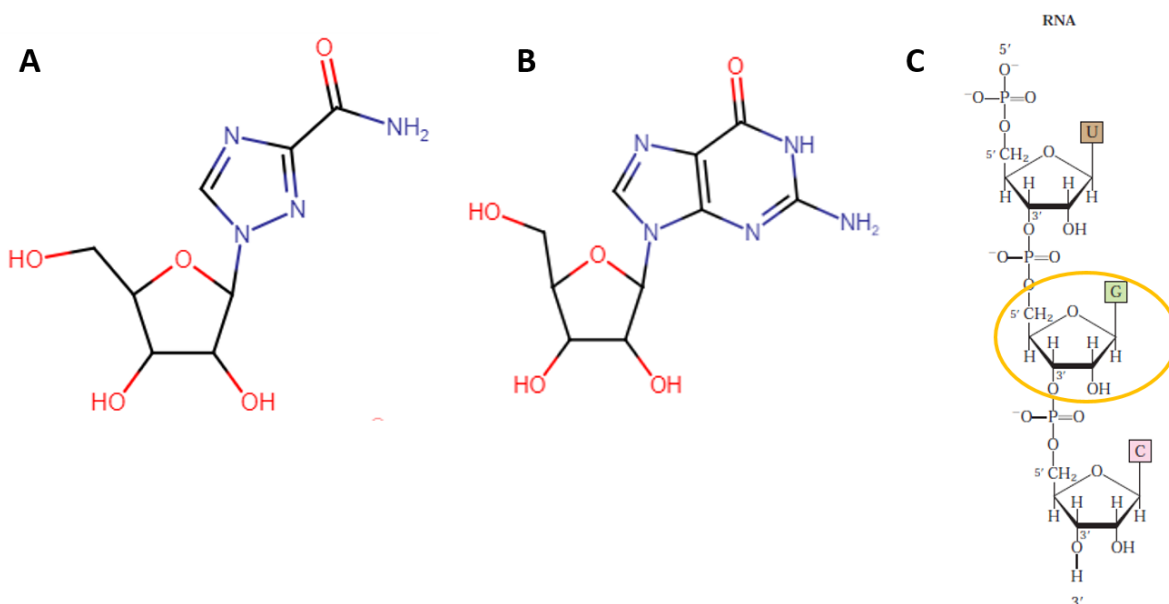


A proteína F é o principal foco de desenvolvimento de estratégias de anticorpos e de inibidores pequenos. Seis sítios antigênicos têm sido descritos, sendo que desses, o palivizumabe se liga ao sítio II (em amarelo) presente em ambas conformações.

Fonte: Adaptado [112].

A ribavirina [Figura 20] é uma molécula análoga ao nucleosídeo guanosina, disponível no mercado para ser usado sob a forma de aerossol. No entanto, apresenta efeitos potencialmente tóxicos (mutagênicos, teratogênicos e carcinogênicos), dificuldades na administração e eficácia questionável [3,4,11-13].

Figura 20 – A Ribavirina



A ribavirina (A) é um antiviral de amplo-espectro utilizado no tratamento de infecções por RSV, pelo vírus da hepatite C e pela febre hemorrágica viral. Sua ação consiste em bloquear a síntese de RNA e o encapsulamento de mRNA virais. Por ser um análogo do nucleosídeo guanosina (B), um dos mecanismos sugeridos para ação antiviral da ribavirina é que ela seja incorporada ao RNA e, por não possuir o fosfato que liga um nucleotídeo ao outro lateralmente, compromete a integridade desse. Moléculas (A) e (B) elaboradas no servidor *Swiss Model* [155].

Fonte: Adaptado [156, 157]

As medidas paliativas consistem no uso de suporte respiratório por meio de ventilação (fornecimento de oxigênio) e de broncodilatadores, e estes últimos têm sido questionados quanto à sua eficácia [4,12,158].

É importante salientar que crianças recém-nascidas e idosos são os grupos mais acometidos pela infecção por RSV, devido ao seu sistema imune não completamente formado ou debilitado. Até aos 2 anos de idade, virtualmente todas as crianças já foram infectadas e aproximadamente 50% delas serão infectadas ao menos duas vezes pelo RSV [159,160]. Logo, a prevenção dessa infecção e seu tratamento devem prezar pelo uso de compostos terapêuticos com características como baixa toxicidade, segurança na administração, eficácia e, além disso, custo reduzido. Embora a elaboração de uma nova

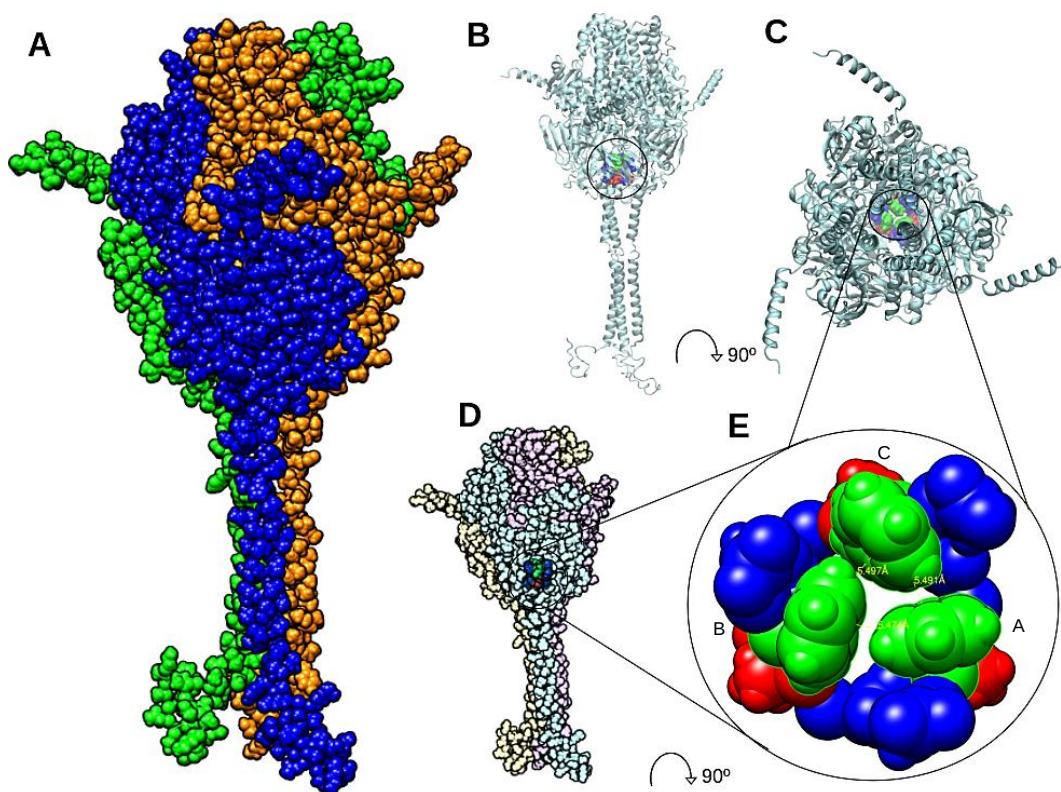
vacina contra RSV tenha entrado em fase 1 de testes clínicos em adultos saudáveis [161], a problemática da resistência justifica o fato de que a busca por compostos ativos continue. Esses compostos devem prevenir a infecção, reduzir a sua gravidade, diminuir o tempo de internação, a taxa de admissão hospitalar bem como os gastos intrínsecos.

Diversas estratégias têm sido estudadas como alternativas ao uso de ribavirina e palivizumabe [para revisão 73,162,163]. Essas abarcam desde a proposta de utilização de cisteinil-leucotrienos para o tratamento de bronquiolite [164], como também o uso de citocinas [165], o desenvolvimento de vacinas [161], outros anticorpos [166], *nanobodies* [167] até a descoberta de pequenas moléculas inibidoras de entrada, fusão ou da replicação desse vírus [3,168-179].

O principal foco das pesquisas, no entanto, é o bloqueio da alteração conformacional de RSV-F. Isso representa uma obstrução à infecção logo em sua etapa inicial. Ainda que o RSV possua o mecanismo de macropinocitose [134], a restrição oferecida por uma proteína F imobilizada configura um grande revés para o sucesso da infecção.

À vista disso, várias regiões de RSV-F têm sido candidatas a alvo de intervenção farmacológica, como sítios antigênicos [257,259], sítios de N-glicosilações [260], resíduos de cisteína [261], assim como diferentes resíduos de RSV F os quais podem interagir com peptídeos e moléculas pequenas capazes de inibir a entrada do vírus [256,258,262]. Battles e colegas [256] descreveram um mecanismo desenvolvido por moléculas pequenas capazes de inibir a entrada do vírus via proteína F. Eles sugerem o inibidor JNJ-2408068 e relatam que sua atividade se deva à ligação com alta afinidade a um bolsão trissimétrico dentro da cavidade central da estrutura de pré-fusão da RSV-F, e, portanto, ocasione a estabilização da região, devido à restrição da movimentação dela. Por conseguinte, mantém o estado meta-estável de conformação pré-fusão [Figura 21], o que evita a alteração conformacional para pós-fusão.

Figura 21 – RSV-F modelada neste trabalho e o bolsão trissimétrico dentro de sua cavidade central



A) Trímero de RSV-F modelada neste trabalho (RSVmF), no formato de esferas. B-D) Diferentes visualizações da estrutura de RSVmF, com o bolsão trissimétrico em realce. Em (B), exibe-se uma vista frontal da proteína, sob formato de *new cartoon*. Em (C), uma vista superior, ainda sob esse mesmo formato. Em (D), vista lateral, sob formato de esferas. Em (E), o bolsão trissimétrico e negativamente carregado isolado da estrutura RSVmF, sob formato de esferas. Esse bolsão é formado por resíduos de fenilalanina (488, em verde, nas cadeias A, B e C), de aspartato (486 e 489, em vermelho, nas cadeias A, B e C) e glutamato (487, em azul, nas cadeias A, B e C). As distâncias entre as fenilalaninas são exibidas, e equivalem a 5,474 Å entre cadeia A (direita abaixo) e B (esquerda abaixo), 5,497 Å entre cadeia B e C (acima, no meio), e 5,491 Å entre cadeia C e A. O VMD [239] e o Chimera [84] foram usados para elaboração e edição da imagem.

Fonte: o autor.

Dentre as muitas moléculas que vêm sendo testadas e, a despeito de todo desenvolvimento da química sintética, os produtos naturais continuam representando um expoente estratégico de exploração [180-183]. Garofalo e colegas [12] e outros [14-18] mostram que o desenvolvimento da infecção por RSV depende de sinalização celular

regulada pela liberação de espécies reativas de oxigênio (EROs ou, em inglês, *Reactive Oxygen Species* – ROS). Os compostos fenólicos, provenientes de metabolismo secundário vegetal, possuem capacidade neutralizante de EROS sendo, portanto, uma alternativa para modulação do processo inflamatório resultante de infecção por RSV, e, possivelmente, para redução da carga viral. Além de sua capacidade antioxidante, os compostos fenólicos interagem com membranas e, por conseguinte, podem interagir com proteínas ali presentes [266-271], sendo que muitas dessas interações explicam seus efeitos farmacológicos. Vejamos algumas dessas moléculas.

1.2 Flavonoides

“... Metafísica? Que metafísica têm aquelas árvores? ...” (Fernando Pessoa)

A história da humanidade perspassa o paradigma saúde-doença e, intrínseco a ele, a forma como a cura é buscada e a doença encarada. Desde os primórdios das civilizações, os registros históricos mostram que, para os povos antigos, as moléstias eram como castigos dos deuses, ao passo que o estar-bem era visto como uma relação harmoniosa para com os mesmos. Se a doença era causada por um desentendimento com as divindades, a cura também deveria ser buscada nelas e no que por elas foi criado, a natureza. Os egípcios inauguram uma medicina mais empírica através não só do desenvolvimento de técnicas de mumificação, como também do uso de ervas medicinais [184].

Assim, os vegetais foram e continuam sendo fontes expressivas de compostos com atividades farmacológicas significativas. Podemos mencionar as civilizações orientais e milenares e seu vasto conhecimento florístico e medicinal. Além disso, o poder ‘curativo’ de alguns chás, os quais, em certas épocas, foram até chamados de “poções mágicas”. De igual forma, o “néctar dos deuses” ou simplesmente “vinho”, tão comumente citado em muitas literaturas, não só como um estimulante mas também como remédio: o primeiro milagre de Jesus foi converter água em vinho; Paulo recomenda a Timóteo que tome vinho como remédio para o estômago; os gregos tinham até um deus consagrado a ele – Dionísio – convertido pelos romanos ao deus Baco. E dessa maneira chegamos a um comprimido de aspirina tão popular atualmente. A produção de compostos com atividades farmacológicas relevantes nada tem a ver com nossas necessidades humanas, mas sim com o custo de sobrevivência requerido daqueles organismos.

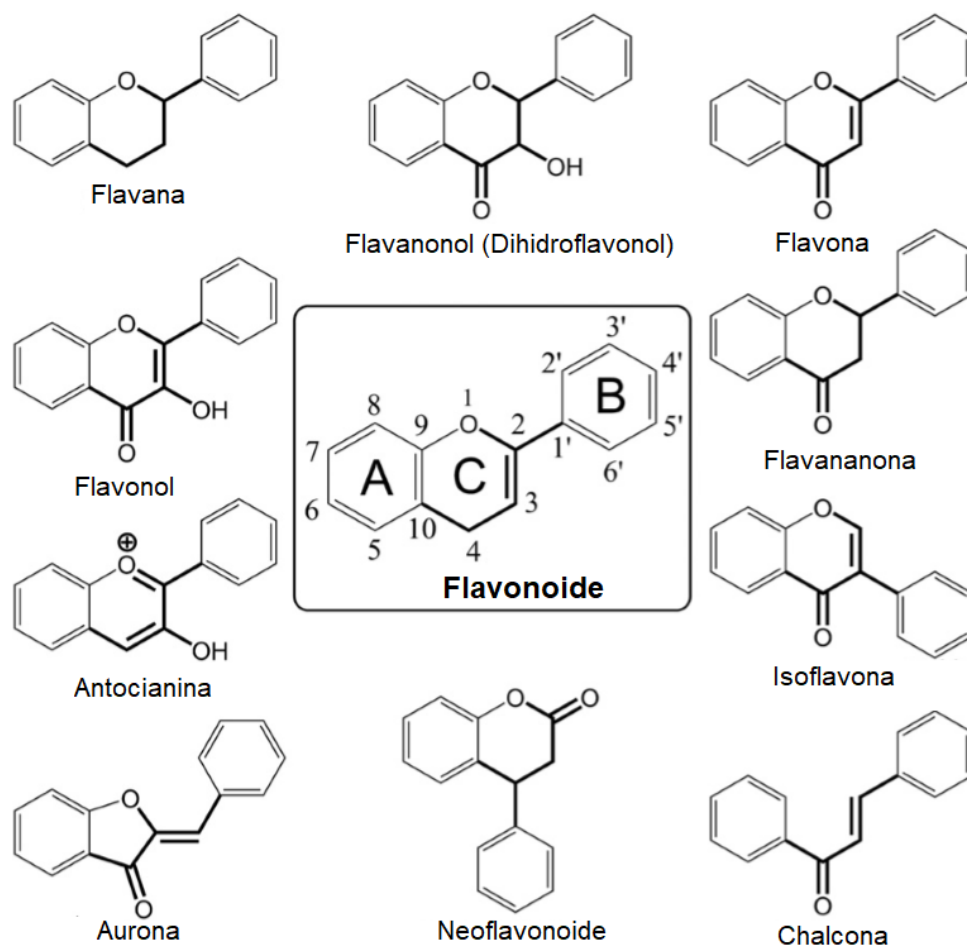
Os vegetais são seres fixos. Isso significa que, independente de como o ambiente ao seu entorno se encontra, eles não possuem como estratégia de sobrevivência a locomoção. Dessa forma, é patente a necessidade de produção de moléculas as quais os possibilitem resistir desde aos raios solares, como também a ataques de

microrganismos, de nematódeos e à herbivoria, e, além disso, a poderem garantir seu pedacinho-de-terra. “*Cada um no seu quadrado*”.

Dentre a mais variada gama de compostos produzidos pelos vegetais, assumem proeminência os compostos flavonoídicos. Flavonoides são moléculas pertencentes ao metabolismo secundário [\[185\]](#) vegetal, isto é, ao conjunto de reações químicas que ocorrem no interior das células vegetais com a finalidade de produzir substâncias que auxiliem na adequação ao ambiente e às suas exigências. A importância desse tipo de metabolismo não se reduz, ainda que classificado como secundário, por não se relacionar à nutrição, respiração ou à síntese de moléculas estruturais, tais como proteínas, lipídios e carboidratos. Ele se destaca, entretanto, por ser essencial para assegurar uma maior chance de competição e sobrevivência. Em sua maioria, os flavonoides compartilham uma estrutura química semelhante (esqueleto de difenil propano com dois anéis benzênicos), como mostrado na parte central [\[Figura 22\]](#).

Entre as suas várias classes, os flavonóis tipo quercetina são os mais abundantes entre os flavonóides conhecidos. Podem ser encontrados em uma variedade de alimentos comuns como maçãs, uvas, tomates, cebolas, chás, entre outros. A quercetina corresponde a cerca de 90% da ingestão humana diária de flavonóides (50-500mg/dia) [\[186\]](#).

Figura 22 – Estrutura básica dos flavonoides e alguns de seus grupos



Os flavonoides diferem entre si, em termos de classes, pelo grau de oxidação do anel "C". Dentro de cada classe, as diferenças residem na presença de radicais e/ou açúcares nos anéis "A" e "B". Moléculas de açúcares (carboidratos) ligadas aos carbonos 6 e 8 (anel A), são classificadas como C-heterosídeos. Se estiverem em outros carbonos que não esses dois, exclusivamente, a molécula é classificada como O-heterosídeo. A quercetina possui radical -OH nas posições 3, 5, 7, 3', 4' (figura central) e uma dupla ligação com oxigênio na posição 4.

Fonte: Adaptado [267].

A quercetina é um flavonoide bastante controverso na literatura. Alguns concordam com o seu potencial, enquanto outros não. E os argumentos de ambos os lados são persuasivos.

A literatura relata uma miríade de atividades sobre o potencial farmacológico da quercetina. Estudos *in-vitro* descrevem propriedades antioxidante, hepatoprotetora, antitumoral, anti-inflamatória, nefroprotetora, imunomodulatória, anti-histamínica, regulatória da pressão arterial, antifibriogênica, gastroprotetiva, entre outras. Além disso, possui atividade antibacteriana contra *Actinobacillus actinomycetemcomitans*, *Micrococcus luteus*, *Actinomyces viscosus*, *Porphyromonas gingivalis* [19,20], *Staphylococcus aureus* [21], contra a formação de biofilmes em patógenos presentes nos alimentos [22] e inibição parcial da DNA girase de *Escherichia coli* [23]. Também apresenta atividade antiviral contra *Herpes simplex* tipo 1, *parainfluenza* tipo 3, rinovírus, influenza A, hepatite C, dengue, contra a transcriptase reversa do HIV e outros retrovírus [24-29].

No entanto, essa vastidão de atividades tem sido questionada quanto à sua efetividade, confiabilidade e validade, por certos autores que a classificam, junto a outros compostos e produtos naturais (*natural products* – NP) como composto de pan-interferência em ensaios (*Pan-assay interference compounds* – PAINS [186]), panaceia metabólica inválida (*Invalid Metabolic Panaceas* – IMPs [187]) ou inibidor promíscuo (*Promiscuous Inhibitors* – PI [188]). As críticas residem justamente no fato de a quercetina apresentar uma vastidão de atividades, o que revelaria inespecificidade. Além disso, menciona-se que tal enormidade de efeitos se deve à propensão da molécula em formar agregados entre si e com proteínas celulares, e, portanto, devido à agregação, causar múltiplos impactos nas diversas atividades celulares. É relatada também a susceptibilidade à interferência com reagentes de ensaios baseados em reações de oxirredução e luminescentes, justamente devido ao seu potencial antioxidante bem como de absorção luminosa [186-189].

Muito vem sendo especulado quanto à especificidade de moléculas candidatas à elaboração de novas drogas, especialmente antivirais, sem que ainda se tenha chegado

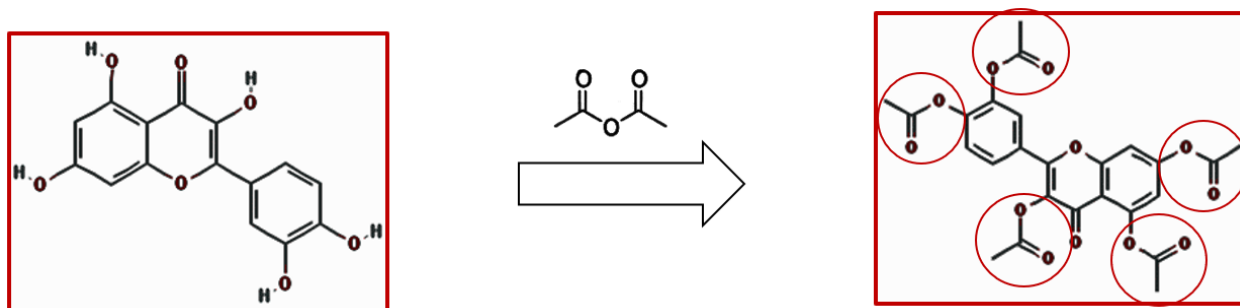
a um consenso. Alguns relatam que a abordagem mais válida seria a da polifarmacologia, ou seja, drogas com múltiplos alvos (*multitarget drugs*), enquanto outros defendem a vertente da especificidade (*one target drug*), a qual, enquanto antivirais, tem sido a predominantemente empregada. Esta última vertente, embora seja bastante eficaz em um primeiro momento, tem produzido medicamentos de alto-custo, ocasionado aumento da resistência viral associado ao baixo índice de resposta em determinados grupos, além de potencializar sua toxicidade na interação com outras drogas [182]. Na verdade, mais recentemente, as nomenclaturas relacionadas a essas duas abordagens vêm sendo revistas, dado que até mesmo compostos tidos “específicos”, como por exemplo, nucleosídeos, não são tão específicos assim. Logo, vêm sendo adotadas as referências *broad-spectrum* e *narrow-spectrum* (amplo e pequeno espectro) [304].

Martinez e colaboradores [182] e outros [190-196] defendem a abordagem polifarmacológica e mostram que, dada a complexidade do organismo em responder às infecções ou mesmo às inflamações, uma abordagem mais abrangente traria como vantagens o combate a diversos patógenos, diminuindo o nível de desenvolvimento de resistência, ao mesmo tempo que seria eficiente para combater coinfeções. Ademais, compostos com múltiplos alvos seriam ideais como uma primeira linha de tratamento ou profilaxia para infecções virais agudas, como aquelas que afetam o trato respiratório. É importante enfatizar que, dadas as diversas possibilidades de interação e alvos, o ideal seria que a proteína pela qual o composto tivesse mais afinidade fosse a responsável pelo seu efeito terapêutico, enquanto que as interações prejudiciais fossem mínimas [193], o que estaria correlacionado com a dose terapêutica [192].

Diferentes estudos têm focado em mostrar a relação entre a estrutura de um flavonoide e sua atividade, baseando-se na avaliação das diferentes posições de radicais no esqueleto dessas moléculas, e também na análise do impacto de algumas modificações – glicosilações ou substituição de radicais hidroxila por radicais metil, acetil, dentre outros – sobre sua respectiva atividade [197-213]. Nesse sentido, Danihelova e colaboradores [209] fazem uma revisão básica, mas interessante, sobre o impacto da lipofilização de flavonoides em algumas de suas atividades farmacológicas, o que inclui,

por exemplo, o aumento da permeabilidade do flavonoide pela membrana. A mesma autora também expõe que a acetilação da quercetina diminui seu efeito citotóxico sobre células de cultura Hela [208]. Um exemplo de acetilação da quercetina pode ser visto na [Figura 23](#) abaixo. Ainda nessa linha de pensamento, Gusdinar e colegas demonstram que a acetilação da quercetina aumenta seu efeito anti-inflamatório em um modelo de edema de pata de rato induzido por carragenina [210]. Park e colaboradores também mostram que a acetilação do resveratrol não afetou a eficiência na inibição de melanogênese celular [211]. Outros pesquisadores, em um estudo combinado de relação estrutura-atividade – SAR – [203] com Ressonância Magnética Nuclear, UV-vis e cromatografia mostram que a a monoacetilação de um glicosídeo aumenta a capacidade de inibição sobre a enzima colinesterase.

Figura 23 – Acetilação da quercetina



A figura ilustra a pentacetilação da quercetina, ou seja, a substituição dos radicais -OH por radicais acetil. Fonte: o autor.

Finalmente, Harvey e colegas [181] mostram que o espaço biologicamente relevante, isto é, os potenciais sítios para ligação em proteínas, é bem mais coberto por *screenings* (triagens) de bibliotecas baseadas em moléculas de produtos naturais do que em bibliotecas de compostos sintéticos. Os mesmos autores argumentam que os principais desafios referentes às moléculas naturais dizem respeito às suas massas moleculares e à sua lipofilicidade, as quais devem ser menores que 350 daltons e

apresentarem um coeficiente de partição ($\log P$) que varie de -1 a 3, respectivamente. Moléculas naturais que se encaixem nessa categoria devem ser usadas como um esqueleto, a partir do qual, modificações estruturais possam ser feitas com intuito de otimizar sua respectiva atividade biológica e aumentar sua especificidade. A validade e o impacto dessas modificações podem ser preliminarmente mensurados via uso de ferramentas computacionais, o que torna a experimentação na bancada menos laboriosa e mais seletiva. Adentremos, então, nesse universo computacional...

1.3 Estratégias Computacionais

“El hombre es un milagro químico que sueña.” (Alfredo Conte)

Apesar de muitas vezes ser um pouco difícil acreditar nisso, nós, seres humanos, somos seres pensantes. Descartes, um dos fundadores do método científico (junto a Francis Bacon) e da geometria cartesiana (ou analítica) que entremeia este trabalho, diria que a confirmação da existência vem através da dúvida. Porque quando duvido, penso. E se penso, existo. O famoso *“Cogito ergo sum.”* Assim, o critério básico para existência, e posteriormente aplicado ao desenvolvimento da Ciência seria duvidar.

A dúvida nos leva não só a (ousar) questionar, como também a buscar respostas e, talvez na ausência dessas, a imaginarmos, a sonharmos ou a ansearmos por algo diferente. Diversos foram os personagens que sonharam com algo diferente, ou mesmo aqueles mais *“fora-da-casinha”*, com algo, porventura, inconcebível para o momento. Os egípcios sonhavam com a imortalidade, e, por isso, desenvolveram as técnicas de mumificação e construíram pirâmides. Nesse mesmo sentido, muitos povos da antiguidade sonharam em eternizar seus impérios e isso se traduziu arquitetonicamente em algumas construções que até hoje desafiam nossa engenharia moderna. Os atenienses, ainda que não plenamente, ousaram ao implementar um projeto de sociedade democrática. Leonardo da Vinci, um daqueles totalmente *“fora-da-casinha”* planejava protótipos de aviões e helicópteros, numa época em que nem carros, como concebemos hoje, existiam. Galileu Galilei atreveu-se a desafiar a lógica predominante do geocentrismo, ainda que não com a intrepidez de Giordano Bruno. Joana D’arc, Marie Curie, Ada Lovelace, Katherine Johnson, Dorothy Vaughan, Cecilia Payne, Mary Jackson e Comandante Grace Hopper prodigalizaram o papel da mulher na sociedade. Martim Lutero quebrou paradigmas ao questionar o abuso da fé, ao escrever e publicar suas 95 teses e ao começar a traduzir a Bíblia para que as pessoas comuns a pudessem ler. Por outro lado, John Lennon, em um contexto de guerra, com sua canção *“Imagine”*, sonhou

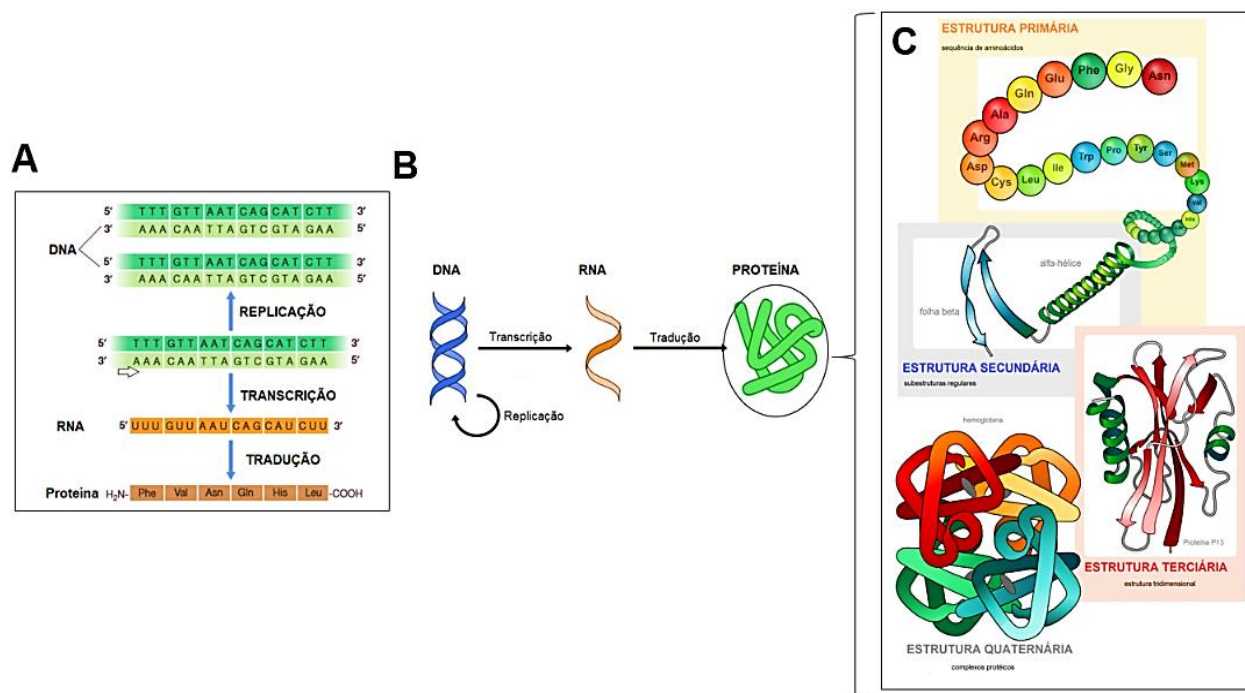
com uma sociedade alternativa, sem religiões. Um outro Martim, desta vez Luther King, com sua política pacifista e seu célebre discurso “*I have a dream*”, também insuflou ânimos que se opuseram a um sistema de ódio racial.

Tantos são os nomes e tantos os sonhos, as causas e os questionamentos advogados por eles... Continuamos (ousando, talvez) a sonhar com uma sociedade mais equilibrada como preconizaram tantas revoluções no passado, e com um Brasil mais educado e, quiçá, menos corrupto... No entanto, certas questões mais elementares atravessam a história da humanidade: o que somos? Para onde vamos? Por que estamos aqui? Se quando pensamos, existimos, qual o sentido da existência? Talvez não encontremos respostas para todas, mas ao menos algumas já começaram a ser solucionadas. E as contribuições são inumeráveis e inomináveis.

O desenvolvimento da filosofia, ou seja, do “*amor ao conhecimento*” é impulsionado pelos gregos na Antiguidade. Filósofos como Platão discorrem, dentre outras coisas, sobre modelos de governo e de sociedade, enquanto alguns como Demócrito e Leucipo estabelecem os conceitos iniciais de “*átomos*”, minúsculas partículas indivisíveis que comporiam a matéria, ou seja, aquilo do que é feito o universo. Esse conceito foi sendo aprimorado ao longo do tempo, por contribuições significativas de personagens como Dalton, Thomsom, Rutherford, Bohr, Heisenberg, Schrödinger, Chadwick e alguns outros. O progresso conceitual sobre do que seria feita a matéria não impediu que outras áreas se expandissem também, como por exemplo, o entendimento das “*leis*” que regulam a matéria, no espaço e no tempo. Um outro grego, Arquimedes, já versava sobre as propriedades da dinâmica dos corpos na água, em 250 a.C. Contribuições significativas relacionadas à astronomia foram feitas por povos antigos como maias, mesopotâmicos e egípcios. Nesse sentido, cabe mencionar o aporte proporcionado pelas pesquisas de Copérnico, Galileu Galilei, Kepler e Newton. As maiores contribuições de Newton se referem ao estabelecimento das bases da mecânica clássica sobre o estudo dos movimentos dos corpos, bem como da gravitação universal e da decomposição luminosa que impactariam os estudos desde então.

Muitos outros cientistas, o que iria muito além da menção no escopo deste trabalho, colaboraram para estabelecer o que hoje temos como a química, a física, a biologia e suas interfaces. Hooke, a partir de um pedaço de cortiça estabelece que a unidade funcional do corpo humano seria a “*célula*”. Pasteur e Koch dão início aos estudos microbiológicos. Mendel contribui com os conceitos de herança genética. Darwin, com a relação evolutiva de desenvolvimento entre os seres vivos. Pauling, Watson, Crick, Wilkins e Rosalind Franklin desvendando a estrutura tridimensional de moléculas fundamentais para a vida [Figura 24]. Sanger inaugura a era dos sequenciamentos com o primeiro sequenciamento de uma proteína. Dayhoff cataloga as sequências protéicas até então existentes, desenvolve o sistema de uma letra para caracterizar os aminoácidos e estabelece algumas das bases para algo que seria equivalente ao “Blast” atual. E a matemática então? O que dizer? Ela acompanha todo o desenrolar civilizatório. Medidas, música, contagem do tempo, trocas, comércio... De mesopotâmicos e sumérios, aos computadores atuais e a nossa realidade virtual, dos quais não abdicamos... Quem hoje imaginaria fazer pesquisa sem o “oráculo” google? Ou se aventurar em uma coleta, ou mesmo uma viagem sem gps? Ou ter de estar fisicamente presente para conversar com alguém, ou comprar produtos que deseja? Ou ainda, não poder ver a pessoa com quem se está falando? Ou talvez, não poder socializar com pessoas de qualquer lugar do mundo e “vender” seus dados em troca de “likes”? *Computador*, máquina de fazer cálculos, ou seja, cálculos, função essa inicialmente desenvolvida por mulheres, por ser considerado um serviço braçal. Para todo esse desenvolvimento foi necessário, entretanto, o estabelecimento de um sistema numérico padrão e que abarcasse o zero. Isso foi propiciado pelas rotas de comércio entre Europa-Ásia (o nosso sistema numérico é indo-arábico).

Figura 24 – Do DNA à Proteína - fluxograma simplificado



Do ácido deoxirribonucléico à proteína. Embora a figura acima exibida seja mais ilustrativa do que explanativa, o intento é ter uma visão global do processo, ou seja, a sequência de passos DNA->RNA->Proteína (A e B), bem como como os níveis de organização dessa estrutura protéica (C). Os trabalhos pioneiros de Linus Pauling com cristalografia, propiciaram o entendimento da estrutura secundária das proteínas e forneceram as bases para que Watson e Crick também desvendassem a estrutura tridimensional do DNA. Há controvérsias sobre a originalidade do trabalho de Watson & Crick, quanto à possível participação colaborativa, mas não inclusa, de cientistas como Wilkins e Franklin. A estrutura linear de uma proteína (C), ou seja, a sequência de aminoácidos da qual ela é composta foi desvendada, primeiramente, por Sanger no sequenciamento da insulina bovina (1958). Margaret Dayhoff auxiliou no estabelecimento de um sistema de uma única letra para catalogar os aminoácidos, bem como em propor matrizes matemáticas para o processo de alinhamento dessas sequências. A relação entre a sequência linear de aminoácidos de uma proteína (estrutura primária) e a determinação de sua estrutura tridimensional continua representando um desafio, não só para área experimental quanto também para área computacional, em que algoritmos combinados a cálculos biofísicos e estatísticos vêm sendo aprimorados com este intuito.

Fonte: Adaptado – (A) [272], (B) [273], (C) [274].

Se outrora o progresso científico foi marcado, majoritariamente, por contribuições individuais, a partir do século XX e, principalmente, com o desmembramento dos diferentes campos científicos e do aprofundamento do conhecimento, as pesquisas se

tornaram mais coletivas e voltadas para a sociedade. Isso é impulsionado pela aplicação dos conhecimentos científicos, ou seja, a tecnologia, voltada às invenções: carros, telefone, rádio, rede elétrica, eletrodomésticos, aviões, vacinas, medicamentos....O que antes era sonho, começa, gradativamente, a se tornar mais e mais palpável, a facilitar a vida do homem contemporâneo, e a prolongá-la, esta última devido, preponderantemente, às descobertas e práticas preventivas na área de saúde.

Contudo, ainda há muito o que descobrir e explorar. Mesmo o desenlace do material genético humano a partir do Projeto Genoma Humano, o qual supostamente desvendaria todo o mistério para cura das doenças que afligem a humanidade, na verdade trouxe muito mais questionamentos do que respostas. Embora os conceitos primitivos de *átomos* tenham sido estabelecidos há mais de 2000 anos, ainda hoje é algo desafiador entender as interações atômicas e moleculares em uma escala experimental. Dentre tantas áreas e aplicações, o desenvolvimento de fármacos, na atualidade, é imprescindível. Não só porque isso nos possibilitou melhorias na qualidade de vida, como por exemplo, tratar doenças como câncer, alzheimer, parkinson, diabetes, como também, porque, à medida que tentamos controlar infecções com antibióticos, antimicóticos e antivirais, teremos de lidar com o surgimento constante de mutantes resistentes. Para isso, é necessário compreender os princípios pelos quais moléculas de interesse interagem ou com os alvos celulares ou com os alvos no microrganismo em questão, nesse caso, proteínas ou enzimas. Todo esse conhecimento é elementar para o desenvolvimento de novas drogas. As principais técnicas utilizadas para essa finalidade incluem **R**essonância **M**agnética **N**uclear (RMN), **R**essonância **P**aramagnética **E**letrônica (RPE) e Cristalografia, as quais apresentam limitações, não só por serem laboriosas e frequentemente custosas, como também por apresentarem dificuldades de representação no que concerne à dinâmica de alterações estruturais dos sistemas biológicos [214].

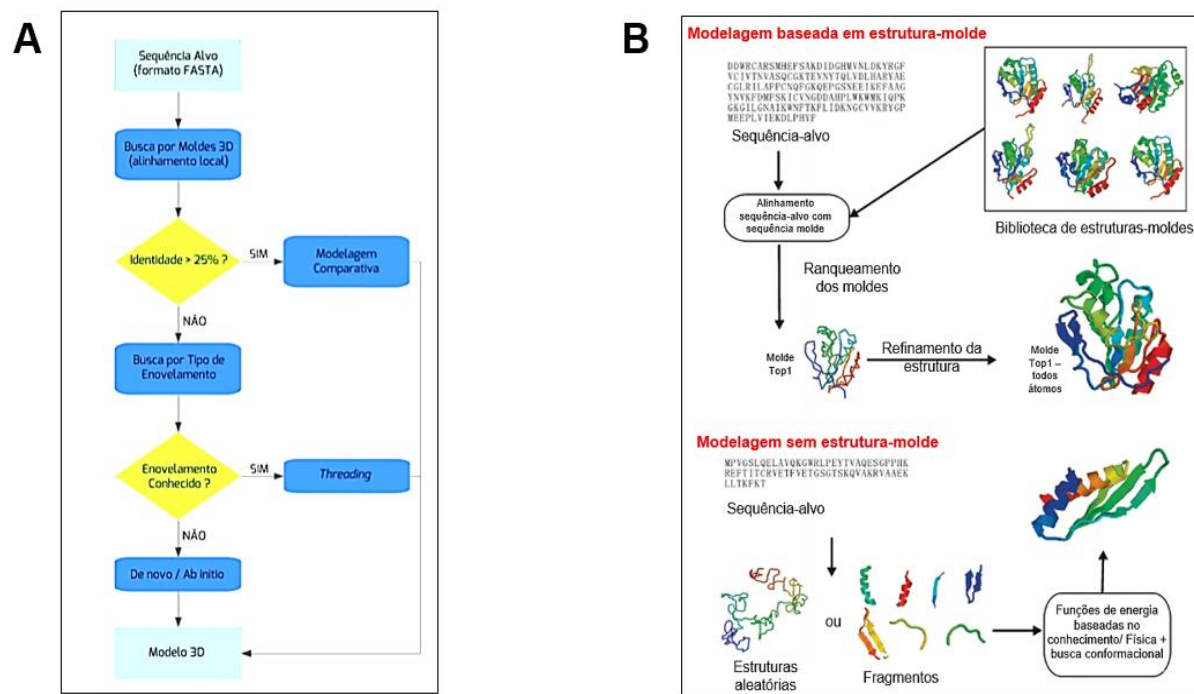
Como alternativa às limitações das técnicas experimentais, diversas ferramentas e estratégias computacionais vêm sendo propostas e utilizadas, assistidas pela constante atualização e aperfeiçoamento de seus algoritmos-base. Essas ferramentas incluem a o

uso de metodologias baseadas em equações físicas e o uso de metodologias fundamentadas numa aprendizagem informática. Cabe ressaltar, entretanto, que o fato de essas metodologias serem baseadas em um ou noutro princípio, não há a exclusão da possibilidade de que elas possam ser combinadas.

As primeiras abordagens são chamadas de *simulações* e embasadas, principalmente, em cálculos biofísicos de estrutura protéica, bem como na utilização de equações provenientes da mecânica clássica, ou seja, em equações de Newton. As segundas, são chamadas *aprendizagem automática* e *data-mining* e se caracterizam pela extração de dados de uma amostragem definida e pela busca por padrões, através de inferências estatísticas [215].

Tanto as simulações quanto as abordagens de aprendizagem e extração de dados têm sido extensivamente utilizadas no processo de descoberta de novas drogas [195-200,202-207,220-222]. Dentro dessa categoria se enquadram a modelagem computacional, o *docking* (“Atracamento”/ “Acoplagem”) e a dinâmica molecular (DM). A proposta da modelagem computacional de proteínas é a de oferecer uma alternativa ao uso de Cristalografia e RMN, e possibilitar que simulemos a estrutura tridimensional (3-D) de uma proteína para o qual temos interesse e cuja estrutura não tenha sido resolvida experimentalmente. Ela pode ser realizada através de métodos independentes de estrutura-molde (*template free* – predição *ab initio* e *de novo*) e aqueles baseados em estrutura-molde (*template based* – *modelagem comparativa* e *threading*) [275, 276]. A adoção de cada uma dessas abordagens diz respeito à qualidade de alinhamentos da sequência primária (a sequência linear de aminoácidos) de uma proteína desejada contra um banco de dados de sequências, bem como à presença de estruturas 3-D de alta qualidade as quais correspondam e representem essas sequências no banco de dados de proteínas (PDB) [275, 276], podendo, ou não, haver relação evolutiva entre essas sequências. Se há relação evolutiva, diz-se que a modelagem é por homologia; em não havendo relação evolutiva, é por enovelamento (*threading*) [Figura 25].

Figura 25 – Resumo das principais técnicas de modelagem de proteínas

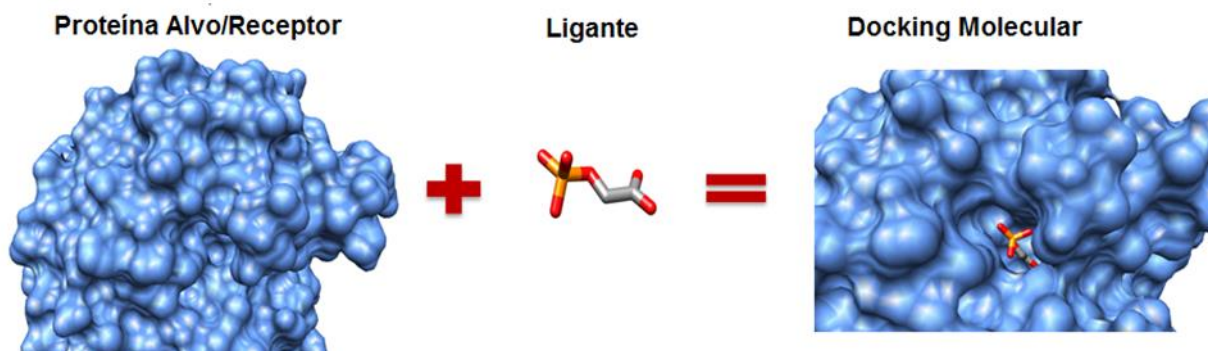


A modelagem nos possibilita simular a estrutura de uma proteína, de uma enzima ou sítio ativo para os quais não existam estruturas tridimensionais resolvidas experimentalmente. A importância da técnica se deve justamente a essa possibilidade, já que experimentalmente tal determinação é um processo laborioso e custoso. A partir da sequência de aminoácidos de uma proteína, é possível verificar se, nos bancos de dados disponíveis (principalmente o “[Blast](#)”) existem sequências similares àquela de interesse, o que pode indicar alguma relação evolutiva entre elas (modelagem comparativa), e/ou para as quais constem estruturas tridimensionais resolvidas. No [Uniprot](#) também é possível averiguar se uma proteína de interesse já foi completamente sequenciada. O Protein Data Bank ([PDB](#)), por outro lado, é o banco de dados mais utilizado para consulta das estruturas tridimensionais de proteínas resolvidas (Biblioteca de estruturas-moldes). A partir do alinhamento e da possível similaridade entre as sequências, a modelagem a ser feita pode ser por homologia ou *threading* (os dois “basicamente” representados na parte superior de (B), pois o *threading* pode relacionar a sequência a um enovelamento para o qual não exista relação evolutiva) ou, em não havendo estruturas 3D disponíveis, a modelagem é feita mais artesanalmente – parte inferior de (B), e envolve *expertise* nos cálculos (Monte Carlo) e refinamentos (bio)físicos, mas também implica em redução da dimensão do tamanho das estruturas estudadas.

Fonte: A [\[b4\]](#), B adaptado [\[275\]](#).

Já quanto ao *docking* e a DM, a proposta é representar macroscopicamente um evento microscópico, isto é, o conjunto de interações interatômicas e intermoleculares entre estruturas tridimensionais (3D) de macromoléculas (*proteínas*, também chamadas *receptores*) com pequenos compostos (*ligantes*), ou de macromoléculas com macromoléculas, através de estratégias como *docking* [Figura 26] e DM. O que mantém e estabiliza esses complexos proteína-ligantes/proteína-proteína são as interações intermoleculares, também chamadas de interações não-covalentes. As ligações de hidrogênio são um exemplo e são bem conhecidas, devido ao seu papel na formação da estrutura secundária de proteínas. Além dessas existem, basicamente, as interações hidrofóbicas, eletrostáticas e de van der Waals [220].

Figura 26 – Representação básica de um Docking Molecular



O *docking* possibilita visualizar a orientação preferencial de interação entre uma proteína alvo e um ligante, ou mesmo entre duas proteínas, através da formação de um complexo. A afinidade dessa ligação pode ser mensurada por meio de funções de pontuação.

Fonte: Adaptado [217].

O *Docking* e a Dinâmica Molecular fazem parte de uma estratégia de desenvolvimento de novas drogas conhecido como “*Modelagem de Drogas Baseada na Estrutura*” (*Structure-Based Drug Design* – SBDD, em inglês). O propósito desse tipo de pesquisa é entender a estrutura de macromoléculas, em termos físico-químicos, e, assim, realizar buscas sistemáticas por ligantes, os quais exibam uma alta afinidade de ligação

com, por exemplo, o sítio ativo dessas macromoléculas [215]. O foco aqui é na estrutura da macromolécula (proteína/receptor). O intuito do *docking* é prever qual a conformação de um determinado ligante dentro de um sítio de interesse na macromolécula [223], de forma a estabelecer qual é o modo de ligação nesse caso, bem como correlacionar isso com a atividade biológica do ligante. A afinidade dessa interação é traduzida em energia de ligação, ou em diferentes funções de pontuação, as quais variam de acordo com o tipo de algoritmo utilizado. Um exemplo de função de pontuação é representado nas seguintes equações (1 e 2) de constante de ligação:

$$K_{\text{ligação}} = e^{\frac{-\Delta G}{RT}} \quad (1)$$

Em que $R \approx 8,314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$;
 T = temperatura em Kelvin;
 ΔG = energia livre de Gibbs ($\text{J} \cdot \text{mol}^{-1}$)

E ΔG equivale a:

$$\Delta G = \Delta H - T\Delta S \quad (2)$$

Em que ΔH = entalpia (J);
 T = Temperatura (Kelvin)
 ΔS = Entropia ($\text{J} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$)

Frequentemente, o que difere a constante de ligação entre alguns tipos de algoritmos voltados a *Docking* é justamente o cálculo da energia livre de Gibbs. Esse cálculo pode variar de forma a considerar diferentemente os tipos de interações não-covalentes, como também a ponderar a presença ou não de solvente, e, além disso, dos efeitos entrópicos [215, 220]. Ademais, outros critérios diferenciam os algoritmos: o tipo

de busca em termos do espaço conformacional que o ligante em questão pode ocupar – dividindo-se em estocático ou sistemático – bem como o critério a ser utilizado no ranqueamento das funções de pontuação – os quais podem ser fundamentados na utilização de um campo de forças, empíricos (como o da [equação 1](#)) ou baseados no conhecimento [revisto por – [215](#), [220](#), [223](#)]. Alguns softwares de *Docking* bastante utilizados são Gold [[224](#)] – estocástico e baseado em campo de forças –, Glide [[225](#)] – sistemático e empírico – e o Autodock [[226](#)] – estocástico, e podendo ser baseado em campo de forças ou empírico.

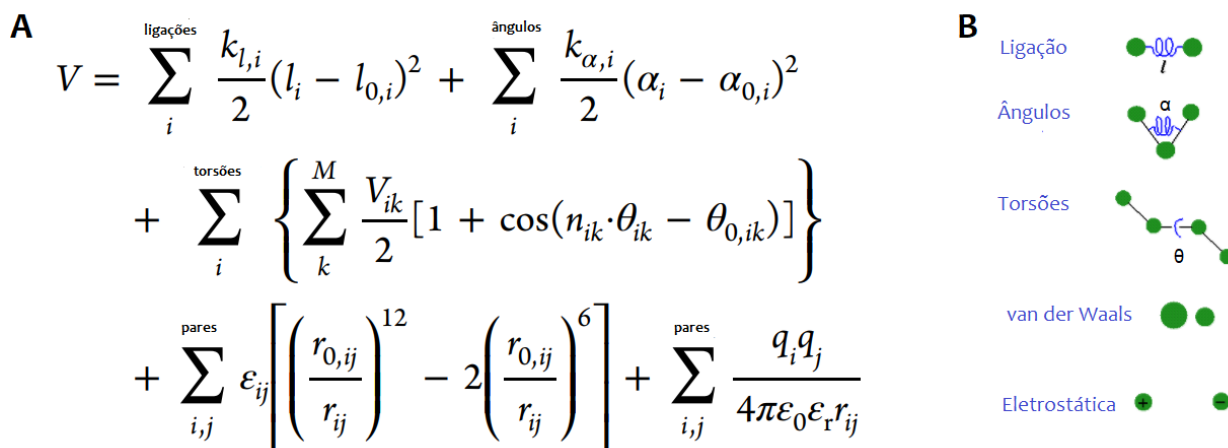
Diferentemente da estratégia do *docking*, o qual tende a considerar uma conformação mais rígida, tanto do receptor quanto do ligante, a DM surge com intuito de ultrapassar essa limitação ao conceder flexibilidade ao complexo proteína-ligante [[214](#),[216](#),[221](#),[222](#),[227](#)]. A ideia por trás dessa estratégia é a aplicação da mecânica Newtoniana (2ª lei de Newton) a um sistema molecular. Consiste no estudo do comportamento de um determinado conjunto de átomos (podendo ser uma proteína, ou um complexo proteína-ligante, ou uma proteína em uma membrana, dentre outros) durante um período de avaliação. Para isso, essa proteína ou complexo é colocado dentro de uma caixa com água, cujo formato depende desse conjunto de átomos e moléculas em questão, formando um sistema (proteína/complexo proteína-ligante + água). Após passar por um processo de minimização de energia, o sistema é gradativamente submetido a um aquecimento até uma temperatura desejada. Isso gera movimento nos e dos átomos, isto é, transforma a energia potencial em energia cinética. Ou seja, as coordenadas espaciais e as velocidades dos átomos vão gradativamente se alterando em função do aquecimento do sistema e do tempo, até a temperatura desejada. Alcançada essa temperatura, o sistema é nela simulado sob determinados parâmetros até atingir o equilíbrio. O comportamento desse sistema é mensurado a partir da resolução da equação de movimento (2) para cada átomo do sistema. Além disso, a interação entre esses átomos e suas energias, ao mesmo tempo, é sistematizada a partir de um determinado campo de forças. A resultante do movimento e das forças atuantes no sistema é atualizada a cada intervalo de tempo, durante um determinado número de

passos. Os cálculos, em relação à equação de movimento (3) e ao campo de forças [Figura 27] são representados abaixo, bem como uma figura esquemática simples de um procedimento de DM clássica [Figura 28]:

$$\vec{F} = m\vec{a} \quad (3)$$

Em que F= força (N);
m= massa (kg);
a = aceleração (m*s⁻²).

Figura 27 – Cálculos envolvidos no campo de forças



Em A, estão representados os termos mais comumente calculados pelos Campos de Forças.

Em que esses representam variações na energia em função:

- do alongamento das ligações (l), de acordo com a lei de Hooke – O parâmetro κ define a rigidez da mola de ligação. $l_{0,i}$ representa a distância de equilíbrio entre dois átomos. Logo, apenas pequenas alterações em l_i são permitidas, já que um valor alto implica em quebra da ligação;
- dos ângulos (α), ou energia de flexão, também baseada na lei de Hooke. O parâmetro κ controla a rigidez da mola de ângulo, enquanto $\alpha_{0,i}$ é o ângulo de equilíbrio;
- das torsões nos ângulos diédricos (θ), ou energia de torsão, modelada por uma função periódica.

V representa a barreira torsional (ou amplitude, isto é, o intervalo de oscilação), θ é o ângulo de torsão ao redor do eixo de ligação, e η é a periodicidade. Esses parâmetros são fundamentados no tipo de ligação entre quatro átomos (e.g. C-C-C-C, C-O-C-N, H-C-C-H, entre outros).

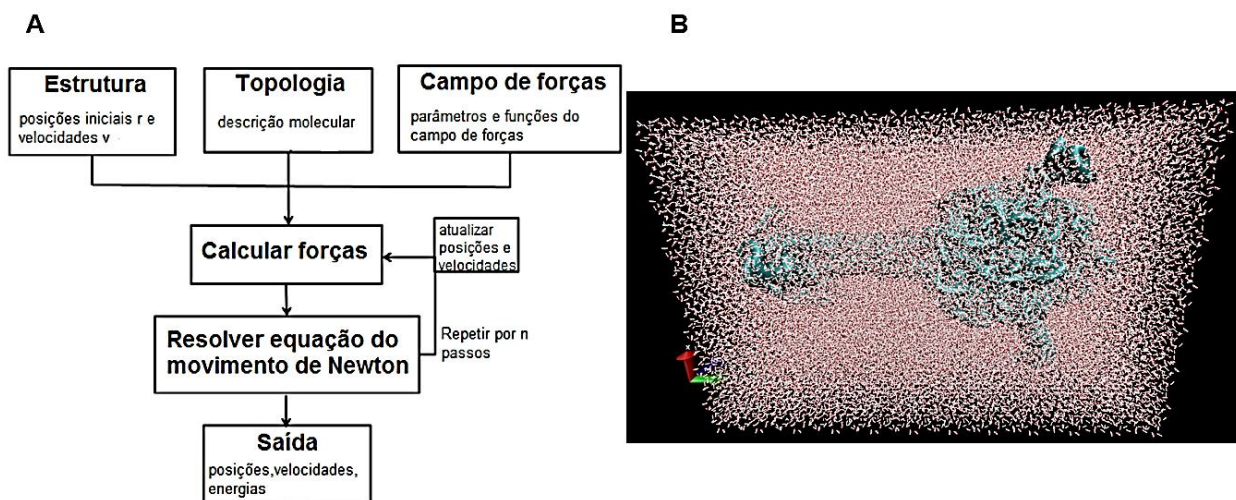
- De forças de van der Waals, ou potencial de Lennard-Jones, em que ϵ é a “profundidade” da curva no gráfico (energia x distância do Potencial de Lennard-Jones) e representa também uma medida de quão fortemente as duas partículas se atraem; σ é a distância em que o potencial intermolecular entre as duas partículas é zero, o que possibilita analisar quão perto duas partículas não-ligadas podem chegar e é, portanto, referido como o raio de van der Waals; r é a distância de separação entre ambas as partículas (medida do centro de massa de uma partícula para o centro de massa da outra partícula)
- e as interações eletrostáticas, baseadas na lei de Coulomb, em que q representa as cargas parciais de dois átomos não-ligados A e B, r é a distância entre eles, e ϵ é a constante dielétrica.

Os termos “ligações”, “ângulos” e “torções” representam os *termos ligados*, isto é, existe uma ligação intramolecular entre os átomos. Já os dois termos “pares” respondem pelos *termos não-ligados* e constituem os cálculos das interações intermoleculares.

Em B há a representação dos termos vistos em A. As esferas verdes representam núcleos dos átomos, e as linhas pretas representam a existência de ligação intramolecular entre eles. No caso dos dois últimos termos, não há ligação intramolecular estabelecida, apenas as interações intermoleculares.

Fonte: Adaptado B [227].

Figura 28 – Fluxograma simplificado de Simulação de Dinâmica Molecular



A – esquema básico dos passos envolvidos e dos tipos de arquivos de entrada numa simulação de Dinâmica Molecular.

B – Visualização da proteína F do RSV modelada neste trabalho, imersa em uma caixa cúbica com solvente, preparada pelo módulo tleap do Amber.

Fonte: Adaptado A [216], B - o autor.

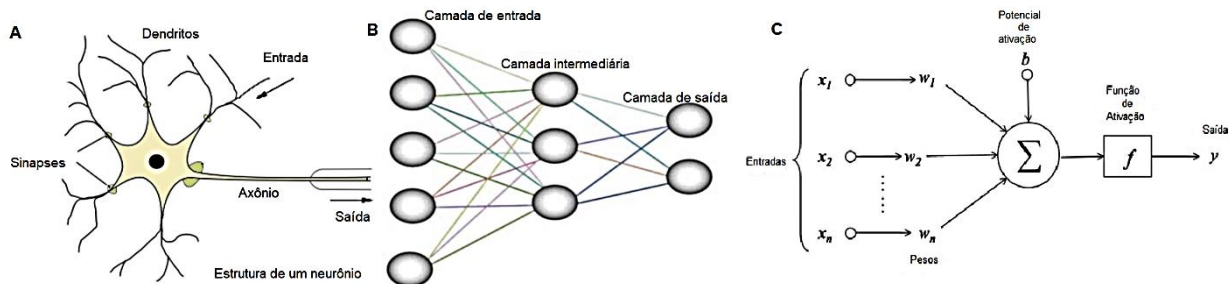
Fica claro que a simulação de DM é baseada em cálculos complexos, fundamentados na resolução da equação de movimento e do campo de forças, a cada passo determinado por um intervalo de tempo. Assim, mesmo para um sistema pequeno, os cálculos para cada átomo e para todos os tipos de interação possíveis (especialmente as forças de van der Waals) se tornam um problema considerável. Logo, esquemas de aproximações são necessários. Isso inclui o uso de “*cutoff*”, isto é, estabelecer um limite de distância para as interações consideradas, bem como de outras abordagens, por exemplo, PME (particle-mesh Ewald) [214]. Os principais pacotes de programas de DM incluem AMBER (www.ambermd.org), GROMACS (www.gromacs.org), CHARMM (www.charmm.org), NAMD (www.ks.uiuc.edu/Research/namd/), Desmond (www.deshawresearch.com/) e Hyperchem (www.hyper.com/) .

Enquanto a DM e o *docking* são estratégias de desenvolvimento de novas drogas voltadas à estrutura, e nesse caso, a estrutura representada é de uma proteína alvo,

existem outras estratégias voltadas aos ligantes. Essas estratégias são chamadas “Triagens virtuais baseadas nos ligantes” (LBVS – *Ligand-Based Virtual Screening*, em inglês). Elas se fundamentam no conhecimento de propriedades físico-químicas de ligantes ativos, de forma a obter um padrão de atividade biológica fundamentada nas características moleculares desses ligantes ativos [215,228,229]. A partir disso, realizam-se buscas sistemáticas por moléculas que apresentem propriedades semelhantes, das quais se espera obter resultados de atividade biológica equivalentes. As ferramentas de *aprendizagem informática* e *data-mining* são expressivamente importantes nesse aspecto. E uma dessas ferramentas são as Redes Neurais Artificiais.

A proposta das Redes Neurais Artificiais (ANN – *Artificial Neural Network*, em inglês) é de simular o complexo sistema neuronal humano, representado pelo cérebro, cujo conhecimento se baseia em um processo de aprendizagem. Um neurônio possui em sua estrutura os dendritos, por onde o impulso nervoso (ou a informação) entra, o corpo celular, por onde a informação é processada, o axônio, por onde a informação passa, e os terminais sinápticos, por onde a informação sai [Figura 29, (A)]. Da mesma forma, a ANN possui a sua camada de entrada, a(s) sua(s) camada(s) de processamento e a camada de saída [____, (B)]. Na ANN, os neurônios são chamados de nós, ou de elementos de processamento [230]. Os nós podem ser interconectados entre si, de forma a gerar multicamadas. A importância dessas múltiplas conexões é mensurada através de um fator peso [230] [____, (C)].

Figura 29 – Redes Neurais



O princípio das redes neurais é o estabelecimento de um sistema interconectado de múltiplos neurônios (neurônios artificiais), os quais operam em paralelo, pelo ajuste (treinamento) dos pesos a eles dados, o que resulta em uma saída com a resposta à alguma questão.

Fonte: Adaptado A e B [218], C [219].

As ANN representam uma estratégia computacional atrativa, não só por possuírem atributos de não linearidade, alto paralelismo, robustez, como também por sua tolerância a falhas, por sua capacidade de aprendizado, habilidade em lidar com imprecisões, e por sua competência em generalizar soluções [277]. Esses atributos propiciam um melhor ajuste aos dados, uma predição adequada ainda que com algum grau de incerteza, além da possibilidade de aplicação de dados novos ao modelo gerado.

Para que uma ANN possa ser utilizada para resolução de algum problema, principalmente aqueles que envolvam sistemas não-lineares (ou seja, nos quais não há uma relação de causalidade direta), é necessário que ela seja treinada. Neste processo de treinamento, ocorre aprendizagem e, então, a rede deve ser capaz de generalizar soluções para o referido problema. Esse processo de treinamento pode ser dividido em duas categorias principais: treinamento supervisionado e treinamento não-supervisionado. O primeiro consiste no conhecimento *a priori* das entradas e das saídas da rede, de forma que os pesos da ANN possam ser corrigidos, para que o erro entre a resposta gerada pela rede e aquela desejada seja mínimo. É como resolver equações e depois checar as respostas. Se você errou, você tem de repetir o processo, até que você se torne apto a resolver e acertar o maior número possível de equações. Por outro lado,

no treinamento não-supervisionado não existem respostas *a priori*, e, portanto, a rede deve se auto-organizar de forma a identificar características comuns entre os elementos do conjunto em questão e categorizá-los. É como organizar o guarda-roupa. Podemos separar as roupas por cor, por tipo, por tecido, por tamanho... Ou seja, algum critério de classificação pode também ser estipulado.

Um treinamento supervisionado pode ser dividido em três etapas: uma fase de treinamento concomitante a uma de validação, e uma de teste, sendo que essas podem seguir a proporção 70:15:15 [278-281]. Para a fase de treinamento é reservada a maior parte da amostragem. A cada ciclo que a rede roda, os resultados esperados são mostrados a ela, de forma que o algoritmo ajuste os pesos (camada por camada) e corrija os erros (camada por camada) [231]. O algoritmo que faz esses ajustes é conhecido como retropropagação de erros (*backpropagation*) [282,283]. A fim de evitar que a rede memorize o conjunto de treinamento, ao mesmo tempo em que esse ocorre, acontece também o processo de validação. A validação é feita com 15% do conjunto amostral. Ou seja, uma parte dos dados os quais não foram mostrados para rede no treinamento, são mostrados e as respostas produzidas pela rede são comparadas com as respostas desejadas, e o erro quadrático dessa comparação também é mensurado. Isso possibilita visualizar se a rede é capaz de generalizar soluções, uma vez que um erro quadrático muito alto na validação indica que não há capacidade de generalização. Os outros 15% restantes da amostra são utilizados para o teste, ou seja, após o treinamento e a validação. Eles nos possibilitam verificar a qualidade final em prover soluções generalizadas da rede. Assim, a rede treinada pode então ser usada para avaliação de um conjunto de dados não-conhecidos.

Entretanto, um dos fatores cruciais para se obter uma boa acurácia para predição executada por uma ANN diz respeito à sua arquitetura (número de camadas e neurônios, bem como os pesos) [284]. Uma das estratégias que tem se provado extremamente eficaz em otimizar a arquitetura de ANN é a adoção de algoritmos genéticos (AG). Os AGs são métodos computacionais os quais combinam princípios de evolução e genética (hereditariedade, mutação, seleção natural e recombinação ou *crossing over*) para

buscar soluções de acordo com operadores probabilísticos [285]. Essa abordagem consiste em produzir uma vasta gama de soluções potenciais e “evoluir” uma solução para o problema. Assim, primeiramente uma população formada por um conjunto aleatório de indivíduos é gerada. Logo, a definição de indivíduos se refere a uma codificação das variáveis relacionadas à arquitetura da rede. Essa população representa possíveis soluções para um dado problema. Durante o processo evolutivo, essa população é avaliada, e, portanto, cada indivíduo recebe uma nota, de acordo com sua habilidade de adaptação ao ambiente. Uma parte dos mais adaptados são mantidos, enquanto o restante é descartado. O grupo mantido pela seleção pode sofrer alterações através de mutações ou recombinação e gerar descendentes para próxima geração. Esse processo de reprodução é mantido até que uma solução ótima, de acordo com algum critério matemático, seja encontrada.

Redes Neurais Artificiais, *Docking*, Dinâmica Molecular e Modelagem Molecular são ferramentas computacionais utilizadas no intento de otimizar estudos, tecnologias, investigações, bem como a descoberta de novos medicamentos, pois, apesar de todos os avanços científicos, a pesquisa contemporânea vem passando por um momento peculiar. A despeito da miríade de desafios a qual o mundo está enfrentando, alguns dos quais incluem explosão populacional, imigração, pobreza, segurança alimentar, crise humanitária, surgimento de microrganismos super-resistentes, reaparecimento de doenças outrora erradicadas, dentre outros, cortes de investimentos cada vez maiores e mais frequentes têm sido anunciados. Há ainda o fato de que, a exigência de soluções profícuas mine toda curiosidade advinda de um resultado malogrado. Erros devem ser considerados como resultados na pesquisa científica? Como lidaríamos hoje com um monge que faz experimentos com ervilhas? Em um artigo, recentemente publicado, os autores revisam que, atualmente, os custos para a produção (do primeiro estudo laboratorial até a liberação no mercado) de uma única droga beira os US\$2.6 bilhões de dólares, em um tempo estimado de 14 anos [221]. Os gastos se devem principalmente a questões de segurança na administração dos compostos candidatos a drogas, bem como a uma baixa atividade numa fase final de avaliação (teste em humanos). Ainda que não

seja possível diminuir todos os custos, essas ferramentas computacionais colaboram no aprimoramento desse processo, ao proporcionarem maior racionalidade, melhor direcionamento, funcionalidade e uniformidade aos procedimentos experimentais, por previamente ranquearem os compostos e excluírem os inativos, o que leva a um aumento de eficiência, e, assim, reduz custos, tempo e favorece maior aproveitamento das pesquisas.

2. OBJETIVOS

O objetivo geral desta pesquisa foi o uso de estratégias computacionais na busca por flavonoides inibidores da replicação do RSV. Assim, avaliou-se o comportamento da quercetina e do seu derivado pentacetilado (Q1) em complexo com a proteína de fusão do RSV por *docking* e Dinâmica Molecular. Além disso, desenvolveu-se uma Rede Neural Artificial assessorada pelo Algoritmo Genético para discriminar a atividade anti-RSV de flavonoides.

Os objetivos específicos incluíram:

- A modelagem da estrutura completa da proteína de fusão do RSV (RSV-F) em seu estado de pré-fusão;
- A dinâmica molecular da proteína RSV-F modelada (RSVmF) (200 nanossegundos);
- A avaliação da energia de interação da quercetina, do Q1 e do inibidor JNJ-2408068 com RSVmF por meio de *Docking* Molecular;
- A formação dos complexos de RSVmF-quercetina, RSVmF-Q1 e RSVmF-JNJ-2408068 e avaliação do comportamento desses ligantes no complexo formado ao longo da trajetória de simulações curtas de Dinâmica Molecular (20 nanossegundos);
- A elaboração de uma base de dados de flavonoides (BD1) para consulta e posterior treinamento da Rede Neural Artificial (ANN);
- A elaboração de uma base de dados de flavonoides (BD2) com a descrição de suas atividades anti-RSV, bem como a descrição das moléculas e de suas propriedades para a implementação da ANN;
- O desenvolvimento de uma Rede Neural Artificial assessorada pelo Algoritmo Genético na plataforma Matlab®, para discriminar a capacidade anti-RSV de flavonoides em “*ativos*” ou “*inativos*”.

3. METODOLOGIA

“Nada é orgânico, é tudo programado...” (Pitty)

3.1 Modelagem Molecular da Glicoproteína F do RSV

O Banco de Dados de Proteínas (*Protein Data Bank* – PDB) contém mais de 20 estruturas cristalizadas da Glicoproteína de Fusão do RSV em complexo com anticorpos ou inibidores em alguns sítios de ligação. No entanto, essas estruturas não possuem as três hélices que formam a espiral na região C-terminal e a maior parte de regiões do tipo mucina. Visando obter a estrutura completa da proteína, foi realizado um procedimento de modelagem, através da submissão da sequência correspondente à Glicoproteína F (UniProtKB: P03420, resíduos 1 ao 574) no servidor I-TASSER, utilizando como *template* o cristal PDB ID: 4JHW. O servidor I-TASSER, que combina procedimentos de modelagem *threading* e *ab-initio*, e que tem sido avaliado como o melhor programa de modelagem *threading* pela comunidade de experimento CASP [232], fornece uma bancada de trabalho *on-line* para a modelagem de alta resolução de estruturas de proteínas e funções (zhanglab.ccmb.med.umich.edu/I-TASSER/). No final do procedimento, I-TASSER gera 5 modelos com diferentes estruturas e C-score. O C-score representa uma pontuação de confiança para estimar a qualidade dos modelos propostos pelo I-TASSER e é calculado com base na significância do alinhamento *threading* com o *template* e os parâmetros de convergência das simulações de montagem de estrutura. O C-score varia tipicamente de -5 a 2 e valores superiores significam um modelo 3D com um alto grau de confiança e vice-versa [232]. Os modelos 3D produzidos por iTASSER (5 modelos) assemelhavam-se à estrutura cristalina 4JHW e o C-score mais alto (igual a 0,08) foi escolhido para realizar a modelagem final. A estrutura modelada foi em seguida importada no UCSF Chimera [84], onde foi triplicada e sobreposta por meio da ferramenta MatchMaker para a montagem da estrutura biológica 4JHW de modo a obter o modelo final.

3.2 Preparação do modelo para simulação de Dinâmica Molecular (MD)

A topologia do sistema e as coordenadas, usadas como entrada para a versão habilitada de GPU do *pmemd*, bem como o mecanismo de MD incluso no pacote AMBER12, foram obtidos através do módulo AmberTools 14 tLeap [233], estabelecendo os padrões para as estruturas através do campo de força AMBERFF14SB [234]. O campo geral de forças AMBER (gaff) foi usado para parametrizar os ligantes usando o antechamber. As estruturas foram então imersas numa caixa retangular, preenchida com moléculas de água TIP3P [235], impondo uma distância mínima entre o soluto e a caixa de 14 Å, ao passo que as cargas foram neutralizadas com a adição de íons Cl⁻, ao sistema solvatado, em posições favoráveis.

3.3 Protocolo de dinâmica molecular (MD)

Uma minimização foi realizada por 5000 passos usando o método *steepest descent*, impondo um *constraint* harmônico de 50 kcal·mol⁻¹·Å⁻² para remover qualquer interação desfavorável introduzida pelo procedimento de modelagem. O sistema foi gradualmente aquecido pelo *ensemble* NVT de 0K a 300K por um período de 500 ps usando o termostato Langevin [236], com o coeficiente de acoplamento de 1.0 ps e um *constraint* fraco de 15 kcal·mol⁻¹·Å⁻². No último passo de equilibração, o sistema foi sujeito a uma simulação de equilibração por 1 ns removendo todos os *constraints*. O sistema otimizado foi então simulado usando o *ensemble* isobárico-isotermal NPT por 200 ns (proteína apenas) e 20 ns (proteína com ligantes), usando condições de contorno periódicas, um intervalo de tempo entre cada passo de 2.0 fs, um *cut-off* de 9 Å para avaliação das interações não-covalentes de curto alcance e o método PME para interações eletrostáticas de longo alcance. O algoritmo SHAKE foi usado para restringir ligações covalentes envolvendo átomos de hidrogênio. A temperatura foi fixada em 300K usando a dinâmica de Langevin, enquanto que a pressão foi mantida constante a 1 atm, através do método de Langevin piston. As posições atômicas foram salvas a cada 500

passos (1.0 ps) para as análises. Todas as simulações foram totalmente realizadas usando 2 NVIDIA TESLA C2075 GPUs.

3.4 Análises da simulação de MD

O desvio quadrático médio (RMSD), a flutuação quadrática média (RMSF), a estrutura secundária, os *eigenvalues* da matriz de covariância, e a análise dos clusters da trajetória da proteína modelada foram avaliados pelos módulos *g_rms*, *g_rmsf*, *do_dssp* [244], *g_anaeig* [242,243] e *g_cluster* [329] do pacote MD GROMACS 4.6.7 [237]. A distância e as ligações de hidrogênio entre a proteína e os ligantes foram avaliadas usando os módulos *g_dist* e *g_hbond* do mesmo pacote. As energias de interação foram avaliadas através do módulo MMPBSA do pacote Amber12 usando a abordagem *gbsa* [238]. As imagens das estruturas foram obtidas usando os programas VMD [239] e UCSF Chimera [84], enquanto que os gráficos foram obtidos através da ferramenta Grace plotting tool (plasma-gate.weizmann.ac.il/Grace/).

3.5 Procedimento de Docking Molecular

Os *dockings* proteína-ligantes foram executados com o programa Auto-Dock Vina 1.1.2, usando o plugin AutoDock/Vina PyMOL [240]. Os arquivos SDF da quercetina, da quercetina penta-acetilada (Q1) e do inibidor JNJ-2408068 foram obtidos da base de dados PubChem, sendo convertidos a arquivos de PDB e completos com hidrogênios usando o programa Open Babel program [241]. O programa SYBYL 6.0 (Tripos Inc. 1699, South Hanley Road St. Louis, Mo, USA) foi usado para minimizar as estruturas dos compostos. Os procedimentos de docking foram realizados usando como receptores os clusters centroides da proteína obtidos pela clusterização prévia da trajetória de MD. A simulação de *docking* foi realizada usando o algoritmo genético como gradiente local de otimização. Para cada simulação de docking, uma caixa de 75 x 75 x 75 (*grid points* nos

eixos X, Y, Z) foi centralizada no sítio de ligação correspondente a cavidade vazia gerada pelos resíduos PHE 488, 1060, 1636. A afinidade dos compostos dockados foi expressa em termos de energia de ligação (kcal/mol). Cada simulação de *docking* levou cerca de 10 minutos em uma estação de trabalho Intel i7 CPU.

3.6 Elaboração da base de dados de flavonoides para Rede Neural Artificial

A etapa de elaboração da base de dados foi dividida nos seguintes passos:

- a) Primeiramente, uma base de dados de flavonoides (BD1) foi elaborada, a partir de revisão de literatura, como também de informações experimentais provenientes do banco de dados públicos Pubchem®, e, além dessas, moléculas da base de dados ChemFaces (www.chemfaces.com). A BD1 contém 690 flavonoides.
- b) Para selecionar os flavonoides com atividade anti-RSV já descrita (BD2), uma revisão sistemática e extensa de literatura foi realizada. Foram considerados apenas os manuscritos cuja parte experimental era bem descrita e consistente, e dos quais se foi possível extrair os dados referentes aos parâmetros estipulados como necessários para compor as entradas da rede, a saber: classe do flavonoide, tipo (aglicona ou glicosídeo), ausência/presença de glicosilações e suas posições, fórmula molecular (C-H-O), peso molecular, coeficiente de partição octanol/água – LogP, a área de superfície topológica polar (*Topological Polar Surface Area* – TPSA), concentração do flavonoide (µg/mL), o tipo viral classificado entre *Long*, *A2* e *B* e a quantidade de vírus em Unidade Formadora de Placas (*Plaque Forming Units* – PFU). O tipo celular foi definido como Hep-2. A BD2 contém 240 dados, referentes a 70 flavonoides.
- c) O banco de dados públicos Pubchem® foi utilizado para consulta das informações concernentes à estrutura do flavonoide, como também para obtenção dos arquivos .sdf. A matriz das coordenadas atômicas descritas no arquivo .sdf representa a estrutura 3D do flavonoide e foi extraída.

- d) Parte dos dados, principalmente quanto às moléculas inativas, foram obtidos por meio de comparação entre a base de dados de flavonoides por nós elaborada (BD1) e o banco de dados de um screening de alta performance [44], através do uso de um programa *in-house* desenvolvido na plataforma Matlab®. Todos os parâmetros de entrada da ANN referente aos flavonoides da BD2, foram devidamente padronizados e as informações textuais receberam numerações afins. A matriz das coordenadas atômicas proveniente do arquivo .sdf de cada flavonoide foi trabalhada na plataforma Matlab por um programa *in-house* de forma a concatenar os dados e escrever um vetor-coluna para cada flavonoide. O tamanho dos vetores foi padronizado, a partir do acréscimo de um valor arbitrário de “-100”. Posteriormente, os dados referentes à cada flavonoide (mencionados no item “b”), receberam numerações afins e foram adicionados ao final desses vetores.
- e) Os flavonoides cujas estruturas 3D não constavam no Pubchem®, foram obtidos pela utilização do software *Swiss ADME* [155] para extração do arquivo *smiles* e caracterização da molécula, para, em seguida, realizar-se a confecção do arquivo .sdf no software OpenBabel [241]

3.7 ANN

Para o processo de aprendizado da ANN, foi utilizado o algoritmo Backpropagation [282,283]. O banco de dados (n = 240 dados, referentes a 70 flavonoides) foi dividido em treinamento, validação e teste, com, respectivamente, 70, 15 e 15% dos dados e seguindo um padrão de validação já estabelecido [278-281]. Diversas funções de transferência que realizam a passagem de informações entre os neurônios foram usadas (logsig, purelin, tansig, hardlim, tribas, radbas e satlin [286], funções de treinamento (trainrp, trainscg, traincgf, traingdx e traingdm [286], camadas ocultas (entre 1 e 3) e o número de neurônios em cada camada (entre 10 e 120, ref [287]).

A precisão da ANN obtida foi verificada de acordo com o erro r^2 entre os valores reais (a saída desejada para rede, convertida em “0” e “1”, de acordo com o critério de “ativo” ou “inativo”) e os valores obtidos pela ANN [282, 287, 288, 289].

No entanto, não existe um método padrão para obter a melhor arquitetura (o número de neurônios em cada camada, número de camadas, funções de treinamento e transferência) da ANN para a solução de um problema [290]. Assim, a técnica de Algoritmo Genético foi usada para melhorar a eficiência em determinar a melhor arquitetura de ANN para o nosso problema [290].

3.8 Algoritmo genético (AG)

Neste estudo, consideramos que a população a ser estudada pelo AG foi formada por diferentes ANNs, e o objetivo foi obter redes que apresentassem o menor erro na classificação de flavonoides em “ativos” ou “inativos”. A técnica de AG desenvolvida em nosso estudo considerou a criação de uma população inicial de ANN com diferentes arquiteturas (indivíduos), que foi gerada aleatoriamente e composta por 100, 200 ou 300 indivíduos.

Cada um foi definido por um “código genético”, ou seja, uma especificação por 9 genes diferentes (o número de neurônios na primeira, segunda e terceira camadas ocultas; a função de transferência para a primeira, segunda e terceira camadas ocultas; função para a camada de saída, a função de treinamento a ser usada e a quantidade de camadas ocultas a serem usadas). Cada ANN foi treinada e testada, e seu percentual de sucesso para o grau de classificação foi avaliado. Portanto, a arquitetura da ANN foi definida em termos de sua precisão. Um fluxograma do AG é mostrado na [Figura 30](#).

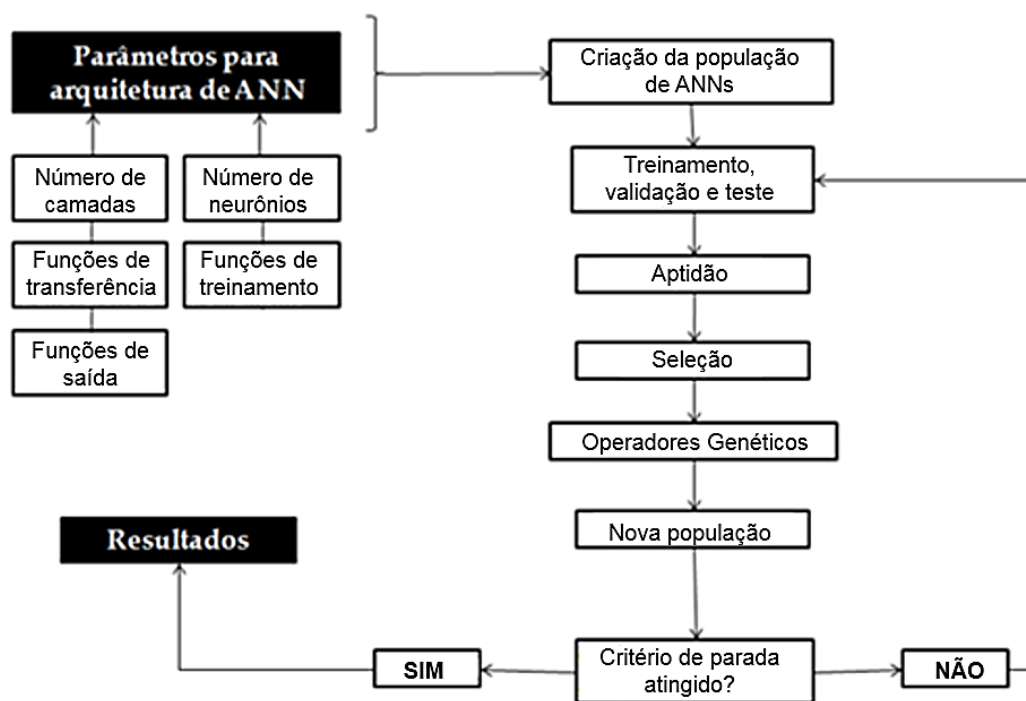
Na etapa seguinte, foi utilizado um processo de seleção de acordo com o elitismo. Nesse processo, os indivíduos foram organizados com base na precisão obtida pela ANN na classificação dos dados dos testes. Após a etapa de seleção, 10% dos indivíduos mais aptos (ou seja, com menor classificação incorreta) foram usados na próxima geração.

Na etapa de replicação e substituição, além dos 10% dos indivíduos mais aptos da geração anterior, 60% da população consistia em novos indivíduos adicionados a

partir da aplicação da técnica de *crossing-over* e mutação entre os indivíduos da população (o que rendeu máximo de 5% para o efeito de mutação). Os indivíduos restantes (30%) foram gerados aleatoriamente, respeitando as condições que resultaram da população inicial. Após a seleção, replicação e substituição, houve uma nova população a partir da qual as melhores ANNs foram salvas. Este ciclo foi denominado uma geração.

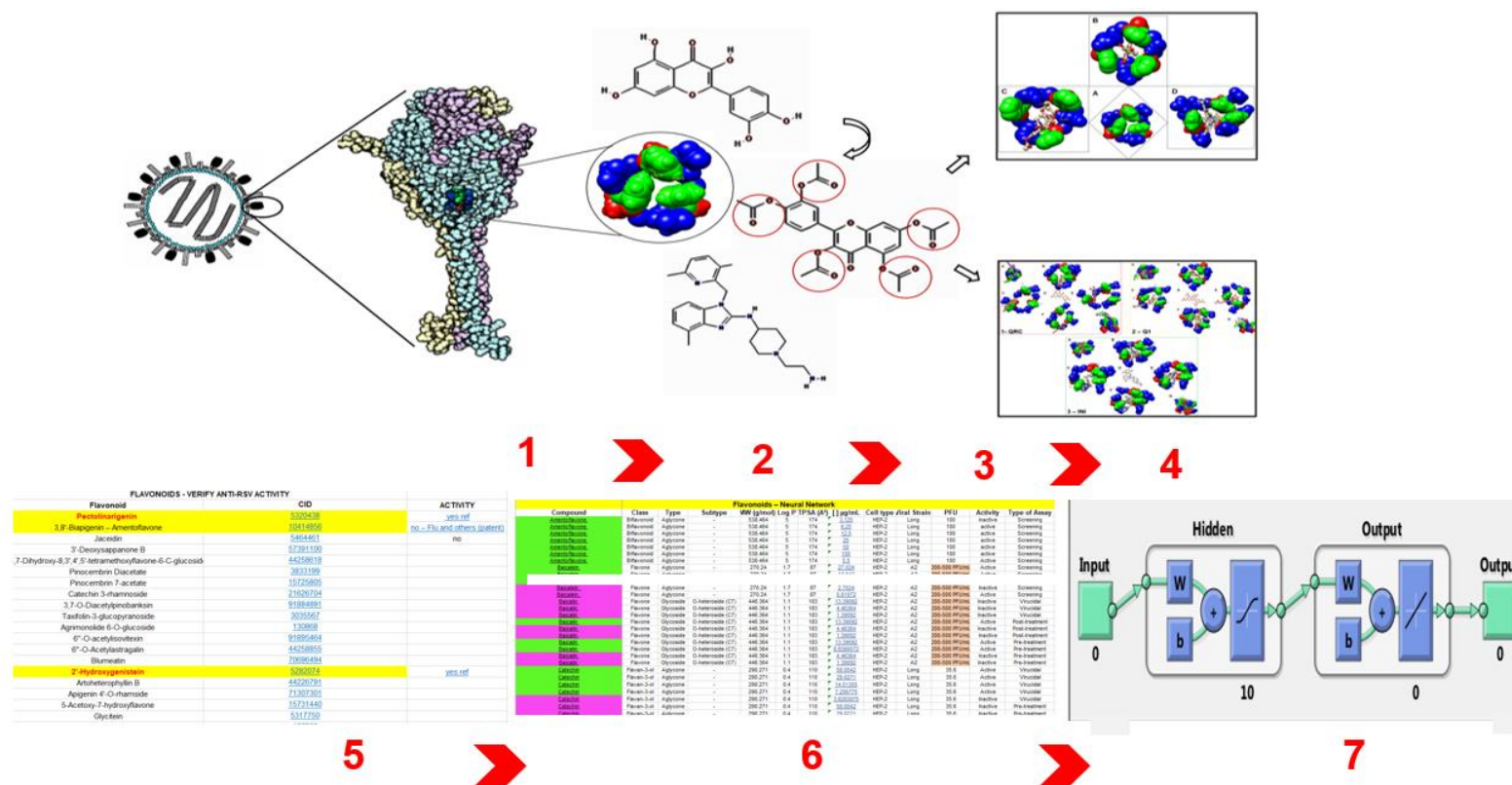
Em média, foram produzidas 1000 gerações para cada população criada. Quando determinamos o número de indivíduos na população inicial (100, 200 ou 300 indivíduos), esse número de indivíduos permaneceu inalterado até o final das 1000 gerações. Todo o procedimento foi realizado 500 vezes diferentes para 100, 200 e 300 indivíduos.

Figura 30 – Algoritmo genético



Algoritmo genético utilizado neste trabalho.
Fonte: o autor.

Figura 31 – Resumo metodologia



Este trabalho pode ser sintetizado em sete eixos principais: 1. Modelagem Completa Proteína de Fusão do RSV no seu estado de pré-fusão (RSVmF); 2. Seleção dos ligantes quercetina, seu derivado acetilado e JNJ como controle; 3. Docking dos ligantes com cavidade central RSVmF; 4. Dinâmica Molecular (DM) cavidade central RSVmF com os ligantes; 5. Elaboração de base de dados de flavonoides (BD1); 6. Elaboração de base de dados de flavonoides para a Rede Neural Artificial (ANN) e 7. Implementação da ANN.

Fonte: o autor.

4. RESULTADOS E DISCUSSÃO

“Eppur si muove!” (Galileu Galilei)

Flavonoides são metabólitos secundários e, ainda que não sejam produzidos com fins de garantir cura às moléstias humanas, suas características químicas e estruturais os conferem de uma vastidão de atividades farmacológicas, devido à sua capacidade de interação com biomoléculas humanas e/ou de outros (micro) organismos. Entender o modo como essas interações acontecem, possibilita-nos propor modificações estruturais com objetivo de refinar determinada atividade. Outra consequência advém da possibilidade de se desenvolverem ferramentas que auxiliem a predizer a capacidade inibitória dessas moléculas.

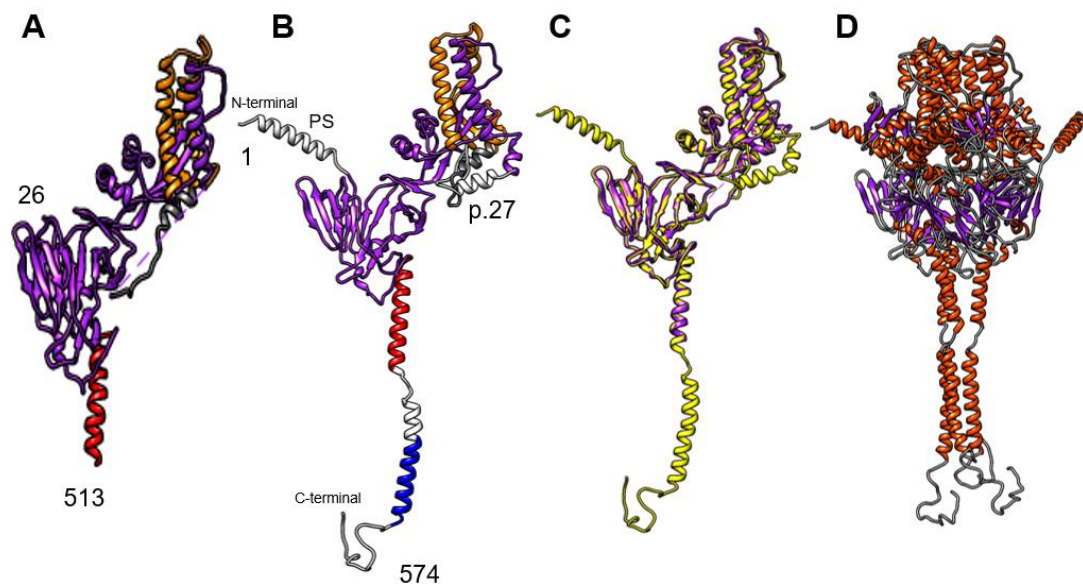
Visamos avaliar a capacidade virucida da quercetina e do seu derivado quercetina pentacetilada (Q1) nos complexos RSVmF-quercetina e RSVmF-Q1 por meio de *Docking* e Dinâmica Molecular (DM), e verificar a influência da acetilação nessa interação. Além disso, estabelecer uma Rede Neural Artificial – ANN – para predizer flavonoides com capacidade anti-RSV. Os resultados apresentados seguem, portanto, a seguinte ordem: Modelagem, DM e análises da proteína F; *Docking* e DM de RSVmF-quercetina/RSVmF-Q1 e análises; ANN: elaboração e implementação.

4.1 Modelagem de RSVmF, dinâmica molecular e análises

O primeiro passo foi elaborar a modelagem de RSV-F, a qual foi realizada com intuito de se obter uma estrutura mais completa da proteína. O PDB 4JHW foi escolhido devido ao fato de, no momento da elaboração da modelagem, ser o cristal mais completo, ou seja, aquele com maior representatividade estrutural da proteína F, a despeito de sua resolução razoável (3,6 Å). O monômero da estrutura obtida – RSVmF – contém 574

resíduos, e seu trîmero 1722. O monômero-molde 4JHW possui 449 resíduos resolvidos (sequências 26-97, 137-513), e o seu trîmero 1347 resíduos. RSVmF apresenta a sequência inicial dos aminoácidos de 1-25, os quais representam o peptídeo sinal, o peptídeo 27 o qual engloba a sequência de aminoácidos 109-137, além dos domínios transmembrânicos (530-550) e a cauda citoplasmática (551-574). Ela é reproduzida na [Figura 32](#):

Figura 32 – RSV-F pré-fusão modelada (RSVmF)



RSVmF foi modelada a partir do PDB 4JHW (A), utilizado como molde, e, a partir do qual foram adicionadas as regiões transmembrânica e C-terminal, além da inclusão do peptídeo sinal (PS). O monômero de RSVmF é visualizado em (B). Em (C) apresenta-se a sobreposição de (A) e (B), e, em (D), visualiza-se o trîmero de RSVmF.

Em (A) e (B) as disposições das cores são as mesmas: em cinza, peptídeo de fusão; em laranja, a região de repetição de padrão “*Heptad*” HRA; em vermelho a região de repetição de padrão “*Heptad*” HRB; em azul, o domínio transmembrânico e, em branco, aparecem a cauda citoplasmática, a sequência dos aminoácidos 513-530, o peptídeo 27, e o peptídeo sinal (sequência de 1 a 25). Os dois últimos são clivados pela furina protease. Em (C), visualiza-se a sobreposição dos monômeros: em amarelo, o monômero de RSVmF, e, em roxo, o monômero de 4JHW. Finalmente, em (D), a coloração varia conforme a disposição de estrutura secundária. Em laranja estão as hélices, em lilás as folhas- β , e em cinza os *random coils* ou alças. RSV-F é uma glicoproteína do tipo I, isto é, seu domínio N-terminal se localiza externamente à membrana do vírus, enquanto que seu domínio C-terminal é citoplasmático (evidenciado em (B)).

Fonte: o autor.

A modelagem de uma estrutura completa de RSV-F implica em um sistema relativamente grande – 1722 resíduos de aminoácidos – distribuídos em uma estrutura de trímero. Uma primeira análise que pode nos fornecer um parâmetro de qualidade da modelagem diz respeito à posição correta dos ângulos de diedros, ou seja, os ângulos *phi* e *psi*, calculados pelo mapa de Ramachandran, fornecido pelo servidor online Procheck [291] e resumidos na [Tabela 1](#).

Tabela 1 – Análise de Ramachandran para RSVmF e 4JHW

Regiões	Cristal 4JHW (449 res)	RSVmF (574 res)
Mais favorecidas	79,80%	84,60%
Adicionalmente permitidas	18,50%	10,20%
Generosamente permitidas	0,50%	3,10%
Não permitidas	1,20%	2,10%

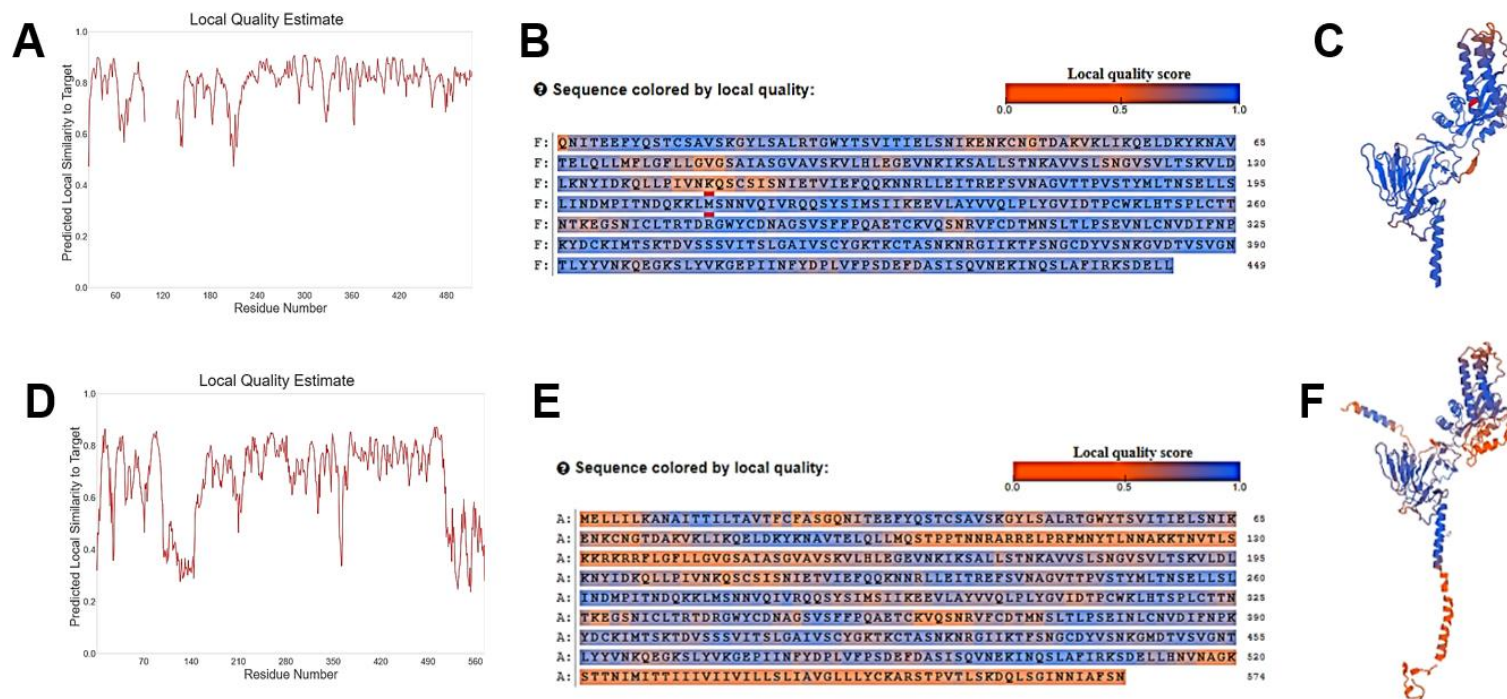
Fonte: O autor.

A análise dos ângulos *phi* e *psi* nos possibilita inferir que, a despeito de RSVmF possuir um maior número de resíduos, seus ângulos em regiões mais favorecidas apresentam um percentual acima de 80%. Por outro lado, a adição das regiões transmembrânicas, da porção C-terminal, bem como do peptídeo sinal (porção N-terminal) e do peptídeo 27, significam a inclusão de hélices e alças, e essas podem colaborar para o aumento na inclusão de regiões generosamente permitidas e não permitidas, devido, entre outros fatores, à flexibilidade das últimas, bem como à presença de eventuais “clashes” (erros e impedimentos estéricos) provenientes do processo de modelagem.

Assim, avaliar a qualidade da modelagem, considerando cada resíduo particularmente é importante, e nos revela informações mais apropriadas. Tal avaliação pode se basear em diferentes abordagens, as quais são possíveis de serem combinadas.

De funções de energia a potenciais estatísticos baseados em aprendizagem, cada abordagem apresenta suas vantagens e desvantagens. Aqui, utilizamos o servidor online Qmeans (“*Qualitative Model Energy ANalysis*” – Análise de energia baseada em modelo qualitativo) [292], o qual combina cálculos de potenciais estatísticos com estratégias físicas de avaliação de estrutura secundária e acesso ao solvente. O resultado é apresentado na [Figura 33](#).

Figura 33 – Avaliação de RSVmF por resíduo



As análises do servidor online QMEANS [292] se referem à qualidade da proteína modelada, em função da qualidade apresentada por cada um de seus resíduos componentes. O software considera tanto parâmetros estatísticos para elaboração de seus cálculos, como também analisa estrutura secundária em função de energia (no caso de se optar por utilizar os cálculos por DSSP) ou por predição baseada em algoritmos de redes neurais (no caso do PSIPRED). Em (A), (B) e (C) aparecem os resultados referentes ao cristal utilizado para modelagem – 4JHW – enquanto que em (D), (E) e (F), aqueles relativos à RSVmF. Em (A) e (D), no eixo X estão listados os resíduos da proteína em ordem numérica, e no eixo Y, revela-se a qualidade desses resíduos em função da estrutura tomada como “referência/molde”. Normalmente, resíduos que evidenciam uma pontuação abaixo de 0,6 podem ser classificados como de baixa qualidade. Em (B) e (E), a mesma análise de qualidade é feita, no entanto dessa vez é exibida a inicial de cada átomo. A pontuação varia de 0 a 1, sendo 0 (laranja) um indicativo de qualidade baixa e 1 (azul), boa qualidade. Em (C) e (F), apresentam-se as estruturas monoméricas de 4JHW e RSVmF, coloridas conforme disposição de qualidade mensurada em (B) e (E).

Fonte: o autor.

A avaliação do software QMEANS permite visualizar aquelas regiões para os quais o processo de modelagem é mais crítico. Essas regiões, evidenciadas na [Figura 33](#), principalmente em (B), (C), (E) e (F) são, justamente, aquelas que foram adicionadas no processo de modelagem (sequência 1-25 – o peptídeo sinal; sequência de 109-137, o peptídeo 27; sequência 530-550, os domínios transmembrânicos; sequência 551-574, a cauda citoplasmática).

O *software* I-Tasser utiliza um procedimento de modelagem chamado de “*Threading*” (Predição de Enovelamento de Proteínas), o qual se configura como um tipo de modelagem intermediária, a partir da combinação de diferentes algoritmos que abarcam uma gama de estratégias. Dentre essas, o alinhamento de múltiplos fragmentos com uma biblioteca de estruturas tridimensionais (3D) conhecidas, o agrupamento desses fragmentos, sob determinadas restrições, o que gera possíveis modelos, e o posterior refinamento desses modelos [\[293-297\]](#). No entanto, existem limitações no que tange à qualidade dos modelos preditos. A presença de múltiplos domínios na estrutura protéica de interesse, a quantidade de resíduos, a existência de bibliotecas com estruturas 3D resolvidas e de boa qualidade (alta resolução ~ 2 Å) para alinhamento com a sequência, o tipo de estrutura ao qual a sequência se refere, a posição da sequência na proteína, dentre outros.

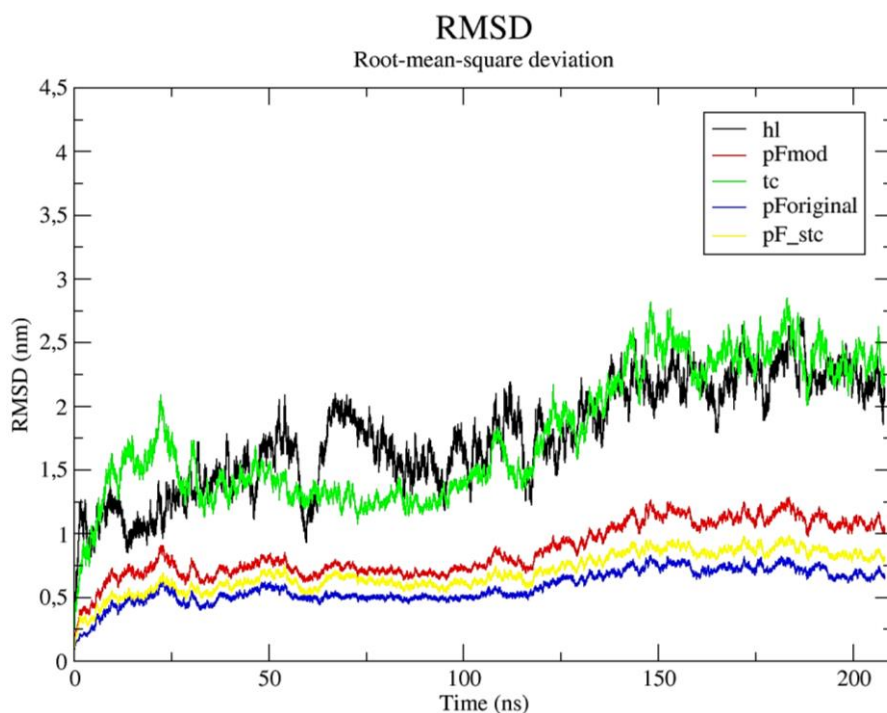
Assim, é inevitável que, para uma boa modelagem, seja necessária a combinação de bons alinhamentos de sequência com bons modelos estruturais. A combinação de domínios extracelulares, transmembrânicos e citoplasmáticos, no caso de RSV-F representa um gargalo para isso, não só por serem regiões distintas, como também pelo fato de se encontrarem em meios distintos. Isso influencia o processo de enovelamento, e, conseqüentemente, de modelagem. Ademais, o procedimento de cristalização de toda essa estrutura protéica é complexo [\[297,298\]](#), de forma que, normalmente, apenas os domínios extracelulares são alvos dele. Isso dificulta a disponibilidade de estruturas tridimensionais (3D) para o alinhamento, principalmente no que tange aos domínios transmembrânicos e à cauda citoplasmática. Além disso, a inclusão de regiões N e C-terminais representam um aumento de distância e, portanto, tendem a diminuir a precisão

dos algoritmos envolvidos. Logo, existe perda de precisão tanto para as funções de estimação de entropia, as quais tendem a favorecer as regiões centrais, onde normalmente se localizam os sítios ativos, quanto no processo de clusterização pelo algoritmo de Monte Carlo, o qual confere mais peso nos cálculos às regiões de maiores domínios [293-295], nesse caso, os domínios extracelulares, representados pela sequência de 27-529.

Uma maneira de refinar os resultados da modelagem, é submeter o modelo à simulação por dinâmica clássica (DM), a fim de observar o comportamento da proteína quando em solução aquosa. A presença de solvente tende a acomodar melhor a proteína e a melhorar a disposição de ângulos desfavoráveis no mapa de Ramachandram. RSVmF foi submetida a uma simulação de DM de 200 nanossegundos (ns) e as análises visaram ponderar sua estabilidade conformacional e alteração de estrutura secundária com o passar do tempo, bem como investigar se as regiões problemáticas da modelagem se manteriam instáveis. De posse desses dados, então, podemos estipular a importância da participação dessas regiões na etapa seguinte, de interação com os ligantes.

Assim, foram realizadas as análises de desvio quadrático médio [Figura 34], de flutuação quadrática média [Figura 35], de estrutura secundária [Figura 36], bem como dos *eigenvalues* (autovalores) da matriz de covariância [Figura 37], as quais são apresentadas a seguir:

Figura 34 – Desvio Quadrático Médio (RMSD) de RSVmF



1nm= 10⁻⁹m.

Legenda:

hl: rmsd da hélice lateral, ou seja, o peptídeo sinal;

pForiginal: rmsd da proteína original;

pF_stc: rmsd de RSVmF sem o triple coil, isto é, a região transmembrânica e C-terminal;

pFmod: rmsd de RSVmf completa;

tc: rmsd do *triple-coil*, isto é, a porção transmembrânica e a cauda C-terminal.

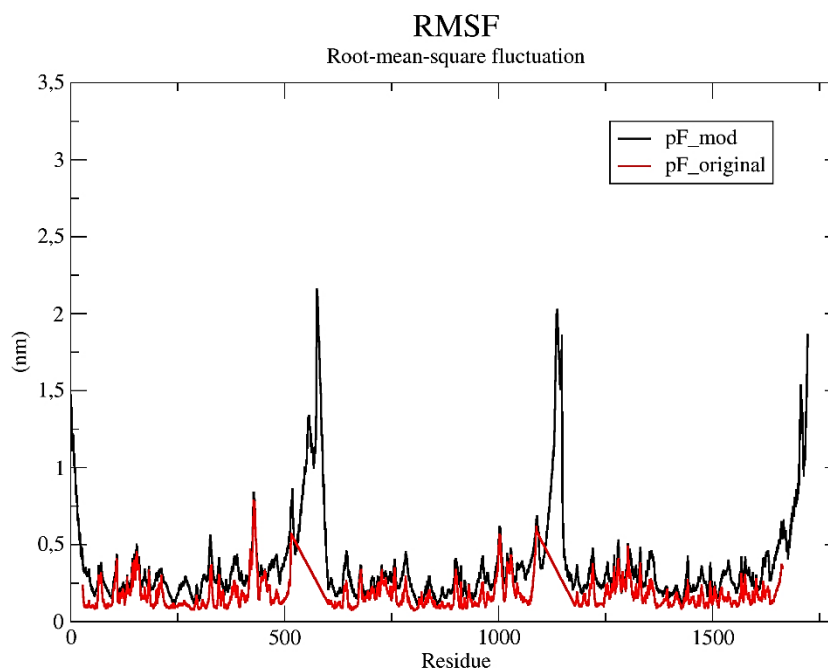
Fonte: o autor.

O RMSD (*Root-mean-square deviation* – desvio quadrático médio) mede a distância entre duas estruturas superpostas de proteínas, no caso, a proteína modelada e a proteína-molde, o cristal 4JHW. Foi calculado a partir do módulo *g_rms* do Gromacs 4.6.7 [237]. A análise do RMSD possibilita visualizar o quanto a estrutura da proteína modelada se distanciou da estrutura molde, bem como avaliar a estabilidade da proteína modelada ao longo do tempo.

Para avaliação do RMSD foi utilizado o desvio relativo ao carbono-alfa, pois esse átomo representa a maior taxa de variação planar do aminoácido. Assim, é possível

observar que RSVmF, no gráfico referida como “*pFmod*” (curva vermelha) apresentou maior desvio da estrutura molde (“*pForiginal*”, curva em azul) isto é, do cristal 4JHW, nos primeiros 40 nanosegundos (ns), atingindo um valor de desvio de, aproximadamente, 0,3 nm. A instabilidade da proteína volta a subir após 110 ns, o que se mantém até 200 ns. Em busca de um maior entendimento dessa variação, fragmentamos a análise, incluindo a observação das regiões adicionadas durante o processo de modelagem, isto é, o peptídeo sinal (sequência de 1 a 25, “*hl*”), e a porção final de RSVmF que engloba os aminoácidos 516-574, 1090-1148, e 1664-1722 (a estrutura característica de *triple coiled coil*, e, por isso, denominamo-no “*tc*”). A fragmentação nos possibilitou visualizar que, de fato, as regiões adicionadas pela modelagem, ou seja, “*hl*” e “*tc*” compõem as maiores fontes de variação conformacional da proteína. Uma vez que se exclui “*tc*” da análise de RMSD (“*pF_stc*”, curva amarela), a taxa de desvio diminui para 0,1 nm. Visando confirmar a influência das diferentes partes na estrutura conformacional de RSVmF ao longo da trajetória de simulação, foi realizada a análise de flutuação quadrática média (RMSF):

Figura 35– Flutuação quadrática média (RMSF) de RSVmF



Legenda:
 pF_original: proteína F molde;
 pF_mod: RSVmF.
 Fonte: o autor.

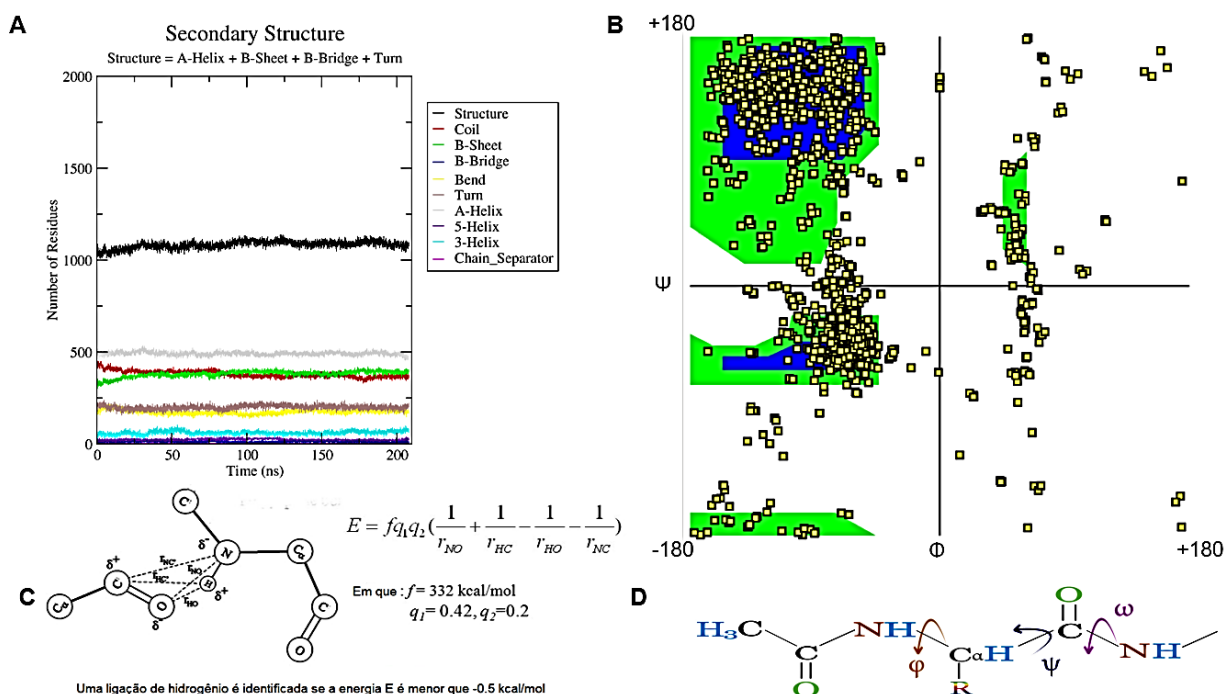
O RMSF (*Root-mean-square fluctuation* – flutuação quadrática média) evidencia a flutuação do movimento dos átomos no decorrer da trajetória de Dinâmica Molecular, a partir da superimposição da proteína modelada sobre a proteína-molde, o que mostra a flexibilidade do sistema.

O módulo *g_rmsf* do Gromacs 4.6.7 [237] foi usado para o cálculo a partir da flutuação dos carbonos- α . A leitura do gráfico nos permite observar que, na proteína modelada – RSVmF – os resíduos de maior flutuação se localizam nas sequências 1-30, 516-600, 1090-1200, e, ao final, 1650-1722. É possível observar também que, esses resíduos não estão incluídos na estrutura da proteína usada como *template*, sendo essa

a razão da existência de uma linha diagonal na “curva” correspondente à sua taxa de flutuação (curva vermelha). As sequências correspondentes aos maiores RMSF se referem às regiões “tc” (516-674, 1090-1148, 1664-1722), como também “hl” (1-25, 575-601, 1149-1175).

Verificar a taxa de desvio e flutuação, quer seja em termos de tempo ou de localização de resíduos de aminoácidos, de uma proteína modelada em relação à proteína-molde é importante. Isso nos possibilita não só visualizar o comportamento em termos de estabilidade da proteína ao longo do tempo, como também verificar quais resíduos colaboram com maior ou menor ênfase nesse sentido. Mas, embora a avaliação do RMSD e do RMSF nos forneçam esse panorama de desvio e flutuação-resíduo sobre a proteína modelada, somente esses parâmetros não são suficientes. É necessário garantir que a estrutura secundária da proteína se mantenha íntegra, isto é, folha- β continue sendo folha- β e α -hélice continue sendo α -hélice, basicamente. Afinal, a interação de cadeias polipeptídicas e o enovelamento protéico corretos ocorrem a partir do arranjo tridimensional determinado pela estrutura secundária. Nesse sentido, a variação da estrutura secundária de RSVmF foi avaliada [\[Figura 36\]](#).

Figura 36 – Estrutura secundária RSVmF



A variação da conformação da estrutura secundária em RSVmF. Em (A), avaliação pelo módulo DSSP (*Definition of Secondary Structure of Proteins* – Definição de Estrutura Secundária em Proteínas); em que o eixo X representa o tempo em ns, e o eixo Y representa o número de resíduos. Em (B), a análise pelo mapa de Ramachandran; o qual verifica a mudança (rotação) dos ângulos phi (no que seria um “eixo X”) e psi (no que seria um “eixo Y”) em torno do plano representado pelo carbono alfa. Em (C), exibe-se a fórmula de cálculo da ligação de Hidrogênio, baseada na Lei de Coloumb e utilizada pelo software DSSP, no qual os termos “ q_1 ” e “ q_2 ” se referem às cargas dos átomos envolvidos, em coloumbs, e “ r ” se refere à distância entre os átomos, a qual comumente é expressa em angstroms ($1\text{\AA} = 10^{-10}\text{m}$) ou nanômetros ($1\text{nm} = 10^{-9}\text{m}$). Em (D), aparece o princípio utilizado pelo mapa de Ramachandran, isto é, o cálculo dos ângulos diedrais, também chamados de ângulos de torção/rotação Ψ e Φ . (C) foi adaptado [301] e (D) [302].

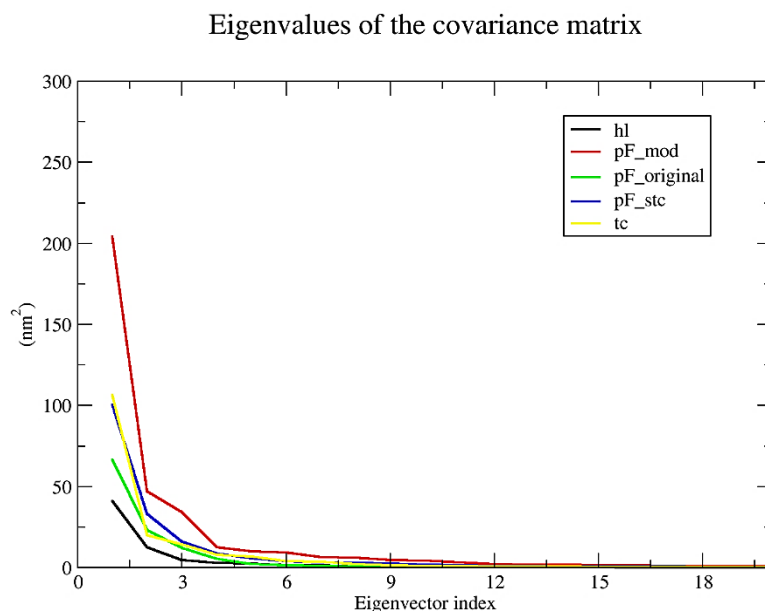
Fonte: o autor.

O módulo *do_dssp* [244] foi empregado na avaliação da estrutura secundária, com base na cadeia principal. Ele caracteriza a estrutura secundária em função do padrão das ligações de hidrogênio, baseado em termos eletrostáticos, como mostrado na Figura 30 (C). Assim, existem oito tipos de estruturas secundárias, hierarquizadas em três categorias principais: hélices (*A-helix*, *5-helix* e *3-helix*), folhas (*B-sheet* e *B-bridge*) e loops (*Turns*, *Bends* e *Coils*), as quais diferem entre si por questões energéticas e de

distância. A análise do gráfico nos mostra que em RSVmF existe um predomínio de conformação α -hélice, devido à inclusão do peptídeo sinal e dos domínios transmembrânicos, os quais possuem esse tipo de estrutura secundária. Além disso, também é possível visualizar uma pequena alternância entre a conformação de *coil* (curva vermelha) e folha-beta (β -sheet, curva verde). Essa alternância se deve ao movimento da proteína durante a DM, uma vez que a conformação de *coil* implica em interações eletrostáticas de alcance mais curto, enquanto que as interações de folhas-beta são de alcance maior [299,300], o que provavelmente é causado pelo efeito hidrofóbico no interior da proteína. A análise do gráfico de Ramachandran evidencia que, aproximadamente, 88% dos resíduos da proteína se encontram dentro das regiões favoráveis e/ou permitidas.

Consecutivamente, foi realizada a análise dos componentes principais (*Principal component Analysis* – PCA), baseando-se no movimento dos Carbonos- α , e pela utilização dos módulos *g_covar* e *g_anaeig* do Gromacs 4.6.7 [243,244,237] .

Figura 37– Eigenvalues da matriz de covariância de RSVmF



Legenda:

hl: eigenvalue referente às “hélices laterais”, isto é, o peptídeo sinal (=looplat);

pF_mod: eigenvalue referente à RSVmF;

pF_original: eigenvalue referente à proteína-molde;

pF_stc: eigenvalue referente à RsmF sem o *triple coil*;

tc: eigenvalue referente ao *triple coil*, isto é, a porção transmembrânica e a cauda C-terminal

Fonte: o autor.

A PCA constitui um conjunto de técnicas matemáticas de transformações lineares para redução de grandes volumes de dados a volumes menores, os quais sejam capazes de descrever as características essenciais daquele sistema. No caso da Dinâmica Molecular, a trajetória de uma proteína se reflete em um conjunto de coordenadas atômicas em 3 dimensões, provenientes da resolução de equações de movimento da mecânica clássica ao longo do tempo de simulação. Para execução da PCA, é necessário remover os movimentos de rotação e translação dos átomos, e, para isso, realiza-se uma normalização dos dados, baseando-se na estrutura da proteína *template* (o cristal) [245]. Após a normalização, uma matriz de covariância da média das posições de x, y e z de

todos os átomos do sistema ao longo da trajetória é construída. A partir dessa matriz de covariância, métodos matemáticos de diagonalização (isto é, construção de uma matriz na qual os elementos não-nulos se encontrem apenas na diagonal principal) podem ser aplicados. Esses métodos se baseiam no princípio da ortonormalização (vetores unitários e ortogonais), de forma a gerar uma matriz diagonalizada com intuito de reduzir a covariação e a hierarquizá-la, tornando apenas as grandes variações como ‘principais’ [246]. As colunas dessa matriz diagonalizada representam os *eigenvectors*, os quais correspondem ao deslocamento dos átomos no espaço e estão associados aos *eigenvalues*, os quais mostram a magnitude desse deslocamento em nm². Logo, a PCA reduz o movimento dos átomos de RSVmF ao longo da simulação de DM a alguns vetores principais – *eigenvectors* – os quais podem ser mensurados por seus valores correspondentes, os *eigenvalues*.

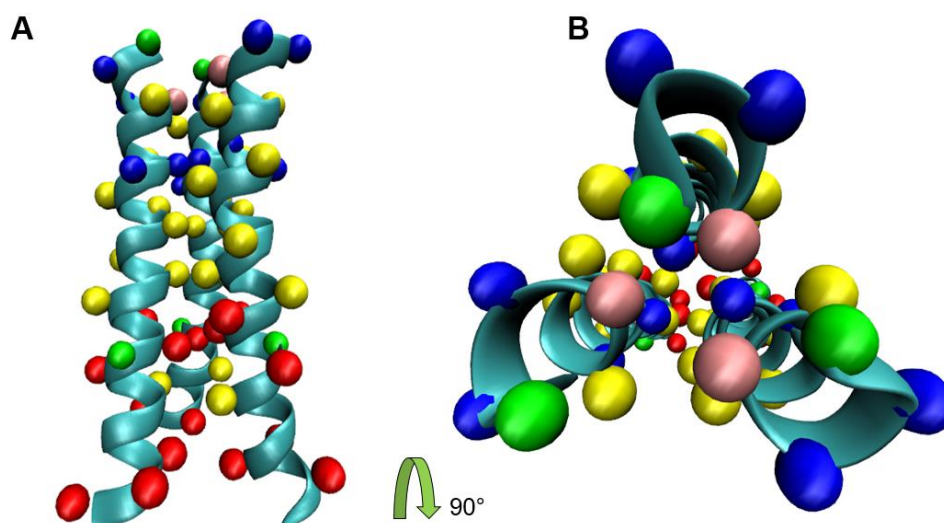
A análise dos *eigenvalues* evidencia que a trajetória da simulação de DM pode ser reduzida a 20 *eigenvectors*, sendo que desses, os quatro primeiros são os mais significativos. Fragmentando essa avaliação, é possível observar a colaboração de “h” e “tc” na dimensão dos *eigenvalues*, e o último, se removido for, reduz a magnitude do deslocamento (*eigenvalue*) do primeiro *eigenvector* de 200 para 100 nm². Os dados de *eigenvalues* corroboram os achados do RMSD e RMSF, indicando que as regiões de maiores flutuações são aquelas abarcadas pelo peptídeo sinal, o que inclui as sequências (1-25, 575-601, 1149-1175) e também pela região característica de *triple coil* (516-674, 1090-1148, 1664-1722).

A primeira descrição cristalográfica da proteína F no seu estado de pré-fusão foi feita por McLellan e colegas em 2013 [247] associada ao anticorpo D25, e depositada no *Protein Data Bank* sob os códigos de identificação PDB 4JHW (proteína + anticorpo) e 4JHA (anticorpo). Essas descrições cristalográficas primaram em descrever o modo de ligação do anticorpo à proteína. Outrossim, a proteína em seu estado pré-conformacional é meta-estável, isto é, não está equilibrada termodinamicamente, e, portanto, altera seu estado pré-conformacional para pós-conformacional facilmente. Elaborar o cristal com o anticorpo foi a estratégia encontrada por esses pesquisadores para buscar estabilizar a

proteína, de forma a reproduzir seu estado pré-conformacional, o que não necessariamente implica em um cristal de alta resolução, devido às dificuldades intrínsecas ao procedimento e às características peculiares da proteína. Desde então, diversas versões de proteína F pré-fusão, associadas aos mais variados tipos de anticorpos e pequenos ligantes têm sido depositadas nesse mesmo banco de dados. No entanto, todas elas, até o momento da escrita desta tese, apresentam-se incompletas, principalmente no que tange à porção final da proteína F, isto é, a porção que abrange dos aminoácidos 513-574 e inclui o domínio transmembrânico e a cauda citoplasmática. Nosso objetivo na modelagem, foi então, completar a estrutura com essa parte além da inclusão do peptídeo sinal, e do peptídeo 27, e, com auxílio da dinâmica molecular, tentar estabilizar a proteína como um todo.

Os nossos achados evidenciam que 200 ns não foram suficientes para estabilizar a estrutura completa da proteína F, principalmente no que tange à região inclusa, relativa ao domínio transmembrânico e à cauda citoplasmática. A região transmembrânica, cuja característica estrutural de α -hélice e de alternância descontínua entre aminoácidos hidrofóbicos e aminoácidos polares, assemelha-se à estrutura de um (*triple*) *coiled coil* [Figura 38]:

Figura 38 – Porção final de RSVmF e seus aminoácidos principais



A estrutura característica de (*triple*) *coiled coil* de RSVmF, representativa da sequência de aminoácidos – LHNVNAGKSTNIMITIVIVILSLIAVGLYC –. É possível observar que sua orientação é paralela. Em vermelho estão representadas as leucinas, em amarelo as isoleucinas, em verde as serinas, em rosa as asparaginas e em azul as treoninas. O VMD [239] foi usado para elaboração e edição da imagem.

Fonte: o autor.

Além dessa região existem também as regiões HRA e HRB, as quais se caracterizam como *coiled coil*. Regiões de *coiled coil* consistem em um desafio peculiar para simulações de DM, apesar de sua relativa simplicidade estrutural [248-250]. A alternância de resíduos hidrofóbicos e polares confere ao *coiled coil* uma característica anfipática, isto é, a possibilidade de interações hidrofóbicas e polares [251]. No contexto da DM, a adição de solvente (em nosso caso, água) e de contra-íons (Cl^-) para a simulação, pode ter causado efeitos consideráveis sobre essa estrutura, desestabilizando as interações hidrofóbicas da parte central do *coil*.

Diversos trabalhos têm trazido um maior entendimento sobre o comportamento da estrutura de *coiled coils* durante as simulações. Um estudo recente de 2017 [248] foca especificamente no papel do aminoácido asparagina (ASN) com relação à manutenção e a estabilidade da conformação da estrutura de um dímero paralelo de *coiled coil*, a partir

da combinação de análise experimental por Ressonância Magnética Nuclear e DM. Esse estudo mostra que a conformação da ASN muda em uma escala de milisegundos, com sua cadeia lateral apontando para dentro ou para fora da estrutura do *coiled coil*. Outro estudo [249] indica que a orientação mais favorável na formação de homodímeros de *coiled coil* é a antiparalela. Nesse mesmo estudo, a estabilidade dos *coils* é atingida com 10 ns de simulação. No entanto, cabe enfatizar que nesses dois estudos, são analisados apenas pequenos peptídeos.

Em um outro estudo, desta vez com a proteína F do RSV [252] os pesquisadores propõem o processo de modelagem da proteína, abarcando a transição de sua conformação do estado de pré-fusão para pós-fusão. Para isso, utilizam um protocolo de DM-orientada a fim de estabilizar a estrutura da proteína. A estabilidade do fragmento referente ao *coiled coil* nesse estudo é obtida após 12 ns de simulação, em um modelo mais simplístico cujo parâmetro principal é força mecânica.

A DM torna possível a visualização de um panorama atômico, energético e estrutural de todo processo de mudanças e conformações protéicas, os quais os experimentos de bancada só vislumbram. Ela facilita a visualização de que uma proteína não possui uma estrutura, mas sim um conjunto de conformações. Entretanto, apesar de toda robustez do método, ele também possui falhas. Seja pelas inúmeras aproximações que tornam os cálculos de trajetória possíveis, seja pela capacidade computacional requerida, seja pelo tempo. Quanto maior um sistema, maiores os cálculos, maior capacidade computacional é requerida e, conseqüentemente, mais tempo de simulação é exigido. Isso pode ser simplificado, pelo uso de sistemas *coarse-grained*, nos quais agrupamento de átomos são representados por *beads*, ou pela remoção do solvente [253,254]. Nós optamos por manter todos os átomos e o solvente justamente para podermos obter uma melhor acurácia da análise. Entretanto, a região transmembrânica, como o nome sugere, é uma região que fica imersa em uma membrana, e, no caso do RSV, essa membrana é proveniente da célula, logo, é uma bicamada lipídica. Isso exigiria a combinação de simulações *coarse-grained* para a membrana e a região transmembrânica, e atomística (todos os átomos) para o restante da proteína, o que, por

ora, estaria um pouco além do escopo do trabalho. Seria incomum se a região transmembrânica estabilizasse em um meio aquoso. Como a região transmembrânica e a cauda citoplasmática não têm implicações diretas no processo de mudança conformacional de pré para pós-fusão, nós escolhemos mantê-la na estrutura da proteína e prosseguirmos com as análises.

4.2 Docking e Dinâmica Molecular RSVmF-quercetina, RSVmF-Q1 e análises

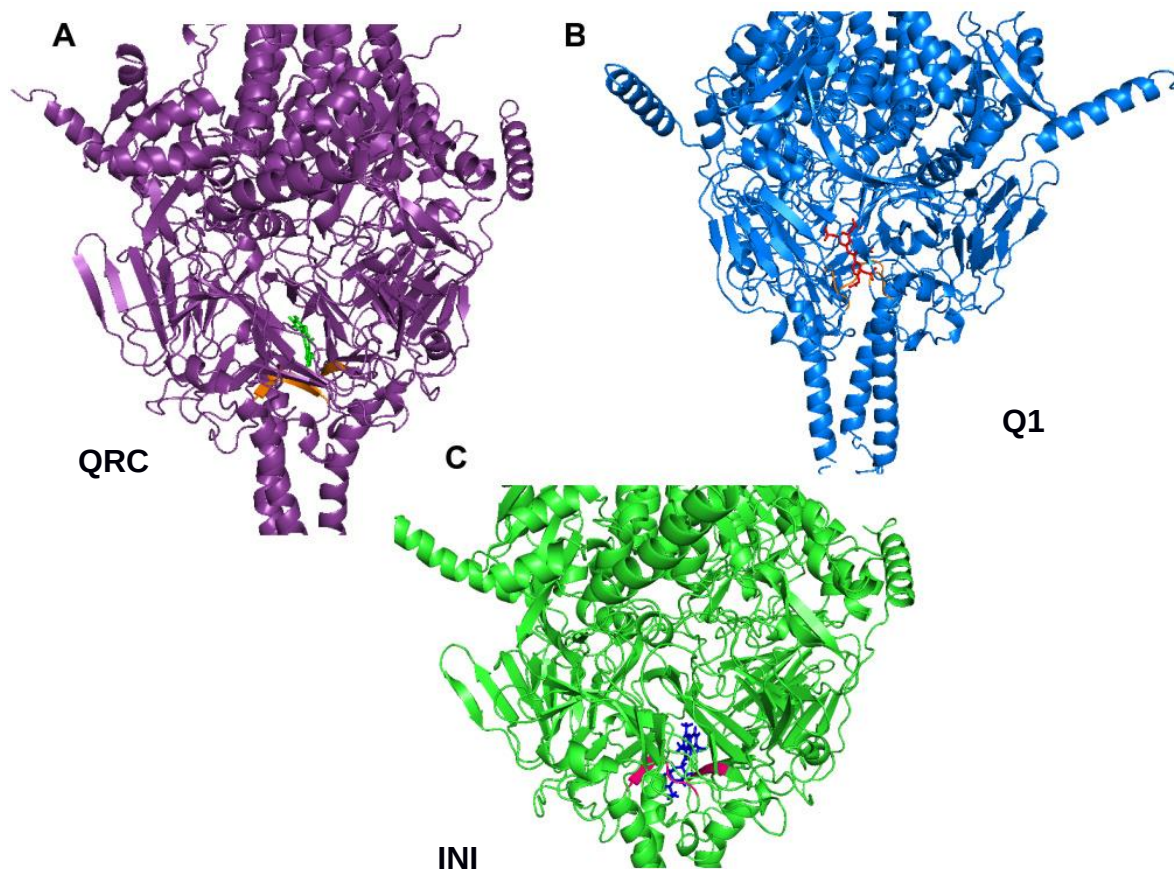
O efeito anti-RSV da quercetina e de seu derivado quercetina penta-acetilada (Q1), e, principalmente, a capacidade de inibição dessas moléculas no protocolo virucida², sugere que elas possam interagir com as proteínas de superfície do RSV, dentre as quais, e, especialmente, a proteína F. Assim, em nosso **segundo passo** decidimos investigar a interação da quercetina em complexo com a cavidade trissimétrica de RSVmF por meio de *Docking* e DM. Posteriormente, verificamos o efeito da acetilação da quercetina, o que resultou no Q1, na interação do complexo RSVmF-Q1. O complexo RSVmF-JNJ-2408068/R170591 foi usado como controle.

Os experimentos de *docking* indicaram que tanto a quercetina, quanto o Q1 e o inibidor preenchem completamente a cavidade previamente descrita [Figura 39]. e foram caracterizados basicamente pela mesma energia de interação (em torno de -6.5 kcal/mol).

²Existem basicamente três tipos de protocolo para avaliação preliminar de atividade antiviral de uma molécula:

- a) Protocolo de pré-tratamento: As células são incubadas com a molécula antes de serem infectadas pelo vírus.
 - b) Protocolo de pós-tratamento: As células são infectadas pelo vírus, e posteriormente incubadas com a molécula.
 - c) Protocolo virucida: O vírus é incubado com a molécula, antes da infecção das células.
- A combinação desses três protocolos permite uma avaliação inicial do mecanismo pelo qual ocorre a ação antiviral de uma molécula de interesse.

Figura 39 – Docking molecular dos complexos RSVmF-quercetina, RSVmF-Q1 e RSVmF-JNJ



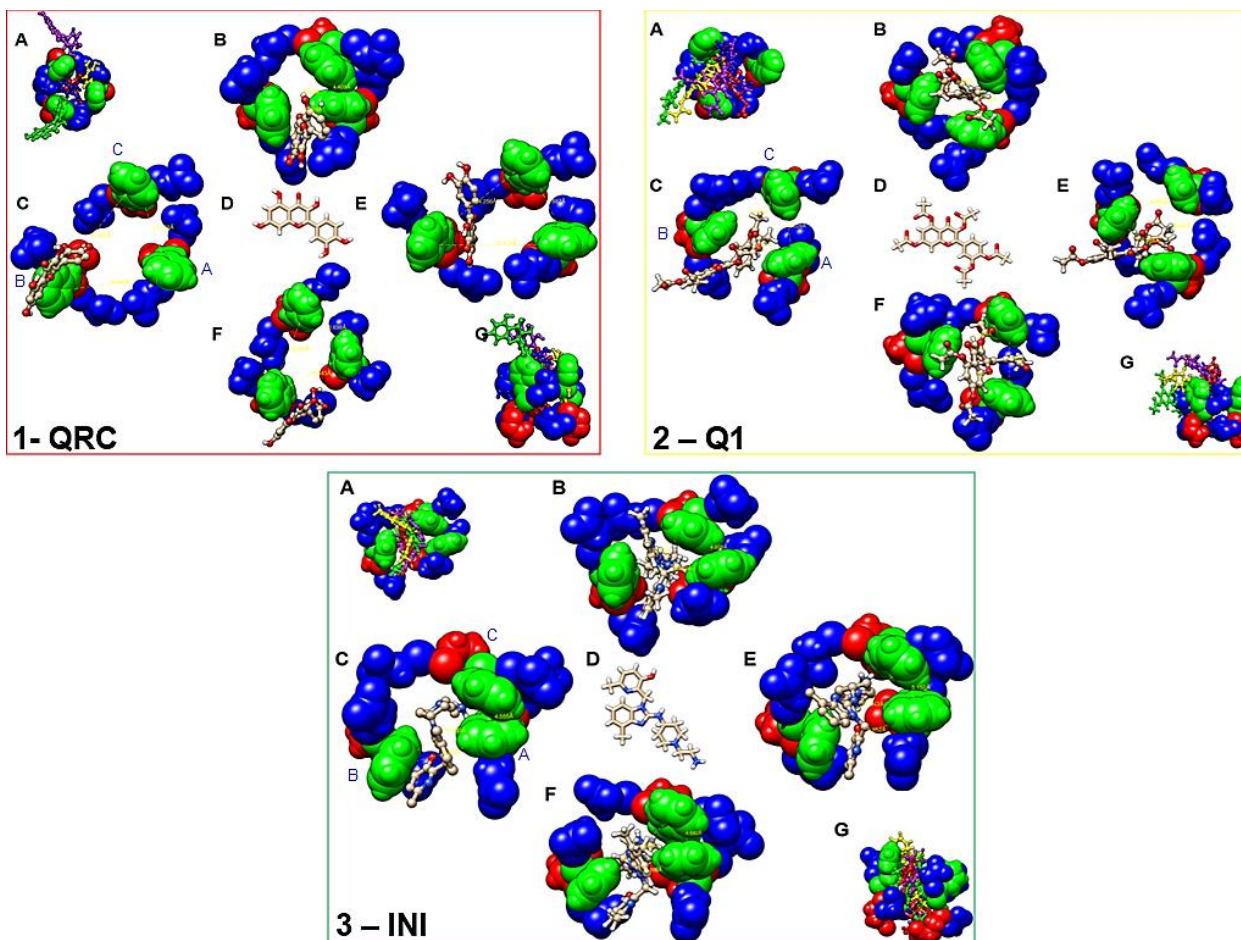
RSVmF em complexo com a quercetina; B) RSVmF em complexo com Q1; C) RSVmF em complexo com inibidor JNJ- 2408068. Todos os ligantes foram alocados no bolsão trissimétrico e negativamente carregado dentro da cavidade central da RSVmF para os experimentos de *Docking Molecular* e obtiveram aproximadamente a mesma energia de interação $\sim -6,5$ Kcal/mol.

QRC = quercetina, Q1= quercetina pentacetilada e INI = inibidor JNJ2408068.

Fonte: o autor.

A fim de superar as limitações do *docking* e explorar o panorama conformacional exibido por cada um dos ligantes, os complexos RSVmF-quercetina, RSVmF-Q1, RSVmF-JNJ2408068 foram simulados através de DM por um tempo total de 20 ns, conforme ilustra a [Figura 40](#):

Figura 40 – Simulação de DM: RSVmF-QCT, RSVmF-Q1, RSVmF-JNJ



A DM nos possibilita visualizar o comportamento de 1– quercetina, 2 – Q1, e 3 – JNJ2408068 (inibidor, usado como controle) ao longo de uma trajetória de simulação de 20 ns. A figura ilustra os três ligantes e suas posições mais representativas na cavidade, formada por resíduos de fenilalanina (488, em verde, nas cadeias A, B e C), de aspartato (486 e 489, em vermelho, nas cadeias A, B e C) e glutamato (487, em azul, nas cadeias A, B e C). Para todos os três, tanto em (A) quanto em (G), podemos observar a alteração de seu posicionamento em relação à cavidade – apenas nesses casos (A) e (G), tomada como “estática”, isto é, o modelo inicial, anterior à DM, sem modificações de movimento, apenas para enfatizar a movimentação do ligante – sob uma visão superior e frontal, respectivamente. Em (B) aparecem todos os primeiros *frames* (conformações, para $t=1\text{ns}$). Em (C), estão discriminadas as cadeias da proteína, em letras azuis. Em (D), são representadas as moléculas em duas dimensões (2D) de cada um dos ligantes. As conformações exibidas em (C), (E) e (F) são referentes à análise de *cluster* da trajetória. Assim, estão representados os três primeiros *frames*, equivalentes aos três primeiros *clusters* hierarquizados da simulação. Logo, (C) retrata a primeira conformação mais representativa, (E), a segunda, e (F), a terceira.

QRC = quercetina, Q1= quercetina pentacetilada e INI = inibidor JNJ2408068.

Obs. As imagens acima ilustradas podem ser vistas com maiores detalhes nos apêndices [II](#), [III](#) e [IV](#).

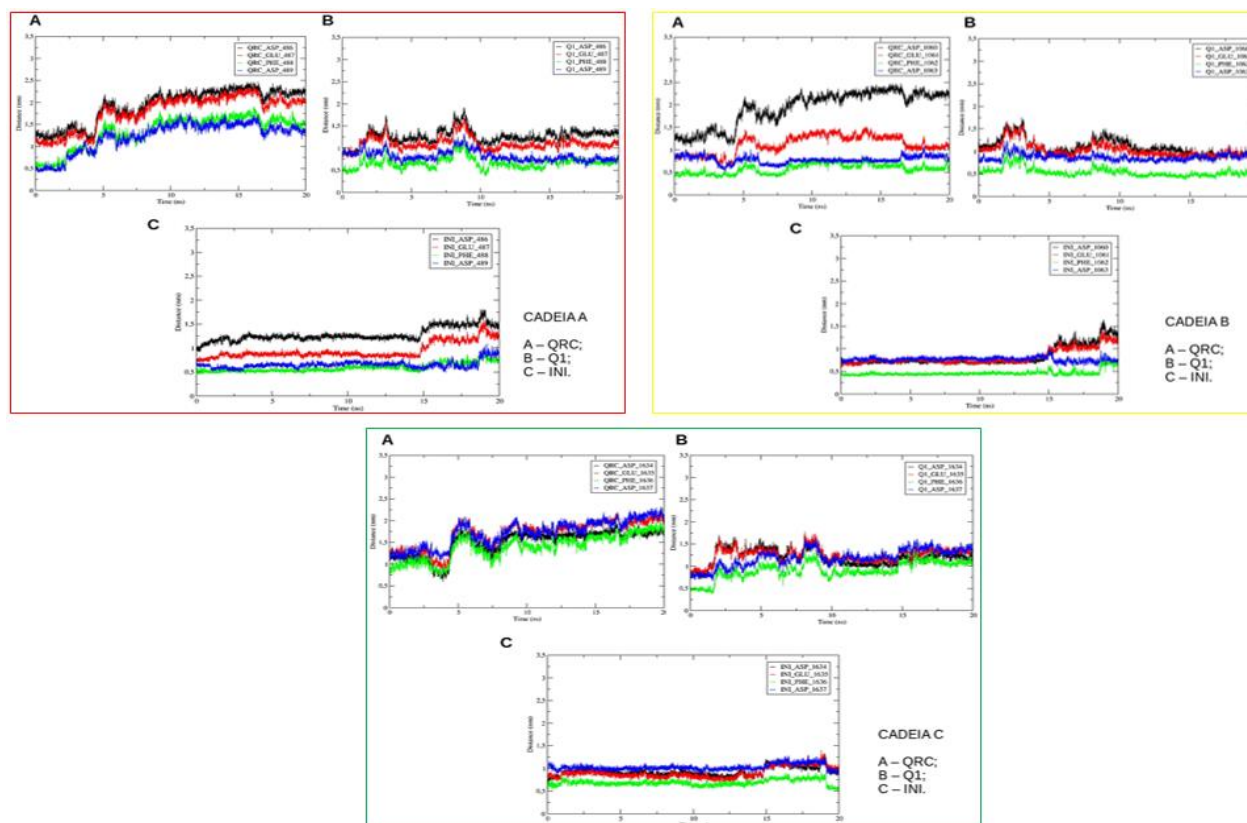
Fonte: o autor.

A avaliação da [Figura 40](#) nos possibilita enxergar comportamentos distintos para os três ligantes. Além da evidente diferença na abertura da cavidade, em relação não só às várias conformações representadas, como também às moléculas em questão, é notório o comportamento evasivo da quercetina. Além disso, é possível visualizar que apesar do distanciamento da cavidade, há uma interação mais estabelecida da quercetina com uma cadeia, isto é, a cadeia B. Após a realização da análise de *cluster* e hierarquização dos principais *frames*, fica evidente que a principal conformação da quercetina é aquela apresentada em (C), ou seja, quando ela praticamente saiu da cavidade. Quanto ao Q1, ainda que seu comportamento ao longo da trajetória não reflita uma total inserção dentro da cavidade, ele permanece ali, e também interage mais com a cadeia B. A análise do inibidor JNJ2408068 mostra que ele se conserva dentro da cavidade. É possível observar também que a quercetina favorece a abertura da cavidade, enquanto que o Q1 e o inibidor JNJ2408068 a mantêm mais fechada. Em relação ao *frame* inicial, exibido em (B), nota-se que o inibidor JNJ2408068 é o que causa um maior contraste na estrutura da cavidade em sua conformação mais representativa (C). Ademais, constata-se o movimento dos três ligantes, evidenciando a maior flexibilidade do Q1 e do inibidor JNJ2408068.

Embora uma imagem valha mais do que mil palavras, cálculos são necessários para explicitar o que pode ser visualizado na [Figura 40](#). Assim, foram realizadas investigações de trajetória das simulações de DM dos complexos RSVmF-quercetina, RSVmF-Q1 e RSVmF-JNJ2408068 para averiguação de distância, de ligações de hidrogênio e de energia.

A avaliação da distância ilustra a evolução tempo-dependente da distância entre o centro-de-massa dos três ligantes e o centro-de-massa dos resíduos da cavidade central, isto é, os resíduos de fenilalanina, de aspartato e de glutamato. Ela é exposta na [Figura 41](#)

Figura 41 – Variação da distância RSVmF-ligantes



Variação da distância entre o bolsão trissimétrico, formado pelos resíduos de aspartato, glutamato e fenilalanina dentro da cavidade central de RSVmF e os ligantes quercetina, Q1 e o inibidor JNJ-2408068, ao longo de uma simulação de 20 ns. Nessa análise as cadeias não foram renumeradas, então os resíduos 486-489 na cadeia A equivalem aos resíduos 1060-1063 na cadeia B e a 1634-1637 na cadeia C, seguindo a mesma ordem: ASP-GLU-FEN-ASP. A distância exibida é referente aos resíduos componentes de cada cadeia. Em (1), (2) e (3) são representadas as cadeias A, B e C, respectivamente. Em (A) aparecem as distâncias referentes à quercetina, em (B), ao Q1, e em (C), ao inibidor JNJ-2408068.

Fonte: o autor.

Obs. As imagens acima ilustradas podem ser vistas com maiores detalhes nos apêndices [V](#), [VI](#) e [VII](#).

A análise de distância ratificou o comportamento das três moléculas, representado na [Figura 40](#). A quercetina continua a revelar uma clara propensão ao escape daquela cavidade. No entanto, esse cálculo nos permitiu ir um pouco além. A quercetina se afasta mais pronunciadamente da cadeia A e da cadeia C, em relação aos resíduos de aspartato e glutamato (486.A, 487.A) e de todos os resíduos da cadeia C (1634.C-1637.C). É interessante notar que a quercetina se distancia da cadeia B, porém com uma menor intensidade para o glutamato (1061.B). Por outro lado, ela praticamente não se distancia dos resíduos de aspartato e fenilalanina (1062.B, 1063.B). Quanto à movimentação do Q1, existe uma tendência similar de distanciamento para as três cadeias, e, assim como a quercetina, esse afastamento é mais notório para A e C. Cabe notar que a variação na distância do Q1-fenilalanina é sempre a menor, comparada aos outros resíduos. No que concerne ao inibidor JNJ2808068, há uma maior distanciação em relação aos resíduos aspartato e glutamato da cadeia A (486.A e 487.A, respectivamente). Essa tendência de afastamento diminui na cadeia B. Quanto à cadeia C, existe pouca variação para todos os resíduos pertencentes a ela (1634.A-1637.A). Ainda com relação a esse inibidor, a distância JNJ2808068-fenilalanina é também a menor, relativa aos outros resíduos da cavidade.

Avaliadas as posições, conformações e distância, a próxima análise foi a investigação das ligações de hidrogênio. Ela nos revelou que a quercetina realiza somente uma ponte de hidrogênio com essa cavidade da proteína, a qual, sem grandes surpresas, é estabelecida justamente com o aspartato (1063.B), para o qual o afastamento na análise de distância foi mínimo. Com relação ao inibidor JNJ-2408068, há o estabelecimento de três ligações de hidrogênio com as cadeias A e B, e uma dessas ligações se manteve por mais de 73% do tempo da simulação ([Tabela 2](#)). Não foi possível verificar ligações de hidrogênio entre o Q1 e os resíduos da cavidade, o que sugere outros tipos de interações como forças de Van der Waals.

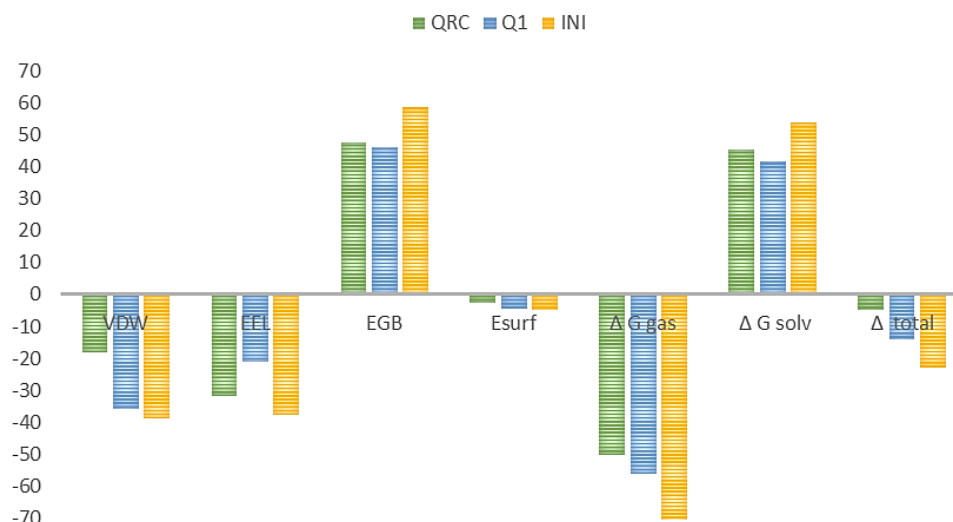
Tabela 2 – Persistência das ligações de hidrogênio entre RSVmF e quercetina/ligante-referência para uma simulação de 20 ns

Resíduo da Proteína	Ligante	Persistência
ASP489 cadeia B	QRC	54.15
GLU487 cadeia A	JNJ-2408068/R170591	41.95
ASP486 cadeia A	JNJ-2408068/R170591	73.29
LYS498 cadeia B	JNJ-2408068/R170591	29.43

ASP= Aspartato; GLU= Glutamato, LYS = Lisina; QRC = Quercetina .
 Fonte: o autor.

Por fim, o método de MM-GBSA foi empregado nos cálculos de energia envolvidos na interação dos ligantes com a proteína. A estratégia e os cálculos abordados pelo método de MM-GBSA são brevemente explanados no [Apêndice VIII](#). Os resultados referentes à análise de MM-GBSA das trajetórias de simulação dos complexos RSVmF-Quercetina, RSVmF-Q1 e RSVmF-JNJ são apresentados no [Gráfico 1](#).

Gráfico 1 – MM-GBSA: Quercetina, Q1 e JNJ-2408068



Avaliação dos diferentes tipos de energia envolvidos na interação da Quercetina, do Q1 e do inibidor JNJ-2408068 com a cavidade central de RSVmF. Os índices apresentados no gráfico correspondem aos valores obtidos pela subtração das energias (complexo – receptor – ligante), que representam a média da interação das energias entre RSVmF-ligantes.

VDW = interações de Van der Waals entre proteína-ligante;

EEL = interações eletrostáticas entre proteína-ligante;

EGB = a contribuição eletrostática para a energia livre de solvatação calculada por GB;

Esurf = a energia da área de superfície - SA (energia de solvatação não polar),

ΔG_{gas} = energia livre de Gibbs na fase gasosa;

ΔG_{solv} = EGB + Esurf;

Δ_{total} = resultado do cálculo envolvendo todos os termos acima (apêndice VIII)

Unidade de todos os termos acima = kcal/mol.

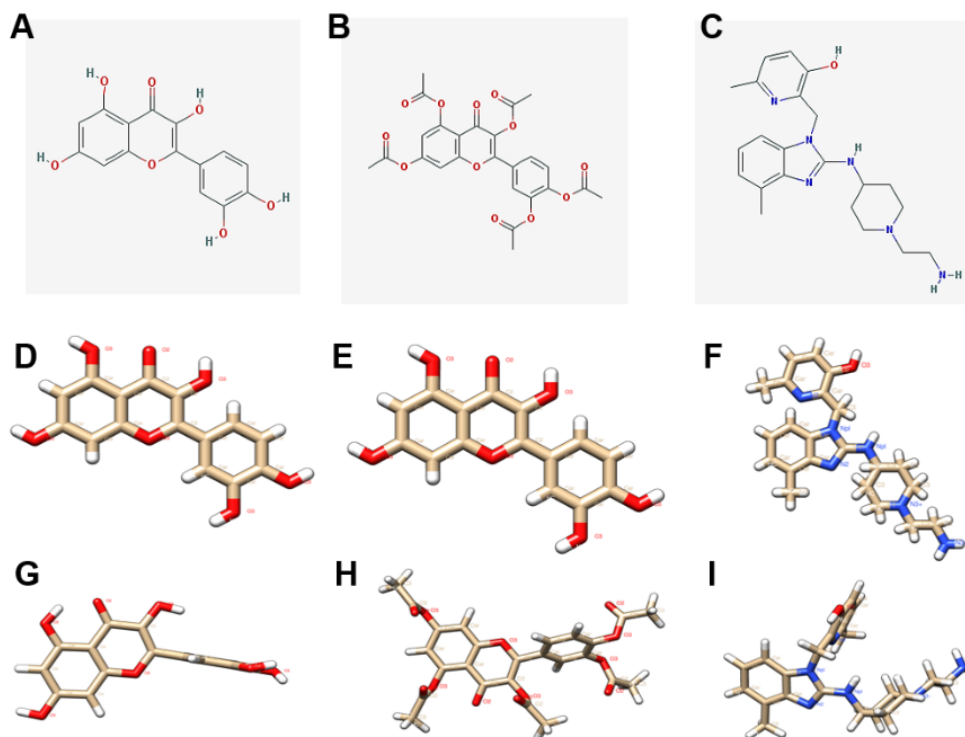
Fonte: o autor.

Respostas adicionais foram obtidas pelos cálculos do método de MM-GBSA. Ele nos permite estimar as contribuições em termos energéticos para a interação proteína-ligante, e correlacionarmos com o que já sabemos. Vimos que a quercetina apresenta uma tendência ao escape da cavidade, confirmada pelo cálculo de distância. Vimos também que a tendência de afastamento do Q1 e do inibidor, no entanto, são menores. Ademais, a quercetina forma uma ligação de hidrogênio e o inibidor três, embora para o Q1 esse tipo de interação com a cavidade não tenha sido constatada. Aqui, podemos observar que para o complexo RSVmF-quercetina, as interações eletrostáticas (EEL)

sobressaem (-31,9703 kcal/mol), embora existam contribuições menores de Van der Waals (VdW) (-18,1968 kcal/mol). Já o complexo RSVmF-Q1 revela uma tendência oposta, com maiores contribuições VdW (-35,9254 kcal/mol) e menores contribuições EEL (-21,0945 kcal/mol). O inibidor JNJ-2408068 apresenta um equilíbrio entre esses dois tipos de interação (-38,8372 kcal/mol VdW e -37,8727 kcal/mol EEL). Pela comparação entre os termos ΔG_{solv} e ΔG_{gas} é possível observar que a quercetina apresenta uma afinidade levemente maior pelo solvente em relação ao Q1. A avaliação do termo Δ_{total} nos revela a energia livre total de Gibbs obtida para formação dos complexos RSVmF-quercetina, RSVmF-Q1 e RSVmF-JNJ-2408068. A energia livre de Gibbs representa uma medida de quão favorável ou também espontâneo a formação de um complexo ou uma reação possam ser. Quanto menor o valor, mais espontânea e favorável a reação/formação do complexo é. Os resultados indicaram que RSVmF-quercetina apresenta um Δ_{total} -4,92 Kcal/mol. RSVmF-Q1 exibe um Δ_{total} de -14,22 kcal/mol e, o inibidor, um Δ_{total} -22,78 kcal/mol.

As informações obtidas pelo método de MM-GBSA corroboram a análise visual da [Figura 40](#), bem como os cálculos de distância e a investigação das ligações de hidrogênio. Apesar de o Q1 apresentar menores contribuições eletrostáticas, e não realizar ligações de hidrogênio com os resíduos formadores da cavidade, nota-se que, no modelo ora proposto, há uma melhor interação da cavidade de RSVmF com essa molécula, em relação à quercetina. Ao examinarmos a [Figura 42](#) podemos esboçar algumas inferências a respeito disso.

Figura 42 – Quercetina, Q1 e inibidor JNJ-2408068



A figura ilustra as estruturas da quercetina, da quercetina pentacetilada (Q1) e do inibidor JNJ2408068, os quais formam o complexo com RSVmF. Em (A), (B), (C), (D), (E) e (F) é possível visualizar a imagem em duas dimensões (2D) desses ligantes, com a diferença de que, nos três primeiros (A, B e C), a qualidade do tipo de ligação entre os átomos é discriminada, enquanto que nos outros três (D, E, e F), preza-se mais por uma representação espacial. Em (G), (H) e (I), pode-se ver o perfil em 3 dimensões (3D) dessas moléculas, evidenciando suas respectivas capacidades de torções e, portanto, seus graus de liberdade de movimento. Na primeira coluna (A, D e G) são representadas as estruturas referentes à molécula quercetina. Na segunda coluna (B, E, e H), aparecem as estruturas relativas ao derivado pentacetilado da quercetina, o Q1. E, por fim, na terceira coluna (C, F e I), exibem-se as estruturas alusivas ao inibidor usado como controle, o JNJ-2408068. Todos os arquivos foram extraídos da base de dados Pubchem, e a edição das imagens das segunda e terceira linhas foram realizadas no Chimera. Em azul aparecem os nitrogênios; em vermelho os oxigênios; em branco os hidrogênios; em bege claro os carbonos.

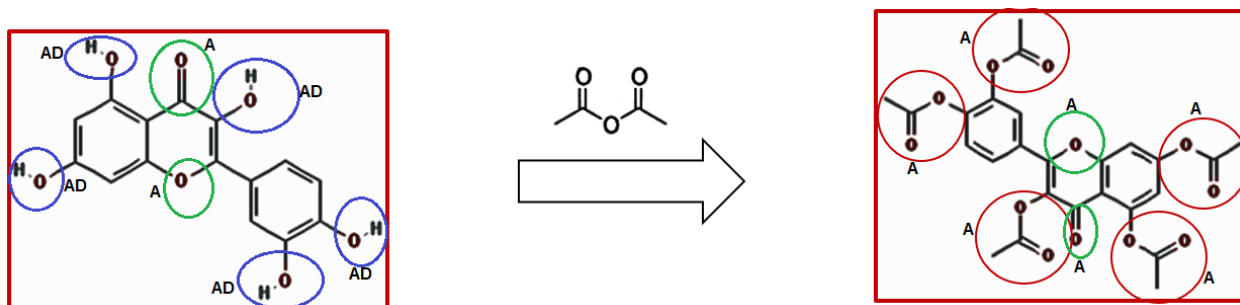
(A), (D) e (G) = quercetina, (B), (E) e (H) = Q1; (C), (F) e (I) = inibidor JNJ2408068. As propriedades físico-químicas dessas moléculas se encontram no apêndice [IX](#).

Fonte: o autor.

A quercetina é um flavonoide com vários grupamentos hidroxila em sua estrutura. A presença desses grupos confere à molécula não só uma alta capacidade antioxidante, como também um caráter mais hidrofílico, em virtude do momento dipolar do grupo OH. Por outro lado, a sua acetilação (Q1) e, por conseguinte, a troca de radicais, ocasiona uma diminuição desse momento dipolar. Isso torna a molécula mais hidrofóbica. Ademais, a saída de um grupo hidroxil para entrada de um grupo acetil tem consequências também para o anel benzênico. O OH é um forte ativador do anel, o que colabora para o aumento da densidade eletrônica dessa estrutura e, por conseguinte, para aumento de sua reatividade. O acetil diminui essa reatividade por induzir um deslocamento de elétrons. Logo, o núcleo aromático fica menos reativo e a estrutura mais estável, o que possibilita interações via *stacking* aromático com outras estruturas possuidoras de anéis benzênicos ricos em elétrons. Já o inibidor JNJ2408068 é um derivado de benzoimidazol e, apesar de ter nitrogênios, é mais hidrofóbico do que a quercetina. Para além da questão de hidrofobicidade, a quercetina contém apenas uma torção em sua estrutura, ao passo que o Q1 e o inibidor possuem 11 e 6 torções, respectivamente [Figura 42, (G), (H), (I)]. As torções atribuem maiores graus de liberdade às moléculas, logo, maiores possibilidades de movimentação, conformação, acomodação, e, conseqüentemente, interação.

A mudança da molécula quercetina para quercetina pentacetilada implica, portanto, em alteração no grau de hidrofobicidade, em mudança na reatividade do anel aromático, em aumento de torção e flexibilidade, e também em diferença no padrão de ligações de hidrogênio a serem realizadas. A troca dos grupos significa a mudança no perfil de “doação/acepção” de hidrogênios para apenas “acepção”. Este fenômeno pode ser visualizado na [Figura 43](#):

Figura 43 – Acetilação da quercetina



A figura ilustra o processo de acetilação da molécula quercetina, a qual se converte em quercetina pentacetilada (Q1).

A = aceitador de hidrogênios; AD = aceitador/doador de hidrogênios.

Fonte: o autor.

O modelo proposto neste trabalho considera uma cavidade formada por resíduos de aspartato, glutamato e fenilalanina. Aspartato e glutamato são aminoácidos com cadeias laterais negativas. A fenilalanina é um aminoácido apolar, com um núcleo aromático. Logo, é desejável que o ligante que vá interagir com essa cavidade combine devidas características hidrofílicas e/ou iônicas, e hidrofóbicas.

O inibidor JNJ-2408068 interage com esse bolsão via *stacking* aromático com as fenilalaninas e por interações eletrostáticas com aspartato e glutamato [256]. A presença de nitrogênios na molécula favorece interações eletrostáticas com os resíduos negativos, bem como a formação de possíveis ligações de hidrogênio. A hidroxila presente na molécula também propicia o estabelecimento de ligações de hidrogênio e de interações eletrostáticas. Os anéis colaboram por fornecerem hidrofobicidade e por favorecerem a interação de *stacking* aromático com as fenilalaninas. Logo, esse inibidor reuniria características hidrofílicas e hidrofóbicas desejáveis para interação com a cavidade. No entanto, em testes *in vivo*, a molécula apresentou um tempo de retenção tecidual bem elevado (principalmente devido à porção amino-etil-piperidina) para diversas espécies de animais testadas, e, consequentemente, não foi aprovada para as etapas posteriores de validação clínica [303].

A quercetina, com seus cinco grupamentos hidroxila, ligados ao anel benzênico, também deveria interagir com a cavidade de forma satisfatória, já que agrega características hidrofílicas (grupos OH) e hidrofóbicas (anéis). No entanto, isso acontece bem timidamente, o que se revela por seu Δ_{total} -4,92 Kcal/mol. Uma possível explicação é que o efeito de solvatação da molécula diminua sua possibilidade de interação via ligação de hidrogênio com os aminoácidos aspartato e glutamato. Devemos lembrar que o grupo OH apresenta capacidade de doação/acepção de elétrons, o que vai ser condicionada ao tipo de interação realizada. Outra possibilidade é que a ligação do grupo OH ao anel aumente sua reatividade, e diminua a interação via *stacking* aromático com a fenilalanina, a qual possui um anel aromático também rico em elétrons. Por outro lado, o Q1, por ser mais apolar, interage menos com o solvente, e exibe maior participação na interação com as fenilalaninas, possivelmente via *stacking* aromático, o que o mantém na cavidade. Os cinco grupamentos acetil da molécula, por sua vez, retratam sua capacidade acceptora de elétrons. Como a cavidade é formada por resíduos de aspartato, glutamato e fenilalanina, não há possibilidades de interação via ligação de hidrogênio com as cadeias laterais desses aminoácidos, o que reduz o panorama desse tipo de interação para esse ligante.

Um trabalho recentemente publicado por Raschka e colaboradores [304] aponta paradigmas de interação proteína-ligante, principalmente focado nos padrões de ligações de hidrogênio estabelecidas nessa interface. Os autores expõem que o reconhecimento proteína-ligante exige a combinação de certos perfis por parte da proteína e por parte do ligante. A partir da análise de 136 complexos de proteínas não-homólogas com três classes de ligantes (análogos de nucleotídeos, peptídeos e ligantes orgânicos), os autores observaram alguns dos seguintes perfis de interação: a proteína tende mais a se comportar como um doador de elétrons e os ligantes como aceptores de elétrons; as cadeias laterais dos aminoácidos arginina e lisina são as que mais participam como doadores, ao passo que as cadeias laterais de aspartato e glutamato atuam como receptores; o pareamento de um grupo amino doador da proteína – oxigênio acceptor do ligante é o tipo de reconhecimento proteína-pequeno ligante mais dominante, entre

outros. Os autores também alegam que selecionar um ligante como aceptor, em vez de um ligante com capacidade de aceitação/doação, estaria relacionado com uma determinação de seletividade para interação proteína-ligante, pois a doação de elétrons pelo grupo amino envolveria uma angulação e um volume mais estreitos do que aqueles referentes a um grupamento hidroxila. O estreitamento diminuiria a possibilidade de interações, e conseqüentemente, propiciaria uma maior especificidade à ligação.

Assim, a alteração do grupamento hidroxila para acetila, ou ainda, de um grupamento com capacidade de aceitação/doação para um com capacidade de aceitação apenas, no caso da quercetina e do Q1, respectivamente, resulta em maior afinidade pela cavidade, ainda que não por meio de ligações de hidrogênio. Nesse sentido, a modificação estrutural também pode ser vista sob a perspectiva de adequação às regras de Lipinski, ou a regra dos 5 [305,306]. Essas normas são chamadas de “a regra dos 5” pelo fato de os valores estipulados por elas se aproximarem de 5, e não porque incluam exatamente 5 diretrizes [305,306]. Elas estabelecem critérios para classificar compostos com baixa absorção e permeabilidade, isto é, compostos que apresentem: mais do que 5 doadores de hidrogênio, mais do que 10 aceptores de hidrogênio, um peso molecular maior do que 500 daltons (1 da ≈ 1 g/mol), ou valor de cálculo de log P maior do que 5. Verber e colaboradores [307], sob outra perspectiva, argumentam que as propriedades mais importantes para determinar a biodisponibilidade oral de candidatos a medicamentos se resumam em: 10 ou menos ligações rotacionáveis, e uma área de superfície polar topológica (TPSA – *Topological Polar Surface Area*) ≤ 140 Å. Dessa forma, ainda que o Q1 tenha revelado uma melhor capacidade de interação com RSVmF, e, por conseguinte, de inibição dessa proteína, outras modificações serão necessárias para se alcançar um composto de padrão “*drug-like*”. Isso estaria relacionado à diminuição de seu peso molecular, de sua TPSA, bem como dos graus de liberdade atribuídos pelas ligações rotacionáveis em sua estrutura. Ademais, a inserção de um grupo com características catiônicas na(s) posição (ões) 3' e 4' (anel B) do Q1, talvez implique em um ligeiro aumento da especificidade do ligante pela cavidade, uma vez que beneficiaria a interação com os resíduos de aspartato e glutamato.

Vários estudos vêm sendo publicados sobre o uso do *Docking* e da DM, aplicados a flavonoides, às proteínas do RSV e às proteínas de outros organismos também. Teixeira e colaboradores [204], em um estudo sobre o potencial de interação da quercetina com a proteína RSV-M2-1, obtiveram valores de energia de docking próximos a -7,3 kcal/mol. Em uma outra pesquisa, com derivados acetilados da quercetina e a proteína RSV-M2-1, Guimaraes e colegas [205] expõem valores energéticos de *docking* na ordem de -6,5 kcal/mol. Souza *et al.* [206] discutem sobre a influência da glicosilação nos flavonoides e a interação com RSV-G, mostrando valores de interação também na ordem de -6, -7 kcal/mol. Gomes e colaboradores [207], entretanto, analisando a interação da quercetina com RSV-N1, propõem valores da ordem de -29 kcal/mol. Todos esses trabalhos tiveram validação de técnicas biofísicas.

Colizzi e colaboradores [263] utilizam DM-orientada para comparar o modo de ligação de flavonoides à proteína β -hydroxyacyl-ACP desidratase de *Plasmodium falciparum* por associação com a quantidade de força necessária para removê-los do sítio de interesse. Assim, de acordo com a força aplicada, os compostos eram classificados em ativos ou inativos. O método teve confirmação experimental.

Em um outro estudo, com *E. coli* [198] o grupo de pesquisa combina análises de docking com a metodologia do 3D-QSAR (*Quantitative Structure-Activity Relationship*, relação quantitativa de estrutura-atividade) e sugerem posições na molécula bem como radicais substituintes que devem ser considerados na busca por agentes antibacterianos. O método também teve confirmação experimental.

A simulação da interação da quercetina com a proteína F, sugere o porquê de seu tímido efeito virucida, em inóculos superiores a M.O.I. 0.1. Embora ela possa interagir com essa proteína, que é a responsável por realizar a fusão do vírus à célula e assim, de fato, iniciar a infecção, essa interação é, em termos termodinâmicos, fraca. Além de ter revelado também que a molécula apresenta uma tendência ao escape daquela cavidade, o que não inibiria efetivamente sua alteração conformacional. A acetilação da molécula, por outro lado, aumenta a sua capacidade virucida, por favorecer energeticamente a

interação com a proteína e mantê-la na cavidade, possivelmente pela sua característica mais hidrofóbica.

Finalmente, dado que uma alteração na estrutura da quercetina amplificou sua capacidade anti-RSV, nós buscamos desenvolver uma ANN para prever moléculas flavonoídicas possuidoras de atividade inibitória contra RSV. Para isso, nós elaboramos uma sequência de base de dados, a fim de implementar uma ANN, acessada pelo algoritmo genético, na plataforma Matlab®.

4.3 ANN: elaboração e implementação

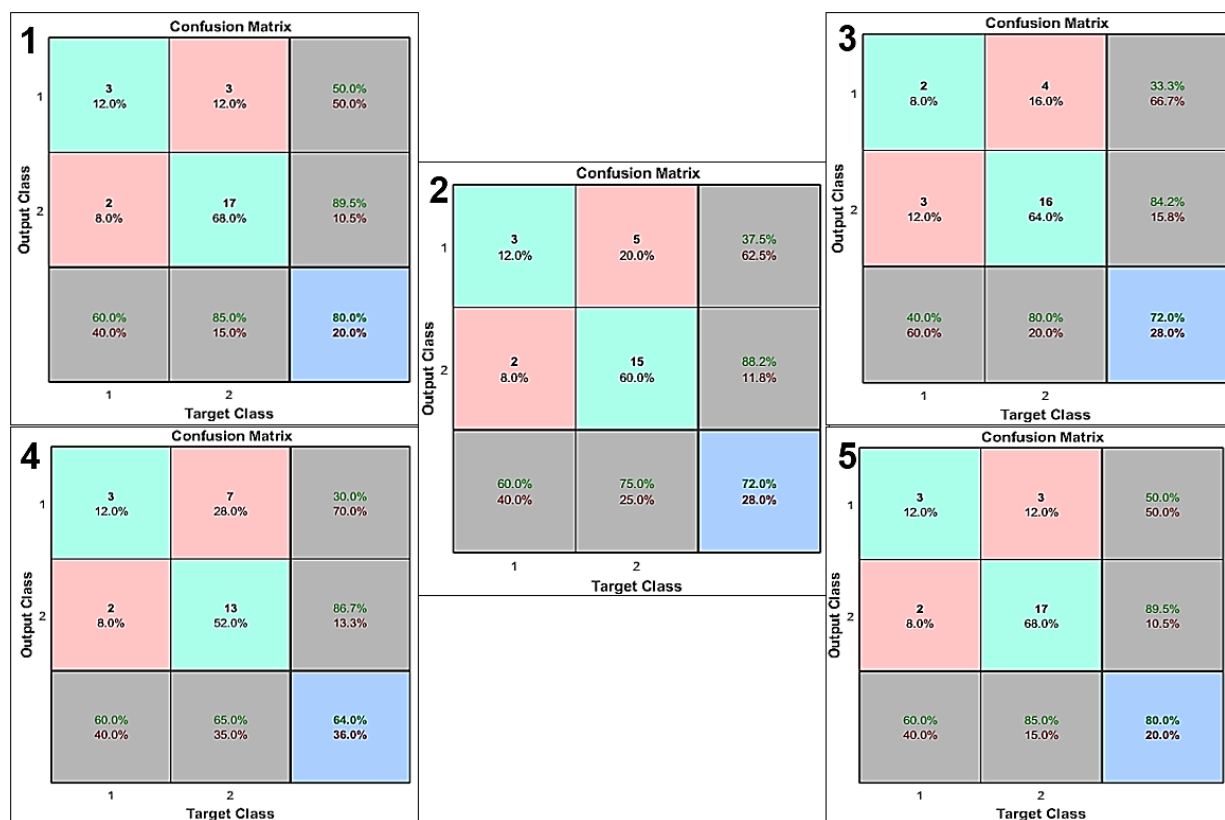
Visamos elaborar uma ANN com propósito de prever flavonoides com capacidade anti-RSV, para nos auxiliar na discriminação de moléculas pertencentes a esse grupo, e possuidoras da referida atividade inibitória, anteriormente à experimentação na bancada.

Para isso primeiramente construímos dois bancos de dados: BD1 para consulta e para posterior aperfeiçoamento da rede, e BD2 para elaboração e implementação dessa primeira versão de rede neural. Além disso, estabelecemos certos parâmetros de entrada, os quais devem ser preenchidos, a fim de que a rede elabore uma decisão sobre a atividade do flavonoide em questão. Esses parâmetros incluem: a classe do flavonoide, a presença ou ausência de glicosilações bem como as (eventuais) posições dessas, o número de átomos de carbono-hidrogênio-oxigênio, o peso molecular, o log P, a TPSA, a concentração a ser testada, a cepa viral do RSV, a quantidade de vírus, e o tipo de tratamento a ser aplicado. A partir desses parâmetros de entrada obrigatórios, a ANN elaborava uma classificação para o flavonoide em “ativo” ou “inativo”.

Nossos primeiros resultados com relação à capacidade preditiva da ANN a partir de BD2, consistiram da divisão da amostra (n=240 dados, referentes a 70 flavonoides) em dois grupos: um grupo (n= 215), para construção da rede, e um grupo de simulação (n= 25), para avaliação de sua capacidade preditiva. Ou seja, as redes foram elaboradas a partir da amostra de 215 flavonoides. Após esse procedimento, o grupo de simulação,

isto é, moléculas desconhecidas pela rede, foram apresentadas a ela e a taxa de acerto foi mensurada, dado que nós já tínhamos as respostas referentes às capacidades “*ativa*” ou “*inativa*” dessas 25 moléculas. Logo, a simulação nos permitiu verificar qual o percentual de acerto da ANN, e, por conseguinte, avaliar sua competência preditiva. Foram selecionados os cinco melhores tipos de redes, e as respostas das simulações são apresentadas abaixo, na [Figura 44](#):

Figura 44 – Primeira simulação ANN



As figuras representam as matrizes de confusão, resultantes da simulação de 25 flavonoides (n=5 ativos, e n=20 inativos), com as cinco melhores ANN obtidas, nesse primeiro passo.

Nas matrizes de confusão exibidas acima, as primeiras duas células da diagonal mostram o número e a percentagem de classificações corretas pela rede treinada. Por exemplo, em (1), três flavonoides ativos foram classificados corretamente como “ativos”. Isso corresponde a 12% dos 25 flavonoides utilizados para simulação. Da mesma forma, 17 flavonoides inativos foram corretamente classificados como “inativos”, o que representa 68% de 25 flavonoides utilizados na simulação. Três flavonoides inativos foram classificados erroneamente como “ativos”, o que corresponde a 12% de 25 flavonoides usados na simulação. De modo similar, 2 flavonoides ativos foram classificados erroneamente como “inativos”, e isso representa oito por cento de 25 flavonoides utilizados na simulação. De seis predições para classificação de “ativos”, 50% foram corretas e 50% incorretas. De cinco flavonoides ativos, 60% foram preditos corretamente como “ativos” e 40% preditos como “inativos”. De 19 predições para classificação de “inativos”, 89,5% foram corretas e 10,5% incorretas. De 20 flavonoides inativos, 85% foram preditos corretamente como “inativos” e 15% preditos como “ativos”. No geral, 80% das predições foram corretas e 20% erradas.

1= matriz de confusão resultante da simulação 1;

2= matriz de confusão resultante da simulação 2;

3= matriz de confusão resultante da simulação 3;

4= matriz de confusão resultante da simulação 4;

5= matriz de confusão resultante da simulação 5.

Fonte: o autor.

OBS: A figura pode ser vista com maiores detalhes no apêndice [X](#).

Nossos primeiros resultados já revelaram uma relevante capacidade preditiva dos algoritmos da rede. Pela análise da [Figura 44](#), podemos observar que as melhores redes ora obtidas, referem-se àquelas representadas em 1 e 5. Elas apresentaram uma acurácia³ geral de acerto de 80%, para nossa primeira simulação. Em relação aos flavonoides classificados como “*ativos*”, o acerto foi estimado em 60% enquanto que para aqueles identificados como “*inativos*”, a estimativa foi de 85%. A acurácia geral das outras três redes variou entre 64 e 72%, e de 40 a 60% de acerto para os flavonoides “*ativos*” e 65 a 80% para os “*inativos*”.

Embora os resultados tenham sido promissores para um início, visto que uma acurácia igual a 80% é significativa, é possível inferir que há divergências quanto ao critério de separação entre “*ativos*” e “*inativos*” para os algoritmos da rede. Em busca de uma melhor compreensão de onde possivelmente estariam as falhas relacionadas à categorização desses flavonoides, nós extraímos os resultados das simulações para podermos avaliar qual tipo de estrutura ou substituinte poderia(m) estar dificultando o acerto na classificação pela rede. Esses dados foram convertidos em binários (0 e 1), e são apresentados na [Tabela 3](#) a seguir.

³ As terminologias “*acurácia*” e “*precisão*”, embora não sejam estritamente sinônimos em um contexto puramente estatístico, são assim utilizados neste trabalho para facilitar a fluência textual.

Tabela 3 – Resultados binários da simulação 1

Alvos		Saídas 1		Saídas 2		Saídas 3		Saídas 4		Saídas 5		Classe molécula
1	0	1	0	1	0	1	0	1	0	1	0	flavone aglycone
1	0	0	1	0	1	0	1	0	1	0	1	flavone glycoside - O-heteroside (C7)
1	0	1	0	0	1	0	1	1	0	0	1	flavan-3-ol aglycone
1	0	1	0	1	0	0	1	0	1	1	0	flavonol aglycone
1	0	0	1	1	0	1	0	1	0	1	0	flavone aglycone
0	1	0	1	1	0	0	1	0	1	0	1	flavone aglycone
0	1	1	0	1	0	1	0	1	0	1	0	flavone aglycone
0	1	0	1	0	1	0	1	0	1	0	1	flavone glycoside - O-heteroside (C7)
0	1	0	1	0	1	0	1	0	1	0	1	flavan-3-ol aglycone
0	1	0	1	0	1	0	1	0	1	0	1	flavan-3-ol aglycone
0	1	0	1	0	1	0	1	0	1	0	1	flavanone aglycone
0	1	1	0	0	1	0	1	1	0	0	1	flavone glycoside O-heteroside (C3)
0	1	0	1	0	1	0	1	0	1	0	1	flavanone glycoside O-heteroside (C7)
0	1	0	1	1	0	0	1	1	0	0	1	isoflavone glycoside C-heteroside (C8)
0	1	0	1	0	1	0	1	0	1	0	1	flavonol aglycone
0	1	0	1	0	1	0	1	0	1	0	1	flavonol aglycone
0	1	0	1	0	1	0	1	0	1	0	1	flavonol aglycone
0	1	0	1	1	0	1	0	1	0	1	0	flavone aglycone (acetyl)
0	1	0	1	0	1	1	0	1	0	0	1	flavone glycoside O-heteroside (C7)
0	1	0	1	0	1	0	1	0	1	0	1	flavanone aglycone
0	1	0	1	0	1	0	1	0	1	0	1	flavanone aglycone
0	1	1	0	1	0	0	1	0	1	0	1	flavone aglycone (methyl)
0	1	0	1	0	1	0	1	1	0	0	1	flavone glycoside C-heteroside (C6)
0	1	0	1	0	1	1	0	1	0	1	0	flavone aglycone (methyl)
0	1	0	1	0	1	0	1	0	1	0	1	flavone glycoside O-heteroside (C7)
Ativo	Inativo	Ativo	Inativo	Ativo	Inativo	Ativo	Inativo	Ativo	Inativo	Ativo	Inativo	

A tabela mostra os resultados concernentes à primeira simulação (n=25 flavonoides) com as cinco melhores redes obtidas em um primeiro passo de avaliação. Os “alvos” se referem às respostas que devem ser obtidas, em termos de “ativos” e “inativos”, isto é, equivalem ao “ideal”. As “saídas” de cada rede foram numeradas de um a cinco e mostram os resultados obtidos na simulação para cada uma dessas redes, logo, o “real”. Em verde, aparecem as classes dos flavonoides “ativos” utilizados para simulação (n=5). Em rosa, exibem-se as classes dos flavonoides “inativos” utilizados para a simulação (n=20). Em vermelho, são destacados os resultados mais proeminentemente incorretos. Em amarelo, são enfatizadas as melhores redes, ou seja, aquelas que obtiveram as maiores percentagens de acerto na simulação. Em negrito, são expostos flavonoides cujas classes e tipos/subtipos apresentaram maiores frequências de erros na predição na simulação das cinco ANNs.

Fonte: o autor.

Pela avaliação da [Tabela 3](#), notam-se as principais fontes de erro em relação à classificação das redes para os critérios de flavonoides “*ativos*” e “*inativos*”. Podemos visualizar que todas elas erraram na classificação de um flavonoide glicosilado “*ativo*”. Ademais, todas também erraram em classificar como “*inativo*” uma flavona aglicona. E, de uma forma geral, constatamos que os equívocos se relacionam às moléculas com grupamentos substituintes (glicosídeos e radicais acetil/metil). Além disso, houve uma frequência maior de erros na classificação dos “*ativos*”. Todavia, é necessário ponderar os pontos positivos também. Os algoritmos da rede se revelaram capazes de distinguir como “*ativo*” uma flavona aglicona, para todas as cinco redes exibidas. Em relação aos “*inativos*”, oito daqueles cujo subtipo é aglicona e sem radicais, foram categorizados corretamente. E, além desses, duas flavonas glicosiladas e uma flavanona glicosilada foram identificadas como “*inativas*”, sem erros.

Um dos fatores mais críticos para que uma predição seja acurada diz respeito à representatividade do espaço amostral utilizado. No caso de nossa ANN, isso se relaciona não só ao conjunto de dados empregado para elaboração da rede, como também àquele utilizado na simulação. É, prudente, portanto, observar a representação do grupo utilizado para construção da rede no conjunto de simulação. Por melhores que sejam os algoritmos, a apresentação de uma molécula totalmente diversa daquelas com as quais a rede foi treinada configura uma provável fonte de erro, dado que não houve esse aprendizado.

Assim, nós retomamos o nosso BD2 para avaliarmos a representatividade das diferentes classes e subtipos nesse conjunto de dados, a qual é ilustrada de forma hierarquizada nas [Tabela 4](#) abaixo

Tabela 4 – Hierarquia Ativos e Inativos

ATIVO			INATIVO		
Flavona	aglicona	27	Flavona	aglicona	39
	glicosídeo	8		glicosídeo	19
	TOTAL	35		TOTAL	58
Flavonol	aglicona	8	Flavonol	aglicona	58
	glicosídeo	0		glicosídeo	0
	TOTAL	8		TOTAL	58
Flavan-3-ol	aglicona	7	Flavanona	aglicona	17
	glicosídeo	0		glicosídeo	18
	TOTAL	7		TOTAL	35
Biflavonoide	aglicona	6	Flavan-3-ol	aglicona	13
	glicosídeo	0		glicosídeo	0
	TOTAL	6		TOTAL	13
Flavanon-3-ol	aglicona	3	Isoflavona	aglicona	8
	glicosídeo	0		glicosídeo	2
	TOTAL	3		TOTAL	10
Flavan	aglicona	2	Isoflavan	aglicona	1
	glicosídeo	0		glicosídeo	0
	TOTAL	2		TOTAL	1
Flavanona	aglicona	2	Biflavonoide	aglicona	1
	glicosídeo	0		glicosídeo	0
	TOTAL	2		TOTAL	1
TOTAL ATIVO		63	Aurona	aglicona	1
				glicosídeo	0
				TOTAL	1
			TOTAL INATIVO		17

A Tabela 4 (A) revela a hierarquia das classes e tipos (aglicona/glicosídeo) relativos ao grupo dos flavonoides classificados como “*ativos*”. Já a Tabela 4 (B), exibe a hierarquia das classes e tipos dos flavonoides classificados como “*inativos*”. Em rosa claro, estão enfatizadas as classes desses compostos que só aparecem no quesito “*inativos*”.

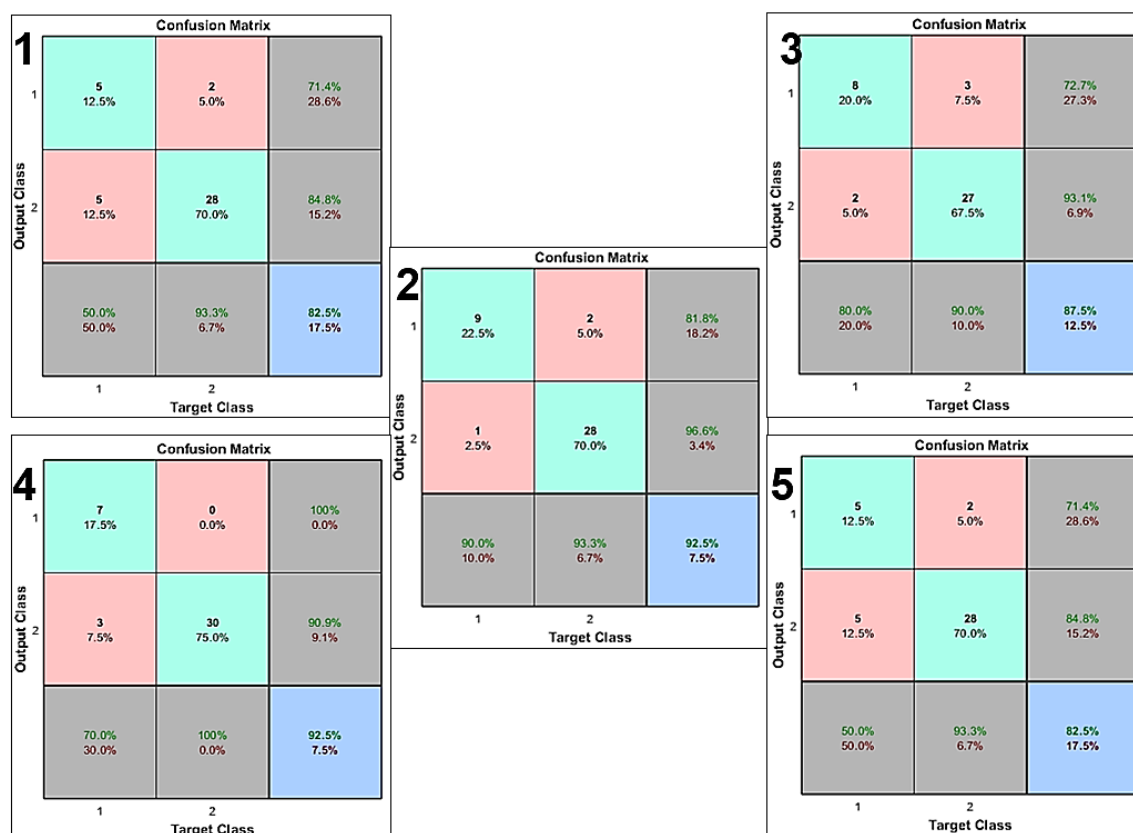
Fonte: o autor.

A análise da [Tabela 4](#) nos revela a proporção numérica de cada classe de flavonoides em nosso conjunto amostral. Tanto para os “*ativos*” quanto para os “*inativos*” as flavonas representam a classe numericamente maior, seguida pelos flavonóis, e, em ambos, a proporção de agliconas é superior a de glicosídeos. Excluídas essas classes, existem divergências quanto à hierarquia das outras. Convém notar que nos “*ativos*” os glicosídeos aparecem apenas na classe flavonas, enquanto que para os “*inativos*”, esses estão presentes nas classes flavonas, flavanonas e isoflavonas. Ademais, as classes flavanon-3-ol e flavan têm representantes apenas nos “*ativos*”, enquanto que o oposto é verdadeiro para as classes isoflavan, isoflavona e aurona em relação aos “*inativos*”. Além disso, é evidente a diferença numérica entre as duas categorias de flavonoides, em que a proporção entre elas é aproximadamente de 1 flavonoide ativo para cada 3 inativos.

À vista de todas essas peculiaridades relacionadas a BD2, nós buscamos uma estratégia que reproduzisse melhor a realidade do banco de dados na simulação, bem como nos permitisse avaliar a interferência de diferentes conjuntos de dados no ajuste dos parâmetros da rede. Essa estratégia é chamada validação cruzada [\[308,309\]](#) em κ -partições e consiste em dividir o conjunto amostral em subconjuntos, de forma a analisar a aptidão de cada subconjunto em definir as características da rede, e, por conseguinte, em determinar sua taxa de acerto. Assim, fragmentamos todo o nosso conjunto amostral de dados de flavonoides em cinco grupos diferentes ($n=200$), e, também, construímos cinco grupos diversos entre si para a simulação ($n=40$). Ou seja, o agrupamento1 reuniu 200 flavonoides para construção da rede, e 40 flavonoides para simulação, e assim semelhantemente para os outros quatro grupos. Essa composição variou entre os cinco conjuntos, tanto para elaboração de cada rede ($n=200$), quanto para cada simulação ($n=40$). Ademais, os subconjuntos de simulação respeitaram os critérios de hierarquia exibidos na [Tabela 4](#). Além disso, preocupamo-nos em aumentar nossa amostra de simulação, dado que apenas cinco flavonoides ativos não respaldam uma estatística significativa. Por conseguinte, os grupos de simulação contêm 10 flavonoides ativos e 30 flavonoides inativos ($n=40$, no total). O resultado de cada agrupamento para a simulação é exibido abaixo, na [Figura 45](#).

Com base nos cinco subconjuntos, houve treinamento e elaboração de cinco redes diversas entre si, as quais foram utilizadas posteriormente para realizar as simulações. Os resultados obtidos a partir dessas simulações são ilustrados abaixo, na [Figura 46](#):

Figura 46 – Matrizes de confusão da simulação 2



As figuras representam as matrizes de confusão, resultantes de simulação de 40 flavonoides (n=10 ativos, e n=30 inativos). Esses resultados se referem, cada um, a um tipo diferente de arquitetura de ANN. Além disso, cada um também provém de um subconjunto distinto de simulação [\[Figura 45\]](#), seguindo o princípio abaixo:

1= matriz de confusão resultante da ANN elaborada a partir do grupo 1 (n=200) e da simulação do grupo 1 (n=40);

2= matriz de confusão resultante da ANN elaborada a partir do grupo 2 (n=200) e da simulação do grupo 2 (n=40);

3= matriz de confusão resultante da ANN elaborada a partir do grupo 3 (n=200) e da simulação do grupo 3 (n=40);

4= matriz de confusão resultante da ANN elaborada a partir do grupo 4 (n=200) e da simulação do grupo 4 (n=40);

5= matriz de confusão resultante da ANN elaborada a partir do grupo 5 (n=200) e da simulação do grupo 5 (n=40);

Fonte: o autor.

OBS: A figura pode ser vista com maiores detalhes no apêndice [XI](#).

A partir da simulação 2, podemos notar que três redes apresentaram predições significativas, principalmente no que tange aos flavonoides “*ativos*”. Se na primeira simulação obtivemos um percentual máximo de acerto em torno de 60%, na segunda alcançamos uma taxa de 70 a 80%. Houve avanços também quanto à predição dos “*inativos*”, a qual, inclusive, atingiu patamares de acerto de 100%. Assim, é possível inferir que a estratégia de validação cruzada em κ -partições foi eficaz em auxiliar a aumentar a taxa de sucesso, entendida aqui como acerto. Dentre as cinco redes geradas e simuladas, as três melhores se referem àquelas representadas em “2”, “3”, e “4”, cujas características e parâmetros são expostos na [Tabela 5](#) abaixo:

Tabela 5 – Parâmetros das melhores arquiteturas ANN

Parâmetros		Arquitetura ANN		
		2	3	4
Neurônios	1ª camada	183	66	40
	2ª camada	123	110	142
	3ª camada	82	23	78
Função de transferência	entre entrada e 1ª camada	radbas	tribas	radbas
	saída	tansig	tribas	purelin
Função de treinamento		trainscg	traincgf	trainscg
Taxa de acerto (treinamento)		87,14%	86,47%	87,86%
Taxa de acerto (teste)		100%	100%	100%
Taxa de acerto (validação)		76,67%	80%	76,67%
Taxa de acerto simulação		92,50%	87,50%	92,50%
Erro quadrático médio		0,0976	0,0978	0,0856

As taxas de acerto se referem às taxas gerais de acerto (por exemplo, célula em azul nas Figuras [44](#) e [46](#)), enquanto que os erros quadráticos médios se relacionam ao erro quadrático médio obtido para o treinamento de cada ANN.

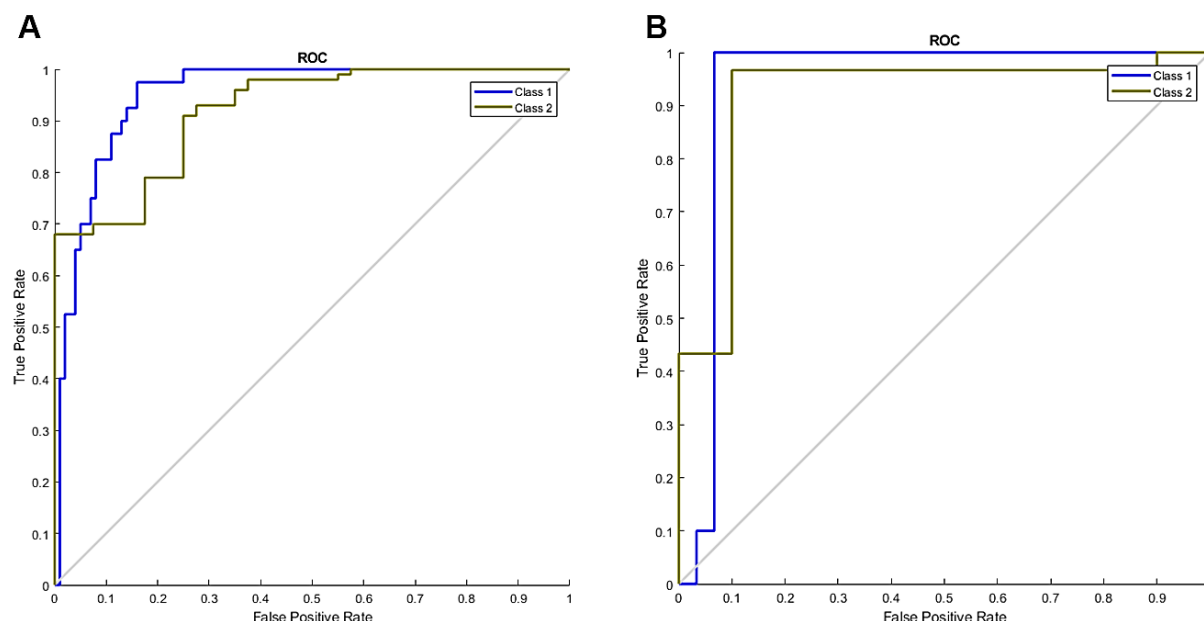
Fonte: o autor.

A avaliação da [Tabela 5](#), bem como da [Figura 46](#) nos informa que a rede com maior acurácia para as duas categorias “*ativos*” e “*inativos*” é a rede 2. A despeito de a rede 4 ter revelado uma acurácia de 100% na simulação dos flavonoides “*inativos*”, ela

errou três flavonoides “*ativos*” enquanto que a rede 2 o fez apenas para um. Ademais, a rede 2 classificou incorretamente dois flavonoides “*inativos*”, mas, apesar disso, a proporção de acurácia para os dois grupos se manteve acima de 90%.

Logo, a partir da seleção de “2” como melhor modelo de ANN, deve-se analisar a acurácia da rede em classificar os dados dentro de um critério binário, isto é, “*ativo*” e “*inativo*”. É importante que a classificação possua sensibilidade e especificidade, ou seja, que categorize “*ativo*” como “*ativo*” e “*inativo*” como “*inativo*”, em relação à totalidade de flavonoides ativos e inativos presentes na amostra avaliada. Isso nos fornece uma probabilidade de acerto a qual pode ser correlata com a validade do teste. Dessa forma, espera-se que haja uma maximização da sensibilidade em detrimento de uma minimização de falsos positivos. Os conceitos de sensibilidade e especificidade são convenientemente avaliados pela curva ROC (*Receiver Operating Characteristic* – “eficiência de operador de recepção de sinais”) [revisto por [330](#)] como ilustrado na [Figura 47](#).

Figura 47 – Curva ROC ANN



A curva ROC mensura a proporção de acertos verdadeiros (sensibilidade) no eixo y, contra a proporção de acertos falsos, no eixo x ($= 1 - \text{sensibilidade}$). Em (A), identifica-se a curva ROC referente ao treinamento da arquitetura de ANN do grupo 2, e, em (B), mostra-se a curva ROC relativa à simulação do grupo 2 de simulação [Figura 46] utilizando essa mesma rede.

Classe 1= “*ativos*”

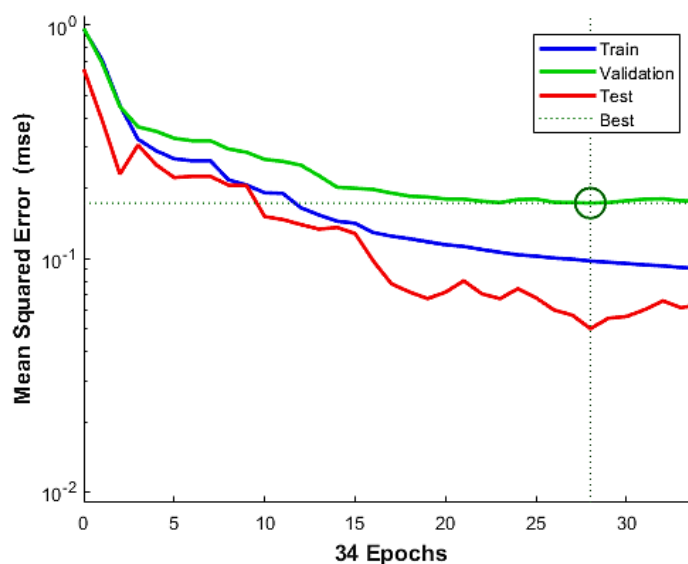
Classe 2= “*inativos*”

Fonte: o autor.

As curvas ROC nos confirmaram a validade da predição obtida pela ANN do grupo 2 (ANN-g2), isto é, de que há acurácia e de que essa é pertinente em relação às duas categorias de flavonoides aqui abordadas: “*ativos*” e “*inativos*”. A precisão nos é revelada a partir da proximidade das curvas com o eixo y (sensibilidade), e também nos evidencia que, inicialmente, existe uma maior imprecisão entre os “*ativos*”, e que essa tendência se altera com o passar da apresentação das amostras no treinamento e na simulação.

Outra maneira de se avaliar a rede é checar o seu desempenho dinâmico (*dynamic performance*), como ilustrado na [Figura 48](#) abaixo.

Figura 48 – Desempenho dinâmico ANN-g2



O gráfico revela a iteração na qual o desempenho de validação atingiu um mínimo, logo, a época 28. O treinamento continuou por mais 6 iterações, mas o melhor desempenho é correlato com a época que apresentou menor erro quadrático na validação.

Em y temos o “erro quadrático médio” (“*Mean Squared Error*”) e em x o “número de época” (“*Epochs*”).

Erro quadrático médio treinamento = 0,0976

Erro quadrático médio teste = 0,0501

Erro quadrático médio validação = 0,1727

Train = Treinamento;

Validation = Validação;

Test = Teste.

Fonte: o autor.

A [Figura 48](#) expõe que as curvas de validação e de treinamento são bem similares, enquanto que a de teste difere dessas. Ademais, exhibe que o erro quadrático médio (mse) mínimo de validação foi maior do que o mse mínimo de treinamento. Além disso, o valor de mse de teste não é superior ao do treinamento.

Os dados de validação refletem a capacidade de generalização da rede frente a amostras até então não-vistas, isto é, valida o processo de treinamento. De forma similar, os dados de teste avaliam o desempenho final de um modelo treinado. Enquanto a

validação ocorre ao mesmo tempo do treinamento, o teste ocorre após o término do treinamento. Isso implica que o desempenho da rede seja avaliado mesmo enquanto o treinamento ocorre, de forma a evitar que o processo de treinamento se torne viciado. Assim, se o erro de validação fosse menor do que o erro de treinamento, isso sugeriria que a rede não possuísse capacidade de generalização. Ademais, o fato de o mse do teste ter sido menor do que o mse de treinamento, indica que não houve memorização excessiva (*overfitting*), ou seja, a rede não decorou as respostas, mas se mostrou capaz de desenvolver solução para classificar as amostras.

Os resultados de nossa ANN descritos acima corroboram não só o potencial preditivo da técnica, como também do método aqui proposto, isto é, redes neurais associadas ao algoritmo genético. O objetivo primário de nosso trabalho era a classificação de flavonoides em “*ativos*” ou “*inativos*”, formulada a partir de parâmetros de entrada. Em um primeiro passo de análise já obtivemos um acerto geral em torno de 80%, e, a partir de uma investigação mais acurada de nosso banco de dados, bem como da busca por um melhor espelhamento desse banco de dados na simulação, conseguimos alcançar patamares de acerto de 92,5%, e propor a utilização da rede do grupo 2 (ANN-g2) como um modelo preditivo. Além disso, os dados de desempenho revelaram que não houve *overfitting*, e a curva ROC nos permite afirmar que a ANN por nós proposta apresenta sensibilidade, e, portanto, é capaz de elaborar a categorização almejada.

Por outro lado, muito embora a tática por nós utilizada de extrair a maior quantidade possível de dados quantitativos dos artigos nos tenha possibilitado chegar a 240 dados referentes a 70 flavonoides, é importante que haja um acréscimo quantitativo de moléculas a esse banco de dados. Quanto maior a amostra, maior a capacidade de generalização da ANN por haver maior representatividade. O que chama a atenção, todavia, é o fato de que estudos como os nossos dependam substancialmente do volume disponível de dados experimentais, o que pode ser um fator limitante. Apesar disso, nós também construímos um BD1, com 690 compostos flavonoídicos, dos quais 70 foram utilizados nessa primeira versão de rede neural para avaliação de flavonoides com

atividade anti-RSV. Explorar esse banco de dados, de forma a avaliar a influência dos diferentes parâmetros de entrada, em um possível modelo de programação dinâmica, constitui uma real possibilidade de incrementos à rede e melhoramento dessa.

Outros trabalhos com redes neurais também ressaltam a capacidade dessa metodologia em auxiliar em diferentes critérios de predição e classificação. Noorizadeh & Farmany [310] utilizam dois modelos de ANN para investigar a correlação entre a capacidade antioxidante e o tipo de estrutura de ácidos benzoicos. Os autores reportam que os coeficientes de correlação para os dois modelos foram de 0,824 e 0,966. Em um outro estudo, por Cerit e colaboradores [311], há também a estimativa de atividade antioxidante em relação à quantidade de compostos fenólicos, pelo uso de redes neurais. Reporta-se um acerto de 90% na predição. Shen e colegas [312] utilizam o algoritmo genético combinado a rede neural baseada em *backpropagation* para relacionar a afinidade de ligação de diferentes tipos de flavonoides a uma P-glicoproteína (uma proteína do tipo transportador ABC multirresistente a drogas, em células tumorais). Os autores relatam a superioridade do método de ANN associada ao algoritmo genético ($r^2 \approx 0,9199$) em relação àquele baseado em regressão linear múltipla ($r^2 \approx 0,855$), e sugerem que seu modelo deva ser usado em triagens futuras. Karawajczyk e colegas [313] sugerem o uso de redes neurais na elaboração de QSAR para investigar o modo de ligação de flavonoides à bilitranslocase (um transportador de membrana). O grupo conclui que o modo de ligação é por ligações de hidrogênio e não por interações eletrostáticas. Por fim, Zuvela e colegas [314] desenvolvem seis métodos baseados em ANN para predição de atividade antioxidante de flavonoides, a partir da utilização de seis descritores moleculares. Os autores sugerem inúmeras correlações não-lineares as quais não poderiam ser visualizadas por regressão, e ousam propor que três das metodologias desenvolvidas no estudo podem ser usadas até mesmo por pessoas não-especialistas, devido à sua facilidade de interpretação.

O uso de redes neurais e de variados tipos de inteligência artificial em pesquisas contemporâneas, especialmente na grande área médico-farmacêutica, tem sido reconhecido e tem despontado [315-323, 198]. Ainda que haja pequenas discórdias ou

perspectivas múltiplas entre os acadêmicos e a sociedade, não só sobre a aplicabilidade delas, mas também quanto à ética inerente às suas diversas finalidades e ainda às próprias metodologias envolvidas, o pensamento comum é de que essas tecnologias auxiliam não só a reduzir a perda de tempo, como também de recursos. Há ainda a vantagem de que elas tragam consigo uma real melhora na eficiência em conduzir pesquisas experimentais, dado sua capacidade de identificação, predição, classificação, e estimação de amostras e eventos.

A primeira aplicação de redes neurais artificiais relacionadas a critérios de estrutura-atividade remonta aos anos 70. Hiller e colaboradores [324] desenvolveram um estudo sobre o uso de perceptrons para classificar em ativos ou inativos 1,3 dioxanos com substituintes. O perceptron foi treinado e testado a partir de um conjunto de 46 compostos com atividades conhecidas, e mostrou-se apto ao reconhecimento tanto no grupo de treinamento ($\approx 87\%$ acerto) quanto no de teste ($\approx 72\%$ acerto). Apesar da sua relevância, pioneirismo e autenticidade, o trabalho não foi muito apreciado na época e permaneceu no anonimato.

Nossa abordagem vai de encontro ao estudo de Hiller, o qual preconizou o desenvolvimento de redes neurais artificiais preditivas, e sana a lacuna de desatenção acadêmica para esse trabalho. Também relatamos o sucesso da ANN associada ao algoritmo genético em determinar, com precisão, a atividade ou inatividade de flavonoides contra RSV. Os ganhos referentes ao nosso estudo respaldam um planejamento racional de procedimentos experimentais, pela exclusão preliminar de flavonoides inativos e a pré-seleção de ativos. Isso poupa tempo e recursos, além de conferir funcionalidade e uniformidade aos procedimentos experimentais, promovendo um aumento de sua eficiência e favorecendo maior aproveitamento das pesquisas. Ademais, provamos o potencial, a relevância e a aplicabilidade de redes neurais artificiais em pesquisa de desenvolvimento de drogas relacionada à virologia.

5. CONCLUSÃO

“... all things are made of atoms, and [that] everything that living things do can be understood in terms of the jiggings and wiggings of atoms”. (Richard Feynman)

Nossos objetivos para este estudo foram elaborar a modelagem da proteína F em seu estado de pré-fusão e então realizar simulações de *docking* e dinâmica molecular dessa proteína em complexo com a quercetina e o Q1, e avaliar a interferência da acetilação nesse processo. Além disso, desenvolver uma ANN associada ao algoritmo genético de forma a prever flavonoides com capacidade anti-RSV.

Nossos resultados podem ser sumarizados da seguinte forma:

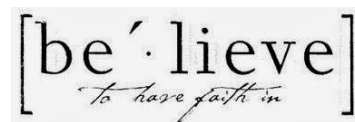
- a) O que deu certo:
 - a proteína F foi modelada completamente (RSVmF), incluindo as porções N-e C-terminais, o peptídeo 27, o domínio transmembrânico e a cauda citoplasmática;
 - os resultados de *docking* evidenciaram que a quercetina, o Q1 e o inibidor-controle ocupam a cavidade central da proteína e possuem basicamente a mesma energia de interação $\sim -6,5$ kcal/mol;
 - a DM com os ligantes evidenciou que a quercetina possui uma tendência ao escape da região da cavidade central, enquanto que o Q1 permanece na cavidade;
 - a acetilação aumentou a capacidade virucida da quercetina por favorecer a interação com a proteína. Isso é traduzido na energia livre de Gibbs, a qual, para o complexo RSVmF-Quercetina foi de $-4,92$ Kcal/mol enquanto que para o Q1 foi de $-14,22$ kcal/mol;
 - o modelo de ANN proposto no trabalho revelou excelente capacidade preditiva ($\approx 92,5\%$ de acerto geral), o que evidencia ser um método válido para o planejamento racional de procedimentos experimentais, pela exclusão preliminar de flavonoides inativos e a pré-seleção de ativos.

b) o que não deu (tão) certo:

- embora RSVmF tenha sido válida para explicar a interação com as moléculas, 200ns não foram suficientes para a estabilizar, principalmente quanto ao seu domínio transmembrânico e cauda citoplasmática. Isso sugere que ou essas regiões devem ser retiradas do modelo, ou inseridas em uma membrana;

- o banco de dados utilizado na elaboração de nossa ANN precisa ser aumentado, de forma a conferir ainda maior capacidade de generalização à rede.

6. PERSPECTIVAS



O desenvolvimento de estratégias computacionais para auxiliar na descoberta de novos medicamentos está alvorecendo. Em nosso trabalho, estipulamos um modelo para avaliação da capacidade virucida de flavonoides, bem como uma proposta de rede neural artificial para seleção de flavonoides com atividade anti-RSV.

No entanto, nenhuma pesquisa é finita em si. Existem pontos no trabalho a serem melhorados e continuados:

- a) Pode-se estabelecer um elo entre a predição gerada pela rede neural artificial e o modelo de avaliação de capacidade virucida (RSVmF);
- b) O modelo de avaliação de capacidade virucida pode ser inserido numa membrana, ou podem ser retirados as porções transmembrânicas, citoplasmáticas, o peptídeo sinal e o peptídeo 27;
- c) O efeito de mutações também pode ser incluso no modelo RSVmF;
- d) Seria interessante se RSVmF fosse inserida em uma membrana e se avaliassem os impactos de flavonoides na membrana e qual efeito disso sobre a proteína;
- e) A ANN pode ser acrescida de alguns outros parâmetros, tais como quantidade de doadores/aceptores de elétrons, e quantidade de ligações rotacionáveis;
- f) O modelo de ANN pode ser estendido a outros vírus respiratórios (e.g. influenza, rinovírus), de forma a se poder ter um comparativo;
- g) O banco de dados BD1 da ANN pode ser usado num modelo de programação dinâmica, de maneira a se avaliar qual parâmetro tem mais peso na determinação de uma possível atividade.

Como um último adendo, gostaria de fazer algumas considerações pessoais.

Muito embora o papel da Ciência seja o de trazer soluções que beneficiem a humanidade como um todo, a dor também é fruto de ações danosas realizadas por ela, e, frequentemente, o desenvolvimento científico esbarra no pecado que tenazmente nos assedia de querermos controlar tudo. Poder. Devemos prezar pela ética, bem como pela

aproximação da Ciência à sociedade, na qual o debate e o esclarecimento devem sempre estar presentes. É desolador ver que, apesar de todo progresso científico (inteligência artificial, *machine learning*, internet das coisas) continuamos tendo como meta a erradicação do analfabetismo em 2018. É desalentador saber que a cada 35 segundos uma criança morre de pneumonia no mundo ou ainda que, a cada 3 segundos, uma morre de fome. 1... 2... 3... Não existe desenvolvimento tecnológico que não exerça impacto na sociedade, beneficiando-a ou sendo usado como subterfúgio para excluir ainda mais. E esse desenvolvimento se inicia, embora não exclusivamente, na Academia. A tecnologia depende essencialmente de quem está por trás dela, seja desenvolvendo, programando ou consumindo. E aqui há um grande perigo em ser apenas consumidor. Em um momento em que máquinas assumem afazeres humanos, no qual algoritmos são utilizados para desvelar nossas afinidades políticas, em que eleições são decididas com base na divulgação massificada de notícias falsas, onde pessoas utilizam democracia para clamar pelo retorno de ditaduras, no qual nos apercebemos envoltos por um noção forjadamente social tão frágil e tênue do poder da informação e do conhecimento, em que mesmo nossos sentimentos mais íntimos são traduzidos por algoritmos de “curtir”, e no qual ferramentas computacionais têm sido usadas com intuito de usurpar nossas consciências, URGE que retomemos nossa posição de seres humanos criativos, críticos e inventivos, que pensam, para repensarmos o direcionamento que temos dado à tecnologia. Sensibilidade, sensatez, dúvida são necessárias para fomentar a empatia e o árduo trabalho socrático de "conhecer-se a si mesmo", e de "saber que nada sabemos". A construção de um SER científico demanda bases de humildade, curiosidade e empatia. Assim, podemos desenvolver uma tecnologia que nos assista na busca pela nossa humanidade, que nos mostre *que não somos apenas algoritmos, mas somos um complexo aglomerado de matemática, lógica, na conjunção, injunção, exclusão da vida, em que a dor é necessária e lapidar essa dor de forma construtiva também*⁴. A dor, aquela de nos sentirmos sozinhos na imensidão do universo, ou talvez na curiosidade imensa

⁴ COSTA, P.F. Em “O sentido da Tecnologia”. Disponível em: <https://unicolegio.com/home/>

de não estarmos, a qual não precisa ser abrandada somente por remédios e controle de hormônios, mas também pela busca conjunta de conhecimento e construção de contextos. URGE que usemos os algoritmos para podermos sentir, sim!, sentir! Mas um *sentir que nos leve a buscar conhecimento, pois conhecimento dá contexto e contexto dá significado*⁴. Um sentir que nos leve a *tentar abrandar a nossa dor e a dor do outro* também. Essa é a proposta desse trabalho. Esse é o *sentido da tecnologia*⁴.

REFERÊNCIAS

- 1- WORLD HEALTH ORGANIZATION . Global Health Observatory Data (GHO). **Child mortality and causes of death**. Disponível em: http://www.who.int/gho/child_health/mortality/en/ . Acesso em 16 março 2018.
- 2- TREGONING, J.S.; SCHWARZE, J. Respiratory viral infections in infants: causes, clinical symptoms, virology and immunology. **Clinical Microbiology Reviews** v. 23, n.1, pp. 74-98, 2010. DOI: [10.1128/CMR.00032-09](https://doi.org/10.1128/CMR.00032-09).
- 3- FEARNES, R.; DEVAL, J. New antiviral approaches for respiratory syncytial virus and other mononegaviruses: Inhibiting the RNA polymerase. **Antiviral Research** v. 134, pp. 63-76, 2016. DOI: [10.1016/j.antiviral.2016.08.006](https://doi.org/10.1016/j.antiviral.2016.08.006) .
- 4- HUANG, K.; WU, H. Prevention of respiratory syncytial virus infection: from vaccine to antibody. **Microbiology Spectrum** v. 2, n. 4, 2014. DOI: [10.1128/microbiolspec.AID-0014-2014](https://doi.org/10.1128/microbiolspec.AID-0014-2014) .
- 5- NAIR, H. *et al.* Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. **Lancet** v. 375, n. 9725, pp. 1545-1555, 2010. DOI: [10.1016/S0140-6736\(10\)60206-1](https://doi.org/10.1016/S0140-6736(10)60206-1) .
- 6- PELTOLA, V.; RUUSKANEN, O. Respiratory viral infections in developing countries: common, severe, and unrecognized. **Clinical Infectious Disease** v. 46, n. 1, pp. 58-60, 2008. DOI: [10.1086/524020](https://doi.org/10.1086/524020) .
- 7- RUDAN, I. *et al.*, Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. **Journal of Global Health** v. 3, n. 1, 2013. DOI: [10.7189/jogh.03.010401](https://doi.org/10.7189/jogh.03.010401) .
- 8- VISSERS, M.; GROOT, R.; FERWERDA, G. Severe respiratory infections: are bugs bugging? **Mucosal Immunology** v. 7, n. 2, pp. 227-238, 2014. DOI: [10.1038/mi.2013.93](https://doi.org/10.1038/mi.2013.93)
- 9- HOMAIRA, N. *et al.*, Effectiveness of Palivizumab in preventing RSV hospitalization in high risk children: a real-world perspective. **International Journal of Pediatrics** v. 2014. DOI: [10.1155/2014/571609](https://doi.org/10.1155/2014/571609) .
- 10- LOZANO, R. *et al.* Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the global burden of disease study 2010. **Lancet** v. 380, pp. 2095-2128, 2012. DOI: [10.1016/S0140-6736\(12\)61728-0](https://doi.org/10.1016/S0140-6736(12)61728-0) .

- 11- LEYSSSEN, P.; DE CLERCQ, E.; NEYTS, J. Molecular strategies to inhibit the replication of RNA viruses. **Antiviral Research** v. 78, n. 1, pp. 9-25, 2008. DOI: [10.1016/j.antiviral.2008.01.004](https://doi.org/10.1016/j.antiviral.2008.01.004) .
- 12- GAROFALO, R.P.; KOLLI, D.; CASOLA, A. Respiratory Syncytial Virus infection: mechanisms of redox control and novel therapeutic opportunities. **Antioxidants & Redox Signaling** v. 18, n. 2, pp. 186-217, 2013. DOI: [10.1089/ars.2011.4307](https://doi.org/10.1089/ars.2011.4307) .
- 13- SIMÕES, E.A.F. *et al.* Challenges and opportunities in developing respiratory syncytial virus therapeutics. **The Journal of Infectious Diseases** v. 15, n. 211, pp. S1-S20, 2015. DOI: [10.1093/infdis/jiu828](https://doi.org/10.1093/infdis/jiu828) .
- 14- DUMONT, P. Increasing the complexity of respiratory syncytial virus infection: reactive oxygen species, DNA damage and premature senescence. **Virulence** v. 7, n. 4, pp. 372-375, 2016. DOI: [10.1080/21505594.2016.162370](https://doi.org/10.1080/21505594.2016.162370) .
- 15- CERVANTES-ORTIZ, S.L.; CUERVO, N.Z.; GRANDVAUX, N. Respiratory syncytial virus and cellular stress responses: impact on replication and physiopathology. **Viruses** v. 8, n. 124, 2016. DOI: [10.3390/v8050124](https://doi.org/10.3390/v8050124) .
- 16- HARIJITH, A.; EBENEZER, D.L.; NATARAJAN, V. Reactive oxygen species at the crossroads of inflammasome and inflammation. **Frontiers in Physiology** v. 5, n. 352, 2014. DOI: [10.3389/fphys.2014.00352](https://doi.org/10.3389/fphys.2014.00352) .
- 17- HOSAKOTE, Y.M. *et al.* Respiratory syncytial virus induces oxidative stress by modulating antioxidant enzymes. **American Journal of Respiratory Cell and Molecular Biology** v. 41, pp. 348-357, 2009. DOI: [10.1165/rcmb.2008-0330OC](https://doi.org/10.1165/rcmb.2008-0330OC) .
- 18- JAMALUDDIN, M. *et al.* Respiratory syncytial virus infection induces a reactive oxygen species-MSK1-Phospho-Ser-276 RelA pathway required for cytokine expression. **Journal of Virology** v. 83, n. 20, pp. 10605-10615, 2009. DOI: [10.1128/JVI.01090-09](https://doi.org/10.1128/JVI.01090-09) .
- 19- LI, M.; XU, Z. Quercetin in a lotus leaves extract may be responsible for antibacterial activity. **Archives of Pharmacal Research** v. 31, n. 5, pp. 640-644, 2008. DOI: [10.1007/s12272-001-1206-5](https://doi.org/10.1007/s12272-001-1206-5) .
- 20- GEOGHEGAN, F.; WONG, R.W.; RABIE, A.B. Inhibitory effect of quercetin on periodontal pathogens in vitro. **Phytotherapy Research** v. 24, n. 6, pp. 817-820, 2010. DOI: [10.1002/ptr.3014](https://doi.org/10.1002/ptr.3014) .
- 21- HIRAI, I. *et al.* Characterisation of anti-Staphylococcus aureus activity of quercetin. **Food Science and Technology** v. 45, pp. 1250-1254, 2010.

- 22- GOPU, V.; MEENA, C.K.; SHETTY, P.H. Quercetin influences quorum sensing in food borne bacteria: In-vitro and In-silico evidence. **Plos One** v. 6, 2015. DOI: [10.1371/journal.pone.0134684](https://doi.org/10.1371/journal.pone.0134684) .
- 23- PLAPER, A. *et al.* Characterization of quercetin binding site on DNA gyrase. **Biochemical and Biophysical Research Communications** v. 306, n. 2, pp. 530-536. DOI: [10.1016/S0006-291X\(03\)01006-4](https://doi.org/10.1016/S0006-291X(03)01006-4) .
- 24- GONZALEZ, O. *et al.* The heat shock protein inhibitor quercetin attenuates hepatitis C virus production. **Hepatology** v. 50, n. 6, pp. 1756-1764, 2009. DOI: [10.1002/hep.23232](https://doi.org/10.1002/hep.23232).
- 25- KAUL, T.N.; MIDDLETON, E.Jr.; OGRA, P. L. Antiviral effect of flavonoids on human viruses. **Journal of Medical Virology** v. 15, n. 1, pp. 71-79, 1985.
- 26- GANESAN, S. *et al.* Quercetin inhibits rhinovirus replication in vitro and in vivo. **Antiviral Research** v. 94, pp. 258-271, 2012. DOI: [10.1016/j.antiviral.2012.03.005](https://doi.org/10.1016/j.antiviral.2012.03.005) .
- 27- WU, W. *et al.* Quercetin as an antiviral agent inhibits Influenza A Virus (IAV) entry. **Viruses** v. 8, n. 1, 2016. DOI: [10.3390/v8010006](https://doi.org/10.3390/v8010006) .
- 28- ROJAS, A. *et al.* Effect of quercetin on Hepatitis C Virus life cycle: from viral to host targets. **Nature Scientific Reports** v. 6, n. 3177, 2016. DOI: [10.1038/srep31777](https://doi.org/10.1038/srep31777) .
- 29- ZANDI, K. *et al.* Antiviral activity of four types of bioflavonoid against dengue virus type-2. **Virology Journal** v. 8, n. 560, 2011. DOI: [10.1186/1743-422X-8-560](https://doi.org/10.1186/1743-422X-8-560) .
- 30- BUENO, F. S.. **Grande dicionário etmológico-prosódico da língua portuguesa**. São Paulo: Saraiva, 1963-1967. 8 v.
- 31- ABBAS, A.K. **Cellular and molecular immunology**. Canada: Elsevier, 2015, 8^a ed. ISBN: 978-0-323-22275-4.
- 32- **STANDFORD CHILDREN'S CARE**. Disponível em: <http://www.stanfordchildrens.org/en/topic/default?id=anatomy-of-the-respiratory-system-in-children-90-P02950>. Acesso em 10 março 2018.
- 33- FLINT, S. J. *et al.* **Principles of Virology**. Washington D.C : American Society of Microbiology, 2015, 4th Edition.
- 34- HITTINGER, M. *et al.* Preclinical safety and efficacy models for pulmonary drug delivery of antimicrobials with focus on in vitro models. **Advanced Drug Delivery Reviews** v. 85, pp. 44-56, 2015. DOI: [10.1016/j.addr.2014.10.011](https://doi.org/10.1016/j.addr.2014.10.011) .

- 35- HALL, C.D. et al. Infectivity of respiratory syncytial virus by various route of inoculation. **Infection and Immunity** v. 33, n. 3, pp. [779-783](#), 1981.
- 36- TROY, N.M.; BOSCO, A. Respiratory viral infections and host responses; insights from genomics. **Respiratory Research** v. 17, n. 156, 2016. DOI: [10.1186/s12931-016-0474-9](#).
- 37- FILHO, N.A. Bases históricas da Epidemiologia. **Cadernos de Saúde Pública** v. 2, n. 3, pp. 304-311, 1986.
- 38- ALVES, M.V.F.F. et al. Perfil dos pacientes internados em uma unidade de terapia intensiva pediátrica de um hospital escola do interior de São Paulo. **Ciência, Cuidado e Saúde** v. 13, n. 2, pp. 294-301, 2014. DOI: [10.4025/cienccuidsaude.v13i2.21912](#).
- 39- MONTEIRO, C.C.; DEZANET, L.N.C.; FRANÇA, E.B. Monitoramento de vírus respiratórios na região metropolitana de Belo Horizonte, 2011 a 2013. **Epidemiologia, Serviços e Saúde** v. 25, n. 2, pp. 233-242, 2016. DOI: [10.5123/S1679-49742016000200002](#).
- 40- OLIVEIRA, B.R.G. et al. Perfil de morbidade de crianças hospitalizadas em um hospital público: implicações para a Enfermagem. **Revista Brasileira de Enfermagem** v. 65, n. 4, pp. 586-593, 2012. DOI: [10.1590/S0034-71672012000400006](#).
- 41- CARDOSO, A. M. A persistência das infecções respiratórias agudas como problema de Saúde Pública. **Caderno de Saúde Pública** v. 26, n. 7, pp. 1270-1271, 2010. Versão On-line [ISSN 1678-4464](#).
- 42- OLIVEIRA, B.R.G. et al. Causas da hospitalização no SUS de crianças de zero a quatro anos no Brasil. **Revista Brasileira de Epidemiologia** v. 13, n. 2, pp. 268-277, 2010. DOI: [10.1590/S1415-790X2010000200009](#).
- 43- THOMAZELLI, L.M. et al. Surveillance of eight respiratory viruses in clinical samples of pediatric patients in Southeast Brazil. **Jornal de Pediatria** v. 83, n. 5, pp. 422-428, 2007. DOI: [10.2223/JPED.1694](#).
- 44- CAETANO, J.R.M. et al. Fatores associados à internação hospitalar de crianças menores de cinco anos, São Paulo, SP. **Revista Saúde Pública** v. 36, n. 3, pp. 285-291, 2002. DOI: [10.1590/S0034-89102002000300005](#).
- 45- CALY, L.; GHILDYAL, R.; JANS, D.A. Respiratory virus modulation of host nucleocytoplasmic transport; target for therapeutic intervention? **Frontier in Microbiology** v. 6, n. 848, 2015. DOI: [10.3389/fmicb.2015.00848](#).

- 46- LAY, M.K. *et al.* New insights on the viral and host factors contributing to the airway pathogenesis caused by the respiratory syncytial virus. **Critical Reviews in Microbiology** v. 42, n. 5, pp. 800-812, 2015. DOI: [10.3109/1040841X.2015.1055711](https://doi.org/10.3109/1040841X.2015.1055711) .
- 47- LEGAND, A. *et al.*, Addressing the public health burden of respiratory viruses: the battle against Respiratory Viruses (BRaVe) Initiative. **Future Virology** v. 8, n. 10, pp. 953-968, 2010. DOI: [10.2217/fvl.13.85](https://doi.org/10.2217/fvl.13.85) .
- 48- MATHERS, C.D.; LONCAR, D. Projections of global mortality and burden of disease from 2002 to 2030. **Plos Medicine** v. 3, n.11, pp. 2011-2030, 2006. DOI:[10.1371/journal.pmed.0030442](https://doi.org/10.1371/journal.pmed.0030442) .
- 49- WORLD HEALTH ORGANIZATION. **The global burden of disease: 2004 update**. ISBN 978 92 4 156371 0. Disponível em: http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/ . Acesso em 16 março 2018.
- 50- MINISTÉRIO DA SAÚDE. Indicadores e Dados Básicos – Brasil 2012 . IDB -2012. Indicadores de mortalidade – C7. Proporção de óbitos por infecção respiratória aguda em menores de 5 anos de idade. Disponível em: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?idb2012/c07.def> . Acesso em 17 março 2018.
- 51- SILVA, D.A.O.; PEDRON, C.D.; CABERLON, I.C. Perfil epidemiológico dos moradores de 0 a 14 anos da área II do bairro São Vicente da cidade de Gravataí/RS. **Revista Ampliar** v. 2, n. 2, 2015.
- 52- RIO GRANDE DO SUL. Secretaria de Planejamento. Gestão e Participação Cidadã. **Indicadores Sociais**: coeficiente de mortalidade infantil. Porto Alegre: SEPLAG, 2011.
- 53- WORLD HEALTH ORGANIZATION. **Newborns: Reducing Mortality**. Disponível em: <http://www.who.int/mediacentre/factsheets/fs333/en/>. Acesso em 17 março 2018.
- 54- BRASIL. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. **Lei nº 8069, de 13 de julho de 1990. Dispõe sobre o Estatuto da Criança e do Adolescente e dá providências**. Brasília: Presidência da República, 1990.
- 55- UNICEF, One is too many: Ending child deaths from pneumonia and diarrhea. **UNICEF**, New York, 2016. ISBN: [978-92-806-4859-1](https://doi.org/10.2478/978-92-806-4859-1) .
- 56- INTERNATIONAL COMMITTEE ON TAXONOMY OF VIRUSES. **ICTV Taxonomy history: Human orthopneumovirus**. Disponível em:

https://talk.ictvonline.org/taxonomy/p/taxonomy-history?taxnode_id=20161069. Acesso em 24 fevereiro 2018.

- 57- BLOUNT, R.E. Jr.; MORRIS, J.A.; SAVAGE, R.E. Recovery of cytopathogenic agent from chimpanzees with coryza. **Proceedings of the Society for Experimental Biology and Medicine** v. 92, n. 3, pp. 544-549, 1956.
- 58- GRAHAM, B.S. Vaccine development for respiratory syncytial virus. **Current opinion in Virology** v. 23, pp. 107-112, 2017. DOI: [10.1016/j.coviro.2017.03.012](https://doi.org/10.1016/j.coviro.2017.03.012) .
- 59- PARRISH, J. Observations on a peculiar catarrhal complaint in children. **The North American Medical and Surgical Journal** v. 1, pp. 24-31, 1826.
- 60- GOODPASTURE, E.W. *et al.* Virus pneumonia of infants secondary to epidemic infections. **American Journal of Disease of Children** v. 57, n. 5, pp. 997-1011, 1939. DOI: [10.1001/archpedi.1939.01990050003001](https://doi.org/10.1001/archpedi.1939.01990050003001).
- 61- ADAMS, J.M. Primary virus pneumonitis with cytoplasmic inclusion bodies: study of epidemic involving 32 infants with 9 deaths. **JAMA** v. 116, n. 10, pp. 925-933, 1941. DOI: [10.1001/jama.1941.02820100019005](https://doi.org/10.1001/jama.1941.02820100019005) .
- 62- ADAMS, J.M. Primary virus pneumonitis in infants. **JAMA** v. 138, n. 16, pp. 1142-1144, 1948. DOI: [10.1001/jama.1948.02900160010004](https://doi.org/10.1001/jama.1948.02900160010004) .
- 63- CHANOCK, R.; ROIZMAN, B.; MYERS, R. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent. I. Isolation, properties and characterization. **American Journal of Hygiene** v. 66, n. 3, pp. 281-290, 1957.
- 64- **VIRALZONE**. Syncytium formation induced by viral infection. Disponível em: http://viralzone.expasy.org/5957?outline=all_by_species . Acesso em 13 fevereiro 2018.
- 65- OPENSHAW, P.J.M. *et al.* Protective and harmful immunity to RSV infection. **Annual Reviews of Immunology** v. 35, pp. 501-532, 2017. DOI: [10.1146/annurev-immunol-051116-052206](https://doi.org/10.1146/annurev-immunol-051116-052206) .
- 66- BOYCE, T. G. *et al.* Rates of hospitalization for respiratory syncytial virus infection among children in medicaid. **The Journal of Pediatrics** v. 137, n. 6, pp. 865–870, 2000. DOI: [10.1067/mpd.2000.110531](https://doi.org/10.1067/mpd.2000.110531).
- 67- WANG, E. E.; LAW, B.J.; STEPHENS, D. Pediatric Investigators Collaborative Network on Infections in Canada (PICNIC) prospective study of risk factors and outcomes in patients hospitalized with respiratory syncytial viral lower respiratory tract infection. **The Journal of Pediatrics** v. 126, n. 2, pp. [212–219](#), 1995.

- 68- NAVAS, L. *et al.* Improved outcome of respiratory syncytial virus infection in a high-risk hospitalized population of Canadian children. Pediatric Investigators Collaborative Network on Infections in Canada. **The Journal of Pediatrics** v. 121, n. 3, pp. 348–354, 1992. DOI: [10.1016/S0022-3476\(05\)90000-0](https://doi.org/10.1016/S0022-3476(05)90000-0) .
- 69- AUJARD, Y.; FAUROUX, B. Risk factors for severe respiratory syncytial virus infection in infants. **Respiratory Medicine**, 2002. DOI: [10.1053/rmed.2002.1295](https://doi.org/10.1053/rmed.2002.1295) .
- 70- MOYES, J. *et al.* Epidemiology of respiratory syncytial virus-associated acute lower respiratory tract infection hospitalizations among HIV-infected and HIV-uninfected South African children, 2010-2011. **The Journal of Infectious Diseases** v. 208, n. SUPPL. 3, 2013. DOI: [10.1093/infdis/jit479](https://doi.org/10.1093/infdis/jit479) .
- 71- ABBAS, S. *et al.* Respiratory viruses in transplant recipients: more than just a cold. Clinical syndromes and infection prevention principles. **International Journal of Infectious Diseases** v. 62, pp. 86-93, 2017. DOI: [10.1016/j.ijid.2017.07.011](https://doi.org/10.1016/j.ijid.2017.07.011) .
- 72- EISENHUT, M. Extrapulmonary manifestations of severe respiratory syncytial virus infection – a systematic review. **Critical Care** v. 10, n. 4, 2006. DOI: [10.1186/cc4984](https://doi.org/10.1186/cc4984) .
- 73- GRIFFITHS, C.; DREWS, S.J.; MARCHANT, D.J. Respiratory syncytial virus: infection, detection, and new options for prevention and treatment. **Clinical Microbiology Reviews** v. 30, n.1, pp. 227-319, 2017. DOI: [10.1128/CMR.00010-16](https://doi.org/10.1128/CMR.00010-16) .
- 74- GOWER, T.L. *et al.* Rhoa signaling is required for respiratory syncytial virus-induced syncytium formation and filamentous virion morphology. **Journal of Virology** v. 79, n. 9, pp. 5326-5336, 2005. DOI: [10.1128/JVI.79.9.5326-5336.2005](https://doi.org/10.1128/JVI.79.9.5326-5336.2005) .
- 75- MUFSON, M.A. *et al.* Two distinct subtypes of human respiratory syncytial virus. **Journal of General Virology** v. 66, pp. 2111-2124, 1985. DOI: [10.1099/0022-1317-66-10-2111](https://doi.org/10.1099/0022-1317-66-10-2111) .
- 76- SIQUEIRA, M.M.; NASCIMENTO, J.P.; ANDERSON, L.J. Antigenic characterization of respiratory syncytial virus group A and B isolates in Rio de Janeiro, Brazil. **Journal of Clinical Microbiology** v. 29, n. 3, pp. 557-559, 1991.
- 77- SULLENDER, W.A. Respiratory syncytial virus genetic and antigenic diversity. **Clinical Microbiology Reviews** v. 13, n. 1, pp. 1-15, 2000. DOI: [10.1128/CMR.13.1.1-15.2000](https://doi.org/10.1128/CMR.13.1.1-15.2000) .
- 78- CANEDO-MARROQUIN, G. *et al.* Modulation of host immunity by human respiratory syncytial virus virulence factors: a synergic inhibition of both innate and adaptive immunity. **Frontiers in Cellular and Infection Microbiology** v. 7, n. 367, 2017. DOI: [10.3389/fcimb.2017.00367](https://doi.org/10.3389/fcimb.2017.00367) .

- 79- SWEDAN, S. *et al.* Multiple functional domains and complexes of the two nonstructural proteins of human respiratory syncytial virus contribute to interferon suppression and cellular location. **Journal of Virology** v. 85, n. 19, pp. 10090-10100, 2011. DOI: [10.1128/JVI.DD413-11](https://doi.org/10.1128/JVI.DD413-11) .
- 80- SPANN, K.M. *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. **Journal of Virology** v. 78, n. 12, pp. 4363-4369, 2004. DOI: [10.1128/JVI.78.8.4363-4369.2004](https://doi.org/10.1128/JVI.78.8.4363-4369.2004) .
- 81- FEARNES, R.; PEEPLES, M.E.; COLLINS, P.L. Increased expression of the N protein of respiratory syncytial virus stimulates minigenome replication but does not alter the balance between the synthesis of mRNA and antigenome. **Virology** v. 236, pp. 188-201, 1997. DOI: [10.1006/viro.1997.8734](https://doi.org/10.1006/viro.1997.8734) .
- 82- SCHMIDT, M.E.; VARGA, S.M. Modulation of the host immune response by respiratory syncytial virus proteins. **Journal of Microbiology** v. 55, n. 3, pp. 161-171, 2017. DOI: [10.1007/s12275-017-7045-8](https://doi.org/10.1007/s12275-017-7045-8) .
- 83- ROSE, P.W. *et al.* The RCSB protein data bank: integrative view of protein, gene and 3D structural information. **Nucleic Acids Research** v. 45: D271-D281, 2017. DOI: [10.1093/nar/gkw1000](https://doi.org/10.1093/nar/gkw1000) .
- 84- PETTERSEN, E.F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. **Journal of Computational Chemistry** v. 25, n. 13, pp.1605–1612, 2004. DOI: [10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084) .
- 85- ASENJO, A.; GONZÁLEZ-ARMAS, J.C.; VILLANUEVA, N. Phosphorylation of human respiratory syncytial virus P protein at serine 54 regulates viral uncoating. **Virology** v. 380, n.1, pp. 26-33, 2008. DOI: [10.1016/j.virol.2008.06.045](https://doi.org/10.1016/j.virol.2008.06.045) .
- 86- MITRA, R. *et al.* The human respiratory syncytial virus matrix protein is required for maturation of viral filaments. **Journal of Virology** v. 86, n.8, pp. 4432-4443, 2012. DOI: [10.1128/JVI.06744-11](https://doi.org/10.1128/JVI.06744-11) .
- 87- FÖRSTER, A. *et al.* Dimerization of matrix protein is required for budding of respiratory syncytial virus. **Journal of Virology** v. 89, n. 8, pp. 4624-4635, 2015. DOI: [10.1128/JVI.03500-14](https://doi.org/10.1128/JVI.03500-14) .
- 88- McLELLAN, J.S.; RAY, W.C.; PEEPLES, M.E. Structure and function of RSV Glycoproteins. **Current topics in Microbiology and Immunology** v. 373, pp. 83-104, 2013. DOI: [10.1007/978-3-642-38919-1_4](https://doi.org/10.1007/978-3-642-38919-1_4) .

- 89- GODER, V.; SPIESS, M. Topogenesis of membrane proteins: determinants and dynamics. **FEBS Letters** v. 504, n. 3, pp. 87-93, 2001. DOI: [10.1016/S0014-5793\(01\)02712-0](https://doi.org/10.1016/S0014-5793(01)02712-0) .
- 90- GAN, S.W. *et al.* The small hydrophobic protein of the human respiratory syncytial virus forms pentameric ion channels. **Journal of Biological Chemistry** v. 287, pp. 24671-24689, 2012. DOI: [10.1074/jbc.M111.332791](https://doi.org/10.1074/jbc.M111.332791) .
- 91- TRIANTAFILOU, K. *et al.* Human respiratory syncytial virus viroporin SH: a viral recognition pathway used by the host to signal inflammasome activation. **Thorax** v. 68, n. 1, pp. 66-75, 2013. DOI: [10.1136/thoraxjnl-2012-202182](https://doi.org/10.1136/thoraxjnl-2012-202182) .
- 92- LEVINE, S.; KLAIBER-FRANCO, R.; PARADISO, PR. Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. **The Journal of General Virology** v. 68, n. 9, pp. 2521-2524, 1987. DOI: [10.1099/0022-1317-68-9-2521](https://doi.org/10.1099/0022-1317-68-9-2521) .
- 93- ROBERTS, S.R. *et al.* The membrane-associated and secreted forms of the respiratory syncytial virus attachment glycoprotein G are synthesized from alternative initiation codons. **Journal of Virology** v. 68, n. 7, pp. [4538-4546](#), 1994.
- 94- HENDRICKS, D.A.; McINTOSH, K.; PATTERSON, J.L. Further characterization of the soluble form of the G glycoprotein of respiratory syncytial virus. **Journal of Virology** v. 62, n.7, pp. [2228-2233](#), 1988.
- 95- BUKREYEV, A.; YANG, L.; COLLINS, P.L. The secreted G protein of human respiratory syncytial virus antagonizes antibody-mediated restriction of replication involving macrophages and complement. **Journal of Virology** v. 86, n. 19, pp. 10880-10884, 2012. DOI: [10.1128/JVI.01162-12](https://doi.org/10.1128/JVI.01162-12) .
- 96- JOHNSON, P.R. *et al.* The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. **Proceeding of the National Academy of Sciences of the United States of America** v. 84, n.16, pp. [5625-5629](#), 1987.
- 97- RUEDA, P.; GARCÍA-BARRENO, B.; MELERO, J.A. Loss of conserved cysteine residues in the attachment (G) glycoprotein of two human respiratory syncytial virus escape mutants that contain multiple A-G substitutions (hypermutations). **Virology** v. 198, n.2, pp. [653-662](#), 1994.
- 98- CHIRKOVA, T. *et al.* CX3CR1 is an important surface molecule for respiratory syncytial virus infection in human airway epithelial cells. **The Journal of General Virology** v. 96, n.8, pp. 2543-2556, 2015. DOI: [10.1099/vir.0.000218](https://doi.org/10.1099/vir.0.000218) .

- 99- JOHNSON, S.M. *et al.* Respiratory syncytial virus uses CX3CR1 as a receptor on primary human airway epithelial cultures. **Plos Pathogens** v. 11, n.12, e1005318, 2015. DOI: [10.1371/journal.ppat.1005318](https://doi.org/10.1371/journal.ppat.1005318) .
- 100- JEONG, K. *et al.* CX3CR1 is expressed in differentiated human ciliated airway cells and co-localizes with respiratory syncytial virus on cilia in a G protein-dependent manner. **Plos One** v. 10, n. 6, e0130517. DOI: [10.1371/journal.pone.0130517](https://doi.org/10.1371/journal.pone.0130517) .
- 101- CHIRKOVA, T. *et al.* Respiratory syncytial virus G protein CX3C motif impairs human airway epithelial and immune cell responses. **Journal of Virology** v. 87, n. 24, pp. 13466-13479, 2013. DOI: [10.1128/JVI.01741-13](https://doi.org/10.1128/JVI.01741-13) .
- 102- TRIPP, R.A. *et al.* CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. **Nature Immunology** v. 2, n.8, pp. 732-738, 2001. DOI: [10.1038/90675](https://doi.org/10.1038/90675) .
- 103- ZHANG, L. *et al.* Infection of ciliated cells by human parainfluenza virus type 3 in an in vitro model of human airway epithelium. **Journal of Virology** v. 79, n.2, pp. 113-1124, 2005. DOI: [10.1128/JVI.79.2.1113-1124.2005](https://doi.org/10.1128/JVI.79.2.1113-1124.2005) .
- 104- COLLINS, P.L.; HUANG, Y.T.; WERTZ, G.W. Nucleotide sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus. **Proceedings of the National Academy of Sciences of the United States of America** v. 81, n. 24, pp. 7683-7687, 1984.
- 105- LILJEROOS, L. *et al.* Architecture of respiratory syncytial virus revealed by electron cryotomography. **Proceeding of the National Academy of Sciences of the United States of America** v. 110, n. 27, pp. 11133-11138, 2013. DOI: [10.1073/pnas.1309070110](https://doi.org/10.1073/pnas.1309070110) .
- 106- GONZÁLEZ-REYES, L. *et al.* Cleavage of the human respiratory syncytial virus fusion protein at two distinct sites is required for activation of membrane fusion. **Proceedings of the National Academy of Sciences of the United States of America** v. 98, n. 17, pp. 9859-9864, 2001. DOI: [10.1073/pnas.151098198](https://doi.org/10.1073/pnas.151098198) .
- 107- ZIMMER, G.; BUDZ, L.; HERRLER, G. Proteolytic activation of respiratory syncytial virus fusion protein. Cleavage at two furin consensus sequences. **The Journal of Biological Chemistry** v. 276, n. 34, pp. 31642-31650, 2001. DOI: [10.1074/jbc.M102633200](https://doi.org/10.1074/jbc.M102633200) .
- 108- BOLT, G.; PEDERSEN, L.O.; BIRKESLUND, H.H. Cleavage of the respiratory syncytial virus fusion protein is required for its surface expression: role of furin. **Virus Research** v. 68, n.1, pp. 25-33, 2000.

- 109- COLLINS, P.L.; MOTTET, G. Post-translational processing and oligomerization of the fusion glycoprotein of human respiratory syncytial virus. **The Journal of General Virology** v. 72, pp. 3095-3101, 1991. DOI: [10.1099/0022-1317-72-12-3095](https://doi.org/10.1099/0022-1317-72-12-3095) .
- 110- GRUBER, C.; LEVINE, S. Respiratory syncytial virus polypeptides. III. The envelope-associated proteins. **The Journal of General Virology** v. 64, pp. 825-832, 1983. DOI: [10.1099/0022-1317-64-4-825](https://doi.org/10.1099/0022-1317-64-4-825) .
- 111- DAY, N.D. *et al.* Contribution of cysteine residues in the extracellular domain of the F protein of human respiratory syncytial virus to its function. **Virology Journal** v. 3, n. 34, 2006. DOI: [10.1186/1743-422X-3-34](https://doi.org/10.1186/1743-422X-3-34) .
- 112- ROSSEY, I. *et al.* Clinical potential of prefusion RSV F-specific antibodies. **Trends in Microbiology** v. 26, n. 3, pp. 209-219, 2018. DOI: [10.1016/j.tim.2017.09.009](https://doi.org/10.1016/j.tim.2017.09.009) .
- 113- HOLGUERA, J.; VILLAR, E.; MUÑOZ-BARROSO, I. Identification of cellular proteins that interact with Newcastle disease virus and human respiratory syncytial virus by a two-dimensional virus overlay protein binding assay (VOPBA). **Virus Research** v. 191, pp. 138-142, 2014. DOI: [10.1016/j.virusres.2014.07.031](https://doi.org/10.1016/j.virusres.2014.07.031) .
- 114- TAYYARI, F. *et al.* Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. **Nature Medicine** v. 17, pp. 1131-1135, 2011. DOI: [10.1038/nm.2444](https://doi.org/10.1038/nm.2444) .
- 115- BEHERA, A.K. *et al.* Blocking intercellular adhesion molecule-1 on human epithelial cells decreases respiratory syncytial virus infection. **Biochemical and Biophysical Research Communications** v. 280, n.1, pp. 188-195, 2001. DOI: [10.1006/bbrc.2000.4093](https://doi.org/10.1006/bbrc.2000.4093) .
- 116- MARCHANT, D. *et al.* Toll-like-receptor 4-mediated activation of p38 mitogen-activated protein kinase is determinant of respiratory virus entry and tropism. **Journal of Virology** v. 84, n.21, pp. 11359-11373, 2010. DOI: [10.1128/JVI.00804-10](https://doi.org/10.1128/JVI.00804-10) .
- 117- KURT-JONES, E. *et al.* Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. **Nature Immunology** v. 1, pp. 398-401, 2000. DOI: [10.1038/80833](https://doi.org/10.1038/80833) .
- 118- TECHAARPORNKUL, S., BARRETTO, N., AND PEEPLES, M. E. (2001). Functional analysis of recombinant respiratory syncytial virus deletion mutants lacking the small hydrophobic and/or attachment glycoprotein gene. **Journal of Virology** v. 75, n. 15, pp. 6825–6834, 2001. DOI: [10.1128/JVI.75.6825-6834.2001](https://doi.org/10.1128/JVI.75.6825-6834.2001) .

- 119- **UNIPROT**: Coiled Coil. Disponível em: <https://www.uniprot.org/help/coiled> . Acesso em 06 março 2018.
- 120- COLLINS, P. L. *et al.* Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. **Proceedings of the National Academy of Sciences of the United States of America** v. 93, n.1, pp. [81–85](#), 1996.
- 121- LI, D. *et al.* Association of Respiratory Syncytial virus M protein with viral nucleocapsids is mediated by the M2-1 protein. **Journal of Virology** v. 82, n.17, pp. 8863–8870, 2008. DOI: [10.1128/JVI.00343-08](#) .
- 122- 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. **Proceedings of the National Academy of Sciences of the United States of America** v. 96, n.20, pp. [11259–11264](#), 1999.
- 123- WALSH, E.E.; FALSEY, A.R. Respiratory syncytial virus infection in adult populations. **Infectious Disorders Drug Targets** v. 12, n.2, pp. [98-102](#), 2012.
- 124- CROMER, D. *et al.* Burden of paediatric respiratory syncytial virus disease and potential effect of different immunization strategies: a modelling and cost-effectiveness analysis for England. **Lancet Public Health** v. 2, e367-e374, 2017. DOI: [10.1016/S2468-2667\(17\)30103-2](#) .
- 125- MITCHELL, I.; DEFOY, I.; GRUBB, E. Burden of respiratory syncytial virus hospitalizations in Canada. **Canadian Respiratory Journal** v. 2017, 2017. DOI: [10.1155/2017/4521302](#) .
- 126- YOU, D. *et al.* Exposure of neonates to respiratory syncytial virus is critical in determining subsequent airway response in adults. **Respiratory Research** v. 7, 2006. DOI: [10.1186/1465-9921-7-107](#) .
- 127- CARROLL, K.N. *et al.* Season of infant bronchiolitis and estimates of subsequent risk and burden of early childhood asthma. **Journal of Allergy and Clinical Immunology** v. 123, n.4, pp. 964-966, 2009. DOI: [10.1016/j.jaci.2008.12.011](#) .
- 128- SIGURS, N. *et al.* Asthma and allergy patterns over 18 years after severe RSV bronchiolitis in the first year of life. **Thorax** v. 65, n.12, pp. 1045-1052, 2010. DOI: [10.1136/thx.2009.121582](#) .
- 129- BLANKEN, M.O. *et al.* Respiratory syncytial virus and recurrent wheeze in healthy preterm infants. **The New England Journal of Medicine** v. 368, pp. 1791-1799, 2013. DOI: [10.1056/NEJMoa1211917](#) .

- 130- GBD 2015. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. **Lancet Respiratory Medicine** v. 5, pp. 691-706, 2017. DOI: [10.1016/S2213-2600\(17\)30293-X](https://doi.org/10.1016/S2213-2600(17)30293-X) .
- 131- BRASIL. Ministério Da Saúde – Secretaria De Vigilância Em Saúde. **Boletim Epidemiológico: Perfil da morbimortalidade por doenças respiratórias crônicas no Brasil, 2003 a 2013** v. 47, n. 19, 2016. ISSN [2358-9450](https://doi.org/10.1186/2358-9450) .
- 132- COLLINS, P.L.; MELERO, J.A. Progress in understanding and controlling respiratory syncytial virus: still crazy after all these years. **Virus Research** v. 162, n.1-2, pp. 80-99, 2011. DOI: [10.1016/j.virusres.2011.09.020](https://doi.org/10.1016/j.virusres.2011.09.020) .
- 133- COLLINS, P.L.; FEARN, R.; GRAHAM, B.S. Respiratory syncytial virus: virology, reverse genetics, and pathogenesis of disease. **Current Topics in Microbiology and Immunology** v. 372, pp. 3-38, 2013. DOI: [10.1007/978-3-642-38919-1_1](https://doi.org/10.1007/978-3-642-38919-1_1) .
- 134- KRZYZANIAK, M.A. *et al.* Host cell entry of respiratory syncytial virus involves macropinocytosis followed by proteolytic activation of the F protein. **Plos Pathogens** v. 9, n.4, e1003309, 2012. DOI: [10.1371/journal.ppat.1003309](https://doi.org/10.1371/journal.ppat.1003309) .
- 135- LIFLAND, A.W. *et al.* Human respiratory syncytial virus nucleoprotein and inclusion bodies antagonize the innate immune response mediated by MDA5 and MAVS. **Journal of Virology** v. 86, n. 15, pp. 8245-8258, 2012. DOI: [10.1128/JVI.00215-12](https://doi.org/10.1128/JVI.00215-12) .
- 136- TENG, M.N.; COLLINS, P.L. Identification of the respiratory syncytial virus proteins required for formation and passage of helper-dependent infectious particles. **Journal of Virology** v. 72, n. 7, pp. [5707-5716](https://doi.org/10.1128/JVI.72.7.5707-5716), 1998.
- 137- LAY, M.K. *et al.* Advances in understanding respiratory syncytial virus infection in airway epithelial cells and consequential effects on the immune response. **Microbes and Infection** v. 15, n. 3, pp. 230-242, 2013. DOI: [10.1016/j.micinf.2012.11.012](https://doi.org/10.1016/j.micinf.2012.11.012) .
- 138- ZENG, R. *et al.* Pattern recognition receptors for respiratory syncytial virus infection and design of vaccines. **Virus Research** v. 167, n.2, pp. 138-145, 2012. DOI: [10.1016/j.virusres.2012.06.003](https://doi.org/10.1016/j.virusres.2012.06.003) .
- 139- McNAMARA, P.S. *et al.* Production of chemokines in the lungs of infants with severe respiratory syncytial virus bronchiolitis. **The Journal of Infectious Diseases** v. 191, n. 8, pp. 1225-1232, 2005. DOI: [10.1086/428855](https://doi.org/10.1086/428855) .

- 140- YOON, J.S. *et al.* Cytokine induction by respiratory syncytial virus and adenovirus in bronchial epithelial cells. **Pediatric Pulmonology** v. 42, n.3, pp. 277-282, 2007. DOI: [10.1002/ppul.20574](https://doi.org/10.1002/ppul.20574) .
- 141- MURPHY, B.R. *et al.* Serum and nasal-wash immunoglobulin G and A antibody response of infants and children to respiratory syncytial virus F and G glycoproteins following primary infection. **Journal of Clinical Microbiology** v. 23, n.6, pp. [1009-1014](https://doi.org/10.1093/jc.23.6.1009), 1986.
- 142- BRANDENBURG, A.H. Respiratory syncytial virus specific serum antibodies in infants under six months of age: limited serological response upon infection. **Journal of Medical Virology** v. 52, n. 1, pp. [97-104](https://doi.org/10.1093/jmv.52.1.97), 1997.
- 143- FALSEY, A.R.; SINGH, H.K.; WALSH, E.E. Serum antibody decay in adults following natural respiratory syncytial virus infection. **Journal of Medical Virology** v. 78, n. 11, pp. 1493-1497, 2006. DOI: [10.1002/jmv.20724](https://doi.org/10.1002/jmv.20724) .
- 144- ELLION, G.B. Acyclovir: discovery, mechanism of action, and selectivity. *Journal of Medical Virology* Supplement 1, pp. 2-6, 1993.
- 145- CIANCI, C. *et al.* Orally active fusion inhibitor of respiratory syncytial virus. **Antimicrobial Agents and Chemotherapy** v. 48, n.2, pp. 413-422, 2004. DOI: [10.1128/AAC.48.2.413-422.2004](https://doi.org/10.1128/AAC.48.2.413-422.2004) .
- 146- McLELLAN, J.S. *et al.* Structure of respiratory syncytial virus fusion glycoprotein in the post fusion conformation reveals preservation of neutralizing epitopes. **Journal of Virology** v. 85, n. 15, pp. 7788-7796, 2011. DOI: [10.1128/JVI.00555-11](https://doi.org/10.1128/JVI.00555-11) .
- 147- ARBIZA, J. *et al.* Characterization of two antigenic sites recognized by neutralizing monoclonal antibodies directed against the fusion glycoprotein of human respiratory syncytial virus. **Journal of General Virology** v. 73, n. 9, 1992. DOI: [10.1099/0022-1317-73-9-2225](https://doi.org/10.1099/0022-1317-73-9-2225) .
- 148- SUBRAMANIAN, K.N. *et al.* Safety, tolerance and pharmacokinetics of a humanized monoclonal antibody to respiratory syncytial virus in premature infants and infants with bronchopulmonary dysplasia. MEDI-493 Study Group. **The Pediatric Infectious Disease Journal** v. 17, n. 2, pp. [110-115](https://doi.org/10.1097/01.pid.0b013e3180000000), 1998.
- 149- COMMITTEE ON INFECTIOUS DISEASES AND BRONCHIOLITIS GUIDELINES. Updated guidance for palivizumab prophylaxis among infants and young children at increased risk of hospitalization for respiratory syncytial virus infection. **Pediatrics** v. 134, n. 2, 2014.

- 150- ZHAO, X.; SULLENDER, W.M. In vivo selection of respiratory syncytial viruses resistant to palivizumab. **Journal of Virology** v. 79, n. 7, pp. 3962-3968, 2005. DOI: [10.1128/JVI.79.7.3962-3968.2005](https://doi.org/10.1128/JVI.79.7.3962-3968.2005).
- 151- ZHAO, X. *et al.* Variable resistance to palivizumab in cotton rats by respiratory syncytial virus mutants. **The Journal of Infectious Diseases** v. 190, n. 11, pp. 1941-1946, 2004. DOI: [10.1086/425515](https://doi.org/10.1086/425515).
- 152- ZHU, Q. *et al.* Analysis of respiratory syncytial virus preclinical and clinical variants resistant to neutralization by monoclonal antibodies palivizumab and/or motavizumab. **The Journal of Infectious Disease** v. 203, n. 3, pp. 674-682, 2011. DOI: [10.1093/infdis/jiq100](https://doi.org/10.1093/infdis/jiq100).
- 153- SYNAGIS PRICES, COUPONS AND PATIENT ASSISTANCE PROGRAMS. Disponível em: <https://www.drugs.com/price-guide/synagis>. Acesso em 24 fevereiro 2018.
- 154- BREKKER, O.H.; SANDLE, I. Therapeutic antibodies for human diseases at the dawn of the twenty-first century. **Nature Reviews Drug Discovery** v. 2, pp. 52-62, 2003. DOI: [10.1038/nrd984](https://doi.org/10.1038/nrd984).
- 155- DAINA, A.; MICHIELIN, O.; ZOETE, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. **Scientific reports** v. 7, n. 42717, 2017. DOI: [10.1038/srep42717](https://doi.org/10.1038/srep42717).
- 156- FIELDS, B. N., IN KNIPE, D. M., & IN HOWLEY, P. M. (2013). Fields virology. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.
- 157- LEHNINGER, A.L.; DAVID, L. 1942- NELSON E MICHAEL M. COX. Lehninger Principles of Biochemistry. 5^a ed. New York ; New Delhi: W.H. Freeman, 2008.
- 158- GADOMSKI, A.M.; BHASALE, A.L. Bronchodilators for bronchiolitis. **Cochrane Database of Systematic Reviews** v. 19, n. 3, 2006. DOI: [10.1002/14651858.CD001266.pub2](https://doi.org/10.1002/14651858.CD001266.pub2).
- 159- DUDAS, R.A.; KARRON, R.A. Respiratory syncytial virus vaccines. **Clinical Microbiology Reviews**, v. 11, n. 3, pP. 430-439, 1998.
- 160- KIM, J.Y.; CHANG, J. In hot pursuit of the first vaccine against respiratory syncytial virus. **Yonsei Medical Journal** v. 57, n. 4, pp. 809-816, 2016. DOI: [10.3349/ymj.2016.57.4.809](https://doi.org/10.3349/ymj.2016.57.4.809).
- 161- GREEN, C.A. *et al.*, Safety and immunogenicity of novel respiratory syncytial virus (RSV) vaccines based on the RSV viral proteins F, N and M2-1 encoded by simian

adenovirus (PanAd3-RSV) and MVA (MVA-RSV); protocol for an open-label, dose-escalation, single-centre, phase 1 clinic trial in healthy adults. **BMJ Open** v 5, n. 10, 2015. DOI: [10.1136/bmjopen-2015-008748](https://doi.org/10.1136/bmjopen-2015-008748) .

- 162- PAPADOULOS, N.G. *et al.* Promising approaches 1 for treatment and prevention of viral respiratory illnesses. **The Journal of Allergy and Clinical Immunology** v. 140, n. 4, pp. 921-932, 2017. DOI: [10.1016/j.jaci.2017.07.001](https://doi.org/10.1016/j.jaci.2017.07.001).
- 163- JORQUERA, P.A.; TRIPP, R.A. Respiratory syncytial virus: prospects for new and emerging therapeutics. **Expert Review of Respiratory Medicine** v. 11, n.8, pp. 609-615, 2017. DOI: [10.1080/17476348.2017.1338567](https://doi.org/10.1080/17476348.2017.1338567) .
- 164- BISGAARD, H. A, STUDY GROUP ON MONTELUKAST AND RESPIRATORY SYNCYTIAL VIRUS. A randomized trial of montelukast in respiratory syncytial virus postbronchiolitis. **American Journal of Respiratory and Critical Care Medicine** v. 167, n. 3, pp. 379-383, 2003. DOI: [10.1164/rccm.200207-747OC](https://doi.org/10.1164/rccm.200207-747OC) .
- 165- PETERSEN, B.C. *et al.* IL-17E (IL-25) and IL-17RB promote respiratory syncytial virus-induced pulmonary disease. **Journal of Leukocyte Biology** v. 95, n. 5, pp. 809-815, 2014. DOI: [10.1189/jlb.0913482](https://doi.org/10.1189/jlb.0913482) .
- 166- ROSSEY, I. *et al.* Potent single-domain antibodies that arrest respiratory syncytial virus fusion protein in its prefusion state. **Nature Communication** v. 8, 2016. DOI: [10.1038/ncomms14158](https://doi.org/10.1038/ncomms14158) .
- 167- SCHEPENS, B. *et al.* Nanobodies® specific for respiratory syncytial virus fusion protein protect against infection by inhibition of fusion. **The Journal of Infectious Diseases** v. 204, n. 11, pp. 1692-1701, 2011. DOI: [10.1093/infdis/jir622](https://doi.org/10.1093/infdis/jir622) .
- 168- CHUNG, D.H. *et al.* A cell based high-throughput screening approach for the discovery of new inhibitors of respiratory syncytial virus. **Virology Journal** v. 10, n. 19, 2013. DOI: [10.1186/1743-433x-10-19](https://doi.org/10.1186/1743-433x-10-19) .
- 169- LEYSSEN, P.; DE CLERCQ, E., NEYTS, J. Molecular strategies to inhibit the replication of RNA viruses. **Antiviral Research** v. 78, n. 1, pp. 9-25, 2008. DOI: [10.1016/j.antiviral.2008.01.004](https://doi.org/10.1016/j.antiviral.2008.01.004) .
- 170- SHIGETA, S. Recent progress in antiviral chemotherapy for respiratory syncytial virus infections. **Expert Opinion on Investigational Drugs** v. 9, n. 2, pp. 221-235, 2000. DOI: [10.1517/13543784.9.2.221](https://doi.org/10.1517/13543784.9.2.221) .
- 171- BONFANTI, J.F. *et al.* Selection of a respiratory syncytial virus fusion inhibitor clinical candidate. 2. Discovery of a Morpholinopropylaminobenzimidazole derivative

- (TMC353121). **Journal of Medicinal Chemistry** v. 51, n. 4, pp. 875-896, 2008. DOI: [10.1021/jm701284j](https://doi.org/10.1021/jm701284j) .
- 172- SIDWELL, R.W.; BARNARD, D.L. Respiratory syncytial virus infections: recent prospects for control. **Antiviral Research** v. 71: 379-390, 2006. DOI: [10.1016/j.antiviral.2006.05.014](https://doi.org/10.1016/j.antiviral.2006.05.014) .
- 173- SUDO, K. *et al.* YM-53403, a unique anti-respiratory syncytial virus agent with a novel mechanism of action. **Antiviral Research** v. 65, n. 2, pp. 125-131, 2005. DOI: [10.1016/j.antiviral.2004.12.002](https://doi.org/10.1016/j.antiviral.2004.12.002) .
- 174- DEVICENZO, J. *et al.* Evaluation of the safety, tolerability and pharmacokinetics of ALN-RSV01, a novel RNAi antiviral therapeutic directed against respiratory syncytial virus (RSV). **Antiviral Research** v. 77, n. 3, pp. 225-231, 2008. DOI: [10.1016/j.antiviral.2007.11.009](https://doi.org/10.1016/j.antiviral.2007.11.009) .
- 175- BATTLES, M.B. *et al.* Molecular mechanism of respiratory syncytial virus fusion inhibitors. **Nature Chemical Biology** v. 12, n. 2, pp. 87-93, 2016. DOI: [10.1038/nchembio.1982](https://doi.org/10.1038/nchembio.1982) .
- 176- BONAVIDA, A. *et al.* Identification of broad-spectrum antiviral compounds and assessment of the druggability of their target for efficacy against respiratory syncytial virus (RSV). **Proceedings of the National Academy of Sciences of the United States of America** v. 108, n. 17, pp. 6739-6744, 2011. DOI: [10.1073/pnas.1017142108](https://doi.org/10.1073/pnas.1017142108) .
- 177- RASMUSSEN, L. *et al.* A high-throughput screening strategy to overcome virus instability. **Assay and Drug Development Technologies** v. 9, n. 2, pp. 184-190, 2011. DOI: [10.1089/adt.2010.0298](https://doi.org/10.1089/adt.2010.0298) .
- 178- GOBEL, J. *et al.* A phenotypic high-throughput screen with RSV-infected primary human small airway epithelial cells (SAECs). **Journal of Biomolecular Screening** v. 20, n. 6, pp. 729-738, 2015. DOI: [10.1177/1087057115580271](https://doi.org/10.1177/1087057115580271) .
- 179- PLANT, H. *et al.* High-throughput hit screening cascade to identify respiratory syncytial virus inhibitors. **Journal of Biomolecular Screening** v. 20, n. 5, pp. 597-605, 2015. DOI: [10.1177/1087057115569428](https://doi.org/10.1177/1087057115569428) .
- 180- NEWMAN, D.J., CRAGG, G.M. Natural products as sources of new drugs from 1981 to 2014. **Natural Products** v. 79, pp. 629-661, 2016. DOI: [10.1021/acs.jnatprod.5b01055](https://doi.org/10.1021/acs.jnatprod.5b01055) .
- 181- HARVEY, A.L.; EDRADA-EBEL, R.; QUINN, R.J. The re-emergence of natural products for drug discovery in the genomics era. **Nature Reviews Drug Discovery** v. 14, pp. 111-129, 2015. DOI: [10.1038/nrd4510](https://doi.org/10.1038/nrd4510) .

- 182- MARTINEZ, J.P. *et al.*, Antiviral drug discovery: broad-spectrum drugs from nature. **Natural Products Reports** v. 1, 2015. DOI: [10.1039/c4np00085d](https://doi.org/10.1039/c4np00085d) .
- 183- BUTLER, M.S.; ROBERTSON, A.A.B.; COOPER, M.A. Natural product and natural product derived drugs in clinical trials. **Natural Products Reports** v. 31, n. 11, pp. 1612-1661, 2014. DOI: [10.1039/c4np00064a](https://doi.org/10.1039/c4np00064a) .
- 184- ABOELSOU, N.H. Herbal medicine in Egypt. **Journal of Medicinal Plants Research** v. 2, n. 2, pp. [82-86](#), 2010.
- 185- SIMÕES, C. M. O.; *et al.* **Farmacognosia: da Planta ao medicamento**. Porto Alegre/Florianópolis Ed.Universidade UFRGS, 1999.
- 186- BAEEL, J.B.; HOLLOWAY, G.A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. **Journal of Medicinal Chemistry** v. 53, pp. 2719-2740, 2010. DOI: [10.1021/jm901137j](https://doi.org/10.1021/jm901137j) .
- 187- BISSON, J. *et al.* Can invalid bioactives undermine natural product-based drug discovery? **Journal of Medicinal Chemistry** v. 59, pp. 1671-1690, 2015. DOI: [10.1021/acs.jmedchem.5b01009](https://doi.org/10.1021/acs.jmedchem.5b01009) .
- 188- McGOVERN, S.L. *et al.* A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. **Journal of Medicinal Chemistry** v. 45, pp. 1712-1722, 2002. DOI: [10.1021/jm010533y](https://doi.org/10.1021/jm010533y) .
- 189- GLASER, J.; HOLZGRABE, U. Focus on PAINS: false friends in the quest for selective anti-protozoal lead structures from Nature? **Medicinal Chemistry Communication** v. 7, pp. 214-223, 2016. DOI: [10.1039/C5MD00481K](https://doi.org/10.1039/C5MD00481K) .
- 190- SENGER, M.R. *et al.* Filtering promiscuous compounds in early drug discovery: is it a good idea? **Drug Discovery Today** v. 21, n. 6, pp. 868-872, 2016. DOI: [10.1016/j.drudis.2016.02.004](https://doi.org/10.1016/j.drudis.2016.02.004) .
- 191- CHAI, C.L.L.; MÁTYUS, P. One size does not fit all: challenging some dogmas and taboos in drug discovery. **Future Medicinal Chemistry** v. 8, n. 1, pp. 29-38, 2016. DOI: [10.4155/fmc.15.167](https://doi.org/10.4155/fmc.15.167) .
- 192- MEDINA-FRANCO, J.L. *et al.* Shifting from the single to the multitarget paradigm in drug discovery. **Drug Discovery Today** v. 18, n. 9/10, pp. 495-501. DOI: [10.1016/j.drudis.2013.01.008](https://doi.org/10.1016/j.drudis.2013.01.008) .

- 193- MERINO, A. *et al.* Drug profiling: knowing where it hits. **Drug Discovery Today** v. 15, n. 17/18, pp. 749-756, 2010. DOI: [10.1016/j.drudis.2010.06.006](https://doi.org/10.1016/j.drudis.2010.06.006) .
- 194- AHMAD, A. *et al.* Therapeutic potential of flavonoids and their mechanism of action against viral infections – A review. **Food Research International** v. 77, pp. 221-235, 2015. DOI: [10.1016/j.foodres.2015.06.021](https://doi.org/10.1016/j.foodres.2015.06.021) .
- 195- SCOTTI, L. *et al.* Docking studies for multi-target drugs. **Current Drug Targets** v. 17, 2015. DOI: [10.2174/1389450116666150825111818](https://doi.org/10.2174/1389450116666150825111818) .
- 196- ZHANG, W.; PEI, J.; LAI, L. Computational multitarget drug design. **Journal of Chemical Information and Modeling** v. 57, pp. 403-412, 2017. DOI: [10.1021/acs.jcim.6b00491](https://doi.org/10.1021/acs.jcim.6b00491) .
- 197- UTKARSH, R.; VARADWAJ, P.K. Flavonoids as multi-target inhibitors for proteins associated with Ebola virus: *In silico* discovery using virtual screening and molecular docking studies. **Interdisciplinary Sciences, Computational Life Sciences** v. 8, pp. 132-141, 2015. DOI: [10.1007/s12539-014-0246-5](https://doi.org/10.1007/s12539-014-0246-5) .
- 198- FANG, Y. *et al.* 3D-QSAR and docking studies of flavonoids as potent Escherichia coli inhibitors. **Scientific reports** v. 6, n. 23634, 2016. DOI: [10.1038/srep23634](https://doi.org/10.1038/srep23634) .
- 199- OLIVERO-VERBEL, J.; PACHECO-LONDOÑO, L. Structure-activity relationships for the anti-HIV activity of flavonoids. **Journal of Chemical Information and Modeling** v. 42, n. 5, pp. 1241-1246, 2002. DOI: [10.1021/ci020363d](https://doi.org/10.1021/ci020363d) .
- 200- LIU, A.L. *et al.* Structure-activity relationship of flavonoids as influenza virus neuraminidase inhibitors and their in vitro anti-viral activities. **Bioorganic & Medicinal Chemistry** v. 16, pp. 7141-7147, 2008. DOI: [10.1016/j.bmc.2008.06.049](https://doi.org/10.1016/j.bmc.2008.06.049) .
- 201- CUSHNIE, T.P.T.; LAMB, A.J. Antimicrobial activity of flavonoids. **International Journal of Antimicrobial Agents** v. 26, pp. 343-356, 2005. DOI: [10.1016/j.ijantimicag.2005.09.002](https://doi.org/10.1016/j.ijantimicag.2005.09.002) .
- 202- SHANMUGAM, A.; GROMIHA, M.M. Quercetin derivatives as non-nucleoside inhibitors for dengue polymerase: molecular docking, molecular dynamics and binding-free energy calculation. **Journal of Biomolecular Structure and Dynamics** v. 35, n.13, 2016. DOI: [10.1080/07391102.2016.1234416](https://doi.org/10.1080/07391102.2016.1234416) .
- 203- URIARTE-PUEYO, I.; CALVO, M.I. Structure-activity relationships of acetylated flavone glycosides from *Galeopsis ladanum* L. (Lamiaceae). **Food Chemistry** v. 120, pp. 679-683, 2010. DOI: [10.1016/j.food.chem.2009.10.060](https://doi.org/10.1016/j.food.chem.2009.10.060) .

- 204- TEIXEIRA, T.S. *et al.* Biophysical characterization of the interaction between M2-1 protein of hRSV and quercetin. **International Journal of Biological Macromolecules** v. 95, pp. 63-71, 2017. DOI: [10.1016/j.ijbiomac.2016.11.033](https://doi.org/10.1016/j.ijbiomac.2016.11.033) .
- 205- GUIMARÃES, G.C. *et al.* Binding investigation between M2-1 protein from hRSV and acetylated quercetin derivatives: ¹H NMR, fluorescence, spectroscopy and molecular docking. **International Journal of Biological Macromolecules** v. 111, pp. 33-38, 2017. DOI: [10.1080/07391102.2016.1234416](https://doi.org/10.1080/07391102.2016.1234416) .
- 206- SOUZA, F.P. *et al.* Interaction model between hRSV G-protein and flavonoids. **International Journal of Sciences** v. 2, n.10, pp. [12-19](#), 2013.
- 207- GOMES, D.E. *et al.* Experimental evidence and molecular modeling of the interaction between hRSV-NS1 and quercetin. **International Journal of Biological Macromolecules** v. 85, pp. 40-47, 2016. DOI: [10.1080/07391102.2016.1234416](https://doi.org/10.1080/07391102.2016.1234416) .
- 208- DANIHELOVÁ, M. *et al.* Antioxidant action and cytotoxicity on Hela and NIH-3T3 cells of new quercetin derivatives. **Interdisciplinary Toxicology** v. 6, n.4, pp. 209-216, 2013. DOI: [10.2478/intox-2013-0031](https://doi.org/10.2478/intox-2013-0031) .
- 209- DANIHELOVÁ, M.; VISKUPICOVÁ, J.; STURDÍK, E. Lipophilization of flavonoids for their food, therapeutic and cosmetic applications. **Acta Chimica Slovaca** v.5, n. 1, pp. 59-69, 2012. DOI: [10.2478/v10188-012-0010-6](https://doi.org/10.2478/v10188-012-0010-6) .
- 210- GUSDINAR, T. *et al.* Anti-inflammatory and antioxidant activity of quercetin-3',3',4'-Triacetate. **Journal of Pharmacology and Toxicology** v. 6, n.2, pp. 182-188, 2011. DOI: [10.3923/jpt.2011.182.188](https://doi.org/10.3923/jpt.2011.182.188) .
- 211- PARK, J. *et al.* Effects of resveratrol, oxyresveratrol, and their acetylated derivatives on cellular melanogenesis. **Archives of Dermatological Research** v. 306, n.5, pp. 475-487, 2014. DOI: [10.1007/s00403-014-1440-3](https://doi.org/10.1007/s00403-014-1440-3) .
- 212- SAKAO, K.; FUJII, M.; HOU, D.X. Acetyl derivative of quercetin increases the sensitivity of human leukemia cells toward apoptosis. **Biofactors** v. 35, n.4, pp. 399-405, 2009. DOI: [10.1002/biof.53](https://doi.org/10.1002/biof.53) .
- 213- DAYEM, A.A. *et al.* Antiviral effect of methylated flavonol Isorhamnetin against influenza. **Plos One** v. 10, n.3, e0121610, 2015. DOI: [10.1371/journal.pone.0121611](https://doi.org/10.1371/journal.pone.0121611) .
- 214- MURA, C.; McANANY, C.E. An introduction to biomolecular simulations and docking. **Molecular Simulation** v. 40, n. 10-11, 2014. DOI: [10.1080/08927022.2014.935372](https://doi.org/10.1080/08927022.2014.935372) .

- 215- FERREIRA, L.G. *et al.* Molecular docking and structure-based drug design strategies. **Molecules** v. 20, pp. 13384-13421, 2015. DOI: [10.3390/molecules200713384](https://doi.org/10.3390/molecules200713384) .
- 216- SONAVANE, Y. *et al.* Molecular dynamics simulation of functionalized nanofibrillar cellulose. Conference: **International Parallel Conferences on Researches in Industrial and Applied Sciences**, At Dubai, UAE, Volume: Panel II: Life and Biological Studies, Page: 1-9, 2014.
- 217- HERNÁNDEZ-SANTOYO, A. *et al.* **Protein-Protein and Protein-Ligand Docking, Protein Engineering - Technology and Application**, Dr. Tomohisa Ogawa (Ed.), InTech, 2013 DOI: 10.5772/56376. Disponível em: <https://www.intechopen.com/books/howtoreference/protein-engineering-technology-and-application/protein-protein-and-protein-ligand-docking>. Acesso em 15 fevereiro 2018.
- 218- **MODELING ANAEROBIC PROCESS FOR WASTEWATER TREATMENT: NEW TRENDS AND METHODOLOGIES**. Disponível em: https://pdfs.semanticscholar.org/5ce7/ced92f890fd0a7028dd61e18c73b9d63744b.pdf?_ga=2.123199945.330265711.1542743805-346756214.1542743805 . Acesso em 16 novembro 2018.
- 219- **DIAGRAM OF AN ARTIFICIAL NEURAL NETWORK**. Disponível em: <https://tex.stackexchange.com/questions/132444/diagram-of-an-artificial-neural-network> . Acesso em 17 março 2018.
- 220- YUNTA, M.J.R. Docking and ligand binding affinity: uses and pitfalls. **American Journal of Modeling and Optimization** v. 4, n.3, pp. 74-114, 2016. DOI: [10.12691/ajmo-4-3-2](https://doi.org/10.12691/ajmo-4-3-2) .
- 221- GANESAN, A.; COOTE, M.L.; BARAKAT, K. Molecular dynamics-driven drug discovery: leaping forward with confidence. **Drug Discovery Today** v. 22, n.2, pp. 249-269, 2017. DOI: [10.1016/j.drudis.2016.11.001](https://doi.org/10.1016/j.drudis.2016.11.001) .
- 222- PERILLA, J.R. *et al.* Molecular dynamics simulations of large macromolecular complexes. **Current Opinion in Structural Biology** v. 31, pp. 64-74, 2015. DOI: [10.1016/j.sbi.2015.03.007](https://doi.org/10.1016/j.sbi.2015.03.007) .
- 223- MENG, X.Y. *et al.* Molecular Docking: a powerful approach for structure-based drug discovery. **Current Computer-aided Drug Design** v. 7, n.2, pp. [146-157](#), 2011.
- 224- JONES, G. *et al.* Development and validation of a genetic algorithm for flexible docking. **Journal of Molecular Biology** v. 267, n.3, pp. 727-748, 1997. DOI: [10.1006/jmbi.1996.0897](https://doi.org/10.1006/jmbi.1996.0897) .

- 225- FRIESNER, R.A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. **Journal of Medicinal Chemistry** v. 47, n. 7, pp. 1739-1749, 2004. DOI: [10.1021/jm0306430](https://doi.org/10.1021/jm0306430) .
- 226- MORRIS, G.M. *et al.* Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. **Journal of Computer-aided Molecular Design** v. 10, n.4, pp. [293-304](#), 1996.
- 227- BIOCHEMISTRY ONLINE. Disponível em: https://employees.csbsju.edu/hjakubowski/classes/ch331/protstructure/mechdynam2_041416.html . Acesso em 30 março 2018.
- 228- RIPPHAUSEN, P.; NISIUS, B.; BAJORATH, J. State-of-the-art in ligand-based virtual screening. **Drug Discovery Today** v. 16, n.9-10, pp. 372-376, 2011. DOI: [10.1016/j.drudis.2011.02.011](https://doi.org/10.1016/j.drudis.2011.02.011) .
- 229- GEPPERT, H.; VOGT, M.; BAJORATH, J. Current trend in ligand-based virtual screening: molecular representations, data mining methods, new application areas, and performance evaluation. **Journal of Chemical Information and Modeling** v. 50, n.2, pp. 205-216, 2010. DOI: [10.1021/ci900419k](https://doi.org/10.1021/ci900419k) .
- 230- MANDLIK, V.; BEJUGAM, P.R.; SINGH, S. Application of artificial neural networks in modern drug discovery. **Artificial Neural Network for Drug Design, Delivery and Disposition** pp. 123-139, 2016. DOI: [10.1016/B978-0-12-801559-9.00006-5](https://doi.org/10.1016/B978-0-12-801559-9.00006-5) .
- 231- BASKIN, I.I.; WINKLER, D.; TETKO, I.V. A renaissance of neural networks in drug discovery. **Expert Opinion On Drug Discovery** v. 11, n.8, pp. 785-795, 2016. DOI: [10.1080/17460441.2016.1201262](https://doi.org/10.1080/17460441.2016.1201262) .
- 232- YANG, J. *et al.* The I-TASSER Suite: Protein structure and function prediction. **Nature Methods** v. 12, pp. 7-8, 2015. DOI: [10.1038/nmeth.3213](https://doi.org/10.1038/nmeth.3213) .
- 233- CASE, D.A. *et al.* **AMBER 14**, University of California, San Francisco, 2014.
- 234- MAIER, J.M.A. *et al.* ff14SB: Improving the accuracy of protein side chain and Backbone parameters from 99SB. **Journal of Chemical Theory and Computation** v. 11, n.8, pp. 3896-3713, 2015. DOI: [10.1021/acs.jctc.5b00255](https://doi.org/10.1021/acs.jctc.5b00255) .
- 235- JORGENSEN, W.L. *et al.* Comparison of simple potential functions for simulating liquid water. **Journal of Chemical Physics** v. 79, pp. 926-935, 1983. DOI: [10.1063/1.445869](https://doi.org/10.1063/1.445869) .
- 236- IZAGUIRRE, J.A. Langevin stabilization of molecular dynamics. **Journal of Chemical Physics** v. 114, pp. 2090-2098, 2001. DOI: [10.1063/1.1332996](https://doi.org/10.1063/1.1332996) .

- 237- VAN DER SPOEL, D.; E. LINDAHL, B. HESS, and the GROMACS development team. **GROMACS 4.6.7**, Royal Institute of Technology and Uppsala University, Sweden, 2014.
- 238- WEISER, J.; SHENKIN, P.; STILL, W. Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). **Journal of Computational Chemistry** v. 20, pp. 217-230, 1999. DOI: [10.1002/\(SICI\)1096-987X\(19990130\)20:2<217::AID-JCC4>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1096-987X(19990130)20:2<217::AID-JCC4>3.0.CO;2-A) .
- 239- HUMPHREY, W., DALKE, A., SCHULTEN, K. VMD: visual molecular dynamics. **Journal of Molecular Graphics** v. 14, n.1, pp. 33-38, 27-28, 1996. DOI: [10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) .
- 240- TROTT, O.; OLSON, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. **Journal of Computational Chemistry** v. 31, pp. 455-461, 2010. DOI: [10.1002/JCC.21334](https://doi.org/10.1002/JCC.21334) .
- 241- BOYLE, N.M.O. *et al.* Open Babel: An open chemical toolbox. **Journal of Cheminformatics** v. 3, n.33, 2011. DOI: [10.1186/1758-2946-3-33](https://doi.org/10.1186/1758-2946-3-33) .
- 242- AMADEI, A.; LINSEN, A.B.; BERENDSEN, H.J. Essential dynamics of proteins. **Proteins** v. 17, n.4, pp. 412-425, 1993. DOI: [10.1002/prot.340170408](https://doi.org/10.1002/prot.340170408) .
- 243- MU, Y.; NGUYEN, P.H.; STOCK, G. Energy landscape of a small peptide revealed by dihedral angle principal component analysis. **Proteins** v. 58, n.1, pp. 45-52, 2005. DOI: [10.1002/prot.20310](https://doi.org/10.1002/prot.20310) .
- 244- KABSCH, W.; SANDER, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. **Biopolymers** v. 22, n.12, pp. 2577-2637, 1983. DOI: [10.1002/bip.360221211](https://doi.org/10.1002/bip.360221211) .
- 245- PRATHAMESH, M. *et al.* Applications of Principal Component Analysis (PCA) in Materials Science, Principal Component Analysis - Engineering Applications, Dr. Parinya Sanguansat (Ed.), **InTech**, 2012. DOI: 10.5772/37523. Disponível em: <https://mts.intechopen.com/books/principal-component-analysis-engineering-applications/applications-of-principal-component-analysis-pca-in-materials-science> . Acesso em 24 março 2018.
- 246- SHLENS, J. **A tutorial on principle component analysis**. Disponível em: <https://arxiv.org/abs/1404.1100> . Acesso em 24 março 2018.

- 247- McLELLAN, J. *et al.* Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. **Science** v. 340, n. 6136, pp. 113-117, 2013. DOI: [10.1126/science.1234914](https://doi.org/10.1126/science.1234914) .
- 248- THOMAS, F. *et al.* Conformational dynamics of asparagine at coiled-coil interfaces. **Biochemistry** v. 56, n.50, pp. 6544-6554, 2017. DOI: [10.1021/acs.biochem.7b00848](https://doi.org/10.1021/acs.biochem.7b00848) .
- 249- PIÑEIRO, A. *et al.* A molecular dynamics study of the formation, stability, and oligomerization state of two designed coiled coils: possibilities and limitations. **Biophysical Journal** v. 89, n. 6, pp. 3701-3713, 2005. DOI: [10.1529/biophysj.104.055590](https://doi.org/10.1529/biophysj.104.055590) .
- 250- MISSIMER, J.H. *et al.* Molecular-dynamics simulations of C- and N-terminal peptide derivatives of GCN4-p1 in aqueous solution. **Chemistry & Biodiversity** v. 2, n.8, pp. 1086-1104, 2005. DOI: [10.1002/cbdv.200590078](https://doi.org/10.1002/cbdv.200590078) .
- 251- TRUEBESTEIN, L.; LEONARD, T.A. Coiled-coils: the long and short of it. **Bioessays** v. 38, n.9, pp. 903-916, 2016. DOI: [10.1002/bies.201600062](https://doi.org/10.1002/bies.201600062) .
- 252- MARTÍN-GARCIA, F. *et al.* Molecular dynamics analysis of conformational change of paramyxovirus F protein during the initial steps of membrane fusion. **Biochemical and biophysical research communications** v. 420, n.1, pp. 42-47, 2012. DOI: [10.1016/j.bbrc.2012.02.112](https://doi.org/10.1016/j.bbrc.2012.02.112) .
- 253- HADLEY, K.R.; McCABE, C. Coarse-grained molecular models of water: a review. **Molecular Simulation** v. 38, n.8-9, pp. 671-681, 2012. DOI: [10.1080/08927022.2012.671942](https://doi.org/10.1080/08927022.2012.671942) .
- 254- BRADLEY, R.; RADHAKRISHNAN, R. Coarse-grained models for protein-cell membrane interactions. **Polymers (Basel)** v. 5, n.3, pp. 890-936, 2013. DOI: [10.3390/polym5030890](https://doi.org/10.3390/polym5030890) .
- 255- WALSH, E.E., HRUSKA, J. Monoclonal antibodies to respiratory syncytial virus proteins: identification of the fusion protein. **Journal of virology** v. 47, pp. [171-177](#), 1983.
- 256- BATTLES, M.B. *et al.* Molecular mechanism of respiratory syncytial virus fusion inhibitors. **Nature Chemical Biology** v. 16, pp. 87-93, 2016. DOI: [10.1038/nchembio.1982](https://doi.org/10.1038/nchembio.1982) .
- 257- MORTON, C.J. *et al.* Structure characterization of respiratory syncytial virus fusion inhibitor escape mutants: homology model of the F protein and a syncytium formation assay. **Virology** v. 11, pp. [275-288](#), 2003.

- 258- DOUGLAS, J.L., *et al.* Small molecules VP-14637 AND JNJ-2408068 inhibit respiratory syncytial virus fusion by similar mechanisms. **Antimicrobial agents and chemotherapy** v. 49, n.6, pp. 2460-2466, 2005. DOI: [10.1128/AAC.49.6.2460-2466.2005](https://doi.org/10.1128/AAC.49.6.2460-2466.2005) .
- 259- LÓPEZ, J.A. *et al.* Antigenic structure of human respiratory syncytial virus fusion glycoprotein. **Journal of Virology** v. 72, n.8, pp. [6922-6928](https://doi.org/10.1128/JVI.72.8.6922-6928.1998), 1998.
- 260- ZIMMER, G.; TROTZ, I.; HERRLER, G. N-glycans of F protein differentially affect fusion activity of human respiratory syncytial virus. **Journal of Virology** v. 75, n.10, pp. 4744-4751, 2001. DOI: [10.1128/JVI.75.10.4744-4751.2001](https://doi.org/10.1128/JVI.75.10.4744-4751.2001).
- 261- DAY, N.D. *et al.* Contribution of cysteine residues in the extracellular domain of the F protein of human respiratory syncytial virus to its function. **Virology Journal** v. 34, n.3, 2006. DOI: [10.1186/1743-422X-3-34](https://doi.org/10.1186/1743-422X-3-34) .
- 262- SUN, Z. *et al.* Respiratory syncytial virus inhibitors targeting the F protein. **Viruses** v.5, pp. 211-225, 2013. DOI: [10.3390/v5010211](https://doi.org/10.3390/v5010211) .
- 263- COLIZZI, F. *et al.* Single-molecule pulling simulations can discern active from inactive enzyme inhibitors. **Journal of the American Chemical Society** v. 132, n.21, pp. 7361-7366, 2010. DOI: [10.1021/ja101021a001](https://doi.org/10.1021/ja101021a001) .
- 264- ZHANG, Y.; SKOLNICK, J. Scoring function for automated assessment of protein structure template quality. **Proteins: Structure, Function and Bioinformatics** v. 57, n. 4, pp. 702-710, 2004. DOI: [10.1002/prot.20264](https://doi.org/10.1002/prot.20264) .
- 265- _____ TM-align: a protein structure alignment algorithm based on the TM-score. **Nucleic Acid Research** v. 33, n.7, pp. 2302-2309, 2005. DOI: [10.1093/nar/gki524](https://doi.org/10.1093/nar/gki524) .
- 266- SANVER, D. *et al.*, Experimental modeling of flavonoid-biomembrane interactions. **Langmuir** v. 32, n. 49, pp. 13234-13243, 2016. DOI: [10.1021/acs.langmuir.6b02219](https://doi.org/10.1021/acs.langmuir.6b02219) .
- 267- TARAHOVSKY, Y.S. *et al.*, Flavonoid-membrane interactions: involvement of flavonoid-metal complexes in raft signaling. **Biochimica et biophysica acta** v. 1838, n.5, pp. 1235-1246, 2014. DOI: [10.1016/j.bbamem.2014.01.021](https://doi.org/10.1016/j.bbamem.2014.01.021) .
- 268- KOPÉC, W.; TELENÍUS, J.; KHANDELIA, H. Molecular dynamics simulations of the interactions of medicinal plant extracts and drugs with lipid bilayer membranes. **FEBS Journal** v. 280, n.12, pp. 2785-2805, 2013. DOI: [10.1111/febs.12286](https://doi.org/10.1111/febs.12286) .
- 269- CHADAL, B. *et al.* Structure-dependent membrane interaction and bioactivity of flavonoids with lipid bilayers. International Symposium on micro-nanomechanics and human Science (MHS), 2012. DOI:

[10.1109/MHS.2012.6492463](https://doi.org/10.1109/MHS.2012.6492463) .

- 270- HENDRICH, A.B. Flavonoid-membrane interactions: possible consequences for biological effects of some polyphenolic compounds. **Acta Pharmacologica Sinica** v. 27, pp. 27-40, [2006](#).
- 271- OTEIZA, P.I. *et al.* Flavonoid-membrane interactions: a protective role of flavonoids at the membrane surface? **Clinical and Developmental Immunology** v. 12, n.1, pp. 19-25, [2005](#).
- 272- MARDIGAN, M.T. *et al.* **Brock Biology of Microorganisms**. 13th edition, 2011. ISBN 10: 0-321-64963-X.
- 273- **THE CENTRAL DOGMA OF MOLECULAR BIOLOGY**. Disponível em: <http://ib.bioninja.com.au/standard-level/topic-2-molecular-biology/27-dna-replication-transcri/central-dogma.html> . Acesso em 15 novembro 2018.
- 274- **PROTEIN STRUCTURE LEVELS**. Disponível em: https://de.wikipedia.org/wiki/Datei:Main_protein_structure_levels_en.svg . Acesso em 15 novembro 2018.
- 275- WU, S.; ZHANG, Y. **Protein structure prediction**. Disponível em: https://zhanglab.ccmb.med.umich.edu/papers/2009_8.pdf . Acesso em 21 de novembro 2018.
- 276- DORN, M. Three-dimensional protein structure prediction: methods and computational strategies. **Computational Biology and Chemistry** v. 53, Part B, pp. 251-276, 2014. DOI: [10.1016/j.compbiolchem.2014.10.001](https://doi.org/10.1016/j.compbiolchem.2014.10.001) .
- 277- BASHEER, I. A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. **Journal of Microbiological Methods** v. 43, n. 1, pp.3–31, [2000](#).
- 278- BENZER, R.; BENZER, S. Application of artificial neural network into the freshwater fish caught in Turkey. **International Journal of Fisheries and Aquatic Studies** v. 2,n.5, pp. 341–346, [2015](#).
- 279- CHAKRABARTY, A., DE, A.; DUBEY, R. An excursion towards multi-frequency-prediction based flexible asset allocation system. **Global Journal of Flexible Systems Management** v. 16, n.2, pp. 173–190, [2015](#).
- 280- GEBLER, D. *et al.* Artificial neural network modelling of macrophyte indices based on physic-chemical characteristics of water. **Hydrobiologia** v. 737, n.1, pp. 215–224, [2014](#).

- 281- SEGURITAN, V. *et al.* Artificial neural networks trained to detect viral and phage structural proteins. **PLOS Computational Biology** 8, e1002657, 2012. DOI: [10.1371/journal.pcbi.1002657](https://doi.org/10.1371/journal.pcbi.1002657) .
- 282- HAYKIN, S. **Redes Neurais: princípios e prática**. Bookman, 2001.
- 283- KROGH, A. What are artificial neural networks? **Nature Biotechnology** v. 26, n.2, pp. 195–197, 2008. DOI: [10.1038/nbt1386](https://doi.org/10.1038/nbt1386) .
- 284- WALCZAK, S.; CERPA, N. Heuristic principles for the design of artificial neural networks. **Information and Software Technology** v. 41, pp. 107–117, [1999](https://doi.org/10.1016/0950-5849(99)00050-0).
- 285- TANOMARU, J. **Motivação, fundamentos e aplicações de Algoritmos Genéticos**. Curitiba: II Congresso Brasileiro de Redes Neurais, 1995.
- 286- BEALE, M. H.; HAGAN, M. T.; DEMUTH, H. B. **Neural Network Toolbox™ User's Guide 8**, 2015.
- 287- GOETHALS, P. L. M. *et al.* Applications of artificial neural networks predicting macroinvertebrates in freshwaters. **Aquatic Ecology** v. 41, n.3, pp. 491–508, 2007. DOI: [10.1007/s10452-007-9093-3](https://doi.org/10.1007/s10452-007-9093-3) .
- 288- LUDERMIR, T. B.; YAMAZAKI, A.; ZANCHETTIN, C. An optimization methodology for neural networks weights and architectures. **IEEE Transactions on Neural Networks** v. 17, pp.1452–1459, [2006](https://doi.org/10.1109/TNN.2006.2582110).
- 289- KOVÁCS, Z. L. **Redes Neurais Artificiais: fundamentos e aplicações: um texto básico**. Editora Livraria da Física, 2006.
- 290- KHOSRAVI, A. *et al.* Comprehensive review of neural network-based prediction intervals and new advances. **IEEE Transactions on Neural Networks** v. 22, n.9, pp. 1341–1356, 2011. DOI: [10.1109/TNN.2011.2162110](https://doi.org/10.1109/TNN.2011.2162110) .
- 291- LASKOWSKI, R.A. *et al.* PROCHECK - a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography** v. 26, pp. 283-291, 1993. DOI: [10.1107/S0021889892009944](https://doi.org/10.1107/S0021889892009944) .
- 292- BENKERT, P.; TOSATTO, S.C.E.; SCHOMBURG, D. QMEAN: A comprehensive scoring function for model quality assessment. **Proteins: Structure, Function, and Bioinformatics** v. 71, n.1, pp. 261-277, 2008. DOI: [10.1002/prot.21715](https://doi.org/10.1002/prot.21715) .

- 293- ROY, A.; KUCUKURAL, A.; ZHANG, Y. I-TASSER: a unified platform for automated protein structure and function prediction. **Nature Protocols** v. 5, n.4, pp. 725-738, 2010. DOI: [10.1038/nprot.2010.5](https://doi.org/10.1038/nprot.2010.5) .
- 294- ZHANG, Y. I-TASSER server for protein 3D structure prediction. **BMC Bioinformatics** v. 9, n. 40, 2008. DOI: [10.1186/1471-2105-9-40](https://doi.org/10.1186/1471-2105-9-40) .
- 295- ZHANG, Y. Template-based modeling and free modeling by I-TASSER in CASP7. **Proteins** v. 69, n.8, pp. 108-117, 2007. DOI: [10.1002/prot.21702](https://doi.org/10.1002/prot.21702) .
- 296- ZHANG, Y.; SKOLNICK, J. Automated structure prediction of weakly homologous proteins on a genomic scale. **Proceedings of the National Academy of Sciences of the United States of America** v. 101, n.20, pp. 7594-7599, 2004. DOI: [10.1073/pnas.0305695101](https://doi.org/10.1073/pnas.0305695101) .
- 297- ZHANG, Y.; SKOLNICK, J. Tertiary structure predictions on a comprehensive benchmark of medium to large size proteins. **Biophysical Journal** v. 87, n.4, pp. 2647-2655, 2004. DOI: [10.1529/biophysj.104.045385](https://doi.org/10.1529/biophysj.104.045385) .
- 298- WHITE, S.H. Biophysical dissection of membrane proteins. **Nature** v. 459, pp. 344-346, 2009. DOI: [10.1038/nature08142](https://doi.org/10.1038/nature08142) .
- 299- ILKOWSKI, B.; SKOLNICK, J.; KOLINSKI, A. Helix-coil and beta sheet-coil transitions in a simplified, yet realistic protein model. **Macromolecular Theory and Simulations** v. 9, n.8, pp. 523-533, [2000](#).
- 300- YAPA, K.; WEAVER, D.L.; KARPLUS, M. B-sheet coil transitions in a simple polypeptide model. **Proteins** v. 12, n.3, pp. 237-265, 1992. DOI: [10.1002/prot.340120304](https://doi.org/10.1002/prot.340120304) .
- 301- **SECONDARY STRUCTURE ASSIGNMENT: DSSP AND STRIDE**. Disponível em: <https://slideplayer.com/slide/4047493/> . Acesso em 03 novembro 2018.
- 302- **FISIPRED: PHI PSI PREDICTION**. Disponível em: <http://crdd.osdd.net/raghava/fisipred/index.html> . Acesso em 03 novembro 2018.
- 303- HEYLEN, E.; NEYTS, J.; JOCHMANS, D. Drug candidates and model systems in respiratory syncytial virus antiviral drug discovery. **Biochemical Pharmacology** v. 1, n.127, pp. 1-12, 2017. DOI: [10.1016/j.bcp.2016.09.014](https://doi.org/10.1016/j.bcp.2016.09.014) .
- 304- RASCHKA, S. *et al.* Protein-ligand interfaces are polarized: discovery of a strong trend for intermolecular hydrogen bonds to favor donors on the protein side with implications for

- predicting and designing ligand complexes. **Journal of Computer-aided molecular design** v. 32, n.4, pp. 511-528, 2018. DOI: [10.1007/s10822-018-0105-2](https://doi.org/10.1007/s10822-018-0105-2) .
- 305- LIPINSKI, C.A. *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews** v. 23, n.1, pp. 3-25, 1997. DOI: [10.1016/S0169-409X\(96\)00423-1](https://doi.org/10.1016/S0169-409X(96)00423-1) .
- 306- LIPINSKI, C.A. *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews** v. 46, pp. 3-26, [2001](#).
- 307- VEBER, D.F. *et al.* Molecular properties that influence the oral bioavailability of drug candidates. **Journal of Medicinal Chemistry** v. 45, n.12, pp. 2615-2623, 2002. DOI: [10.1021/jm020017n](https://doi.org/10.1021/jm020017n) .
- 308-KOHAVI, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. **International Joint Conference on Artificial Intelligence(IJCAI)**, [1995](#).
- 309- RIPLEY, B.D. **Pattern recognition and neural networks**. Cambridge University Press, Cambridge, UK, 1996.
- 310- NOORIZADEH, H.; FARMANY, A. Investigation of capacity behaviors by linear and nonlinear models chemometrics. **Trends in Applied Sciences Research** v. 6, pp. 1324-1334, 2011. DOI: [10.3923/tasr.2011.1324.1334](https://doi.org/10.3923/tasr.2011.1324.1334) .
- 311- CERIT, I. *et al.* Estimation of antioxidant activity of foods using artificial neural networks. **Journal of Food and Nutrition Research** v. 56, n.2, pp. 138-148, [2017](#).
- 312- SHEN, J. *et al.* A genetic algorithm-backpropagation artificial neural network model to quantify the affinity of flavonoids toward P-glycoprotein. **Combinatorial Chemistry & High Throughput Screening** v. 17, n.2, pp. 162-172, [2014](#).
- 313- KARAWAJCZYK, A. *et al.* Properties of flavonoids influencing the binding to bilitranslocase investigated by neural network modelling. **Biochemical Pharmacology** v. 73, n.2, pp. 308-320, 2007. DOI: [10.1016/j.bcp.2006.09.024](https://doi.org/10.1016/j.bcp.2006.09.024) .
- 314- ZUVELA, P.; DAVID, J.; WONG, M.W. Interpretation of ANN-based QSAR models for prediction of antioxidant activity of flavonoids. **Journal of Computational Chemistry** v. 39, n.16, pp. 953-963, 2018. DOI: [10.1002/jcc.25168](https://doi.org/10.1002/jcc.25168) .

- 315- WU, Y.; WANG, G. Machine learning based toxicity prediction: from chemical structural description to transcriptome analysis. **International Journal of Molecular Sciences** v. 19, n.8, E2358, 2018. DOI: [10.3390/ijms19082358](https://doi.org/10.3390/ijms19082358) .
- 316- RIFAIOGLU, A.S. *et al.* Recent applications of deep learning and machine intelligence on *in silico* drug discovery: methods, tools and databases. **Briefings in Bioinformatics**, bby061, 2018. DOI: [10.1093/bib/bby061](https://doi.org/10.1093/bib/bby061) .
- 317- SCHNEIDER, G. Automating drug discovery. **Nature Reviews Drug Discovery** v. 17, pp. 97-113, [2018](https://doi.org/10.1038/s41573-018-0000-0).
- 318- ROCHA, J.C. *et al.* A method based on artificial intelligence to fully automatize the evaluation of bovine blastocyst images. **Scientific Reports** v. 7, n. 7659, [2017](https://doi.org/10.1038/s41598-017-07659-0).
- 319- MANDLIK, V.; BEJUGAM, P.R.; SINGH, S. Application of artificial neural networks in modern drug discovery. **Artificial Neural Network for Drug Design, Delivery, and Disposition**, pp. 123-139, 2016. DOI: [10.1016/B978-0-12-801559-9.00006-5](https://doi.org/10.1016/B978-0-12-801559-9.00006-5) .
- 320- MERCADER, A.G.; POMILIO, A.B. 2D- and 3D-QSAR Studies of Flavonoids, Biflavones and Chalcones: Antiviral, Antibacterial, Antifungal, and Antimycobacterial Activities. **Anti-infective Agents** v. 10, n.1, 2012. DOI: [10.2174/2211362611201010041](https://doi.org/10.2174/2211362611201010041) .
- 321- AMIC, D. *et al.* SAR and QSAR of the antioxidant activity of flavonoids. **Current Medicinal Chemistry** v. 14, n.7, pp. 827-845, [2007](https://doi.org/10.2174/093964602568888888).
- 322- LANDAVAZO, D.; FOGEL, G.B. Evolved neural networks for quantitative structure-activity relationships of anti-HIV compounds. **Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No.02TH8600)**, 2002. DOI: [10.1109/CEC.2002.1006233](https://doi.org/10.1109/CEC.2002.1006233) .
- 323- BASKIN, I.I.; WINKLER, D.; TETKO, I.V. A renaissance of neural networks in drug discovery. **Expert Opinion on Drug Discovery** v. 11, n.8, pp. 785-795, 2016. DOI: [10.1080/17460441.2016](https://doi.org/10.1080/17460441.2016) .
- 324- HILLER, S.A. *et al.* Cybernetic methods of drug design. I. Statement of the problem—the perceptron approach. **Computers and Biomedical Research, an International Journal** v. 6, n.5, pp. 411-421, [1973](https://doi.org/10.1016/0010-0285(73)90001-0).
- 325- CHENG, T. *et al.* Computation of octanol-water partition coefficients by guiding an additive model with knowledge. **Journal of Chemical Information and Modeling** v. 47, n.6, pp. 2140-2148, 2007. DOI: [10.1021/ci700257y](https://doi.org/10.1021/ci700257y) .

- 326- ERTL, P.; ROHDE, B.; SELZER, P. Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. **Journal of Medicinal Chemistry** v. 43, n. 20, pp. 3714-3717, 2000. DOI: [10.1021/jm000942e](https://doi.org/10.1021/jm000942e) .
- 327- DECHERCHI, S. *et al.* Implicit solvent methods for free energy estimation. **European Journal of Medicinal Chemistry** v. 91, pp. 27-42, 2015. DOI: [10.1016/j.ejmech.2014.08.064](https://doi.org/10.1016/j.ejmech.2014.08.064) .
- 328- ZHANG, Y.; DING, Y. Molecular dynamics simulation and bioinformatics study on chloroplast stromal ridge complex from rice (*Oryza sativa* L). **BMC Bioinformatics** v. 17, n.28, 2016. DOI: [10.1186/s12859-016-0877-0](https://doi.org/10.1186/s12859-016-0877-0) .
- 329- DAURA, X. *et al.* Peptide folding: when simulation meets experiment. **Angewandte International Edition Chemie** v. 38, n. 1-2, pp. 236-240, 1999. DOI: [10.1002/\(SICI\)1521-3773\(19990115\)38:1/2<236::AID-ANIE236>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1521-3773(19990115)38:1/2<236::AID-ANIE236>3.0.CO;2-M) .
- 330- GONÇALVES, L. *et al.* Roc curve estimation: an overview. **Revstat – Statistical Journal** v. 12, n.1, pp. 1-20, [2014](#).

BIBLIOGRAFIA

- 1- ROSA, C.A.P. **História da Ciência: da antiguidade ao renascimento científico.** Brasília: FUNAG, 2012, 2ª ed. ISBN: 978-85-7631-393-9 .
- 2- ROSA, C.A.P. **História da Ciência: o pensamento científico e a ciência no século XIX.** Brasília: FUNAG, 2012, 2ª ed. ISBN: 978-85-7631-395-3 .
- 3- ROSA, C.A.P. **História da Ciência: a ciência e o triunfo do pensamento científico no mundo contemporâneo.** Brasília: FUNAG, 2012, 2ª ed. ISBN: 978-85-7631-396-0 .
- 4- **BIOINFORMÁTICA DA BIOLOGIA À FLEXIBILIDADE.** Organização de Hugo Verli. Porto Alegre, 2014.
- 5- SILVA, I.N.; SPATTI, D.H.; FLAUZINO, R.A. **Redes Neurais Artificiais: para engenharia e ciências aplicadas.** São Paulo: Artliber, 2ª edição, 2016.

APÊNDICE I – Manuscript draft

***In vitro* anti-RSV activity of Quercetin Pentaacetate:
Possible mechanism is through interaction with F protein**

RIBEIRO, A.G.; LOPES, B.R.P., COSTA, M.F.; IACOVELLI, F.; FALCONI, M.; SILVA, T.F.; SOUZA, P.S.S.; ARAÚJO, G.C.; CARUSO, I.P.; SOUZA, F.P.; TOLEDO, K.A.

Abstract

Lower respiratory tract infections are the main cause of infections in childhood, reaching approximately 12 million severe cases and 3 million very serious cases around the world. The human pneumovirus respiratory syncytial virus (RSV) is the major pathogen in these cases which contributes to several cases of bronchiolitis, pneumonia and chronic obstructive pulmonary infections in people of all ages, especially in infants and the elderly. There are currently no vaccines and the therapeutic strategies in use present little efficacy or restricted use. Flavonoids are the focus of several studies that seek for effective and cost-effective antiviral drugs. Quercetin is a flavonoid commonly found in the human diet and has an inhibitory activity against several types of virus, including RSV. Chemical modification of molecules has proven to be a promising strategy to improve the biological activity of natural compounds. The aim of this work was to evaluate the anti-RSV activity of Quercetin and Quercetin Pentaacetate (QPA) in culture of HEP-2 cells in different protocols. The antiviral assays were performed with different MOIs (0.1, 0.5 and 1.0), evaluated by MTT protocol and confirmed by plaque assay. Quercetin and QPA showed anti-RSV activity through virucidal mechanism which was confirmed by adsorption and addition time assays. QPA presented inhibitory activity in the adsorption and addition time assays, what confirmed its activity occurs mainly in the early phases of the viral cycle. The plaque assay corroborated the reduction in the number of viral foci in virucidal protocols as 50%. Classical molecular dynamics (MD) simulations of RSV-F fusion protein in complex with Quercetin suggested hydrophobic interactions and hydrogen bonding between F protein. Our data indicate that QPA possess a promising anti-RSV activity, which is suggested to be related with the interference of RSV-F fusion activity.

Key words: RSV, antiviral, flavonoids, quercetin, quercetin pentaacetate

Introduction

The human pneumovirus Human Respiratory Syncytial Virus (hRSV) was isolated and described as the main pathogen responsible for acute lower respiratory tract infections (ALRTI), and it is worldwide distributed (Weissebach, 1998; Jennings, et al 2004; Iwane, 2004; Guthrie, 2005). RSV infections affect approximately 70% of newborns during their first year of life and 95% of children up to 2 years (Collins, 2002). The infection begins with the inhalation of infectious viral particles or through direct contact with nasopharyngeal secretions expelled by infected persons (Hall, 1981). RSV-ALRTI is considered a public health issue at a global level, due to the hospitalization rates and high treatment costs.

The prophylactic method recommended by the FDA (Food and Drug Administration), is the administration of Palivizumab (Synagis®), a monoclonal antibody which binds to RSV-F (Yu et al., 2006). However, the costs regarding that prophylaxis are very high what limits its application to high-risk children. Currently, many research attempt to develop an effective treatment of general use for patients with RSV infection. The focus of most studies are the surface glycoproteins G and F, as they mediate the initial steps of adhesion and fusion of the virus to the host cell. Another licensed drug (1985) for use in children is Ribavirin (Virazole®), which possesses broad antiviral action. As it is a nucleoside analogue, it inhibits the nucleic acid synthesis of virus, interfering with virus replication. Nevertheless, it is an antiviral agent with potentially mutagenic and teratogenic (VENTRE; RANDOLPH, 2004).

Vegetables are source of a wide range of compounds, usually derived from secondary metabolism, such as alkaloids, flavonoids and others,—with great medicinal potential. (Formica and Regelson 1995, Kinghorn 2002). Flavonoids represent one of the most important and diversified chemical groups produced by plants and are widely distributed in the plant kingdom. Quercetin 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxychromen-4-one is the main flavonoid present in the human diet. It is one of the most potent flavonols with a variety of pharmacological activities described in the literature such as anti-HIV activity through inhibition of its reverse transcriptase as well as antiviral activity against ~~and~~ other retroviruses, herpes simplex type 1, type-1 poliovirus, type-3 parainfluenza, and hepatitis-C (Gonzales et al., 2009). Kaul et al. (1985) demonstrated that Quercetin has anti-RSV activity due to virucidal mechanism, besides being able to inhibit the viral replication in infected cells.

The study of antiviral activity of flavonoids has been investigated since 1940. However, only in the recent decades there have been attempts at synthesizing chemical modifications in these molecules in order to enhance their capacity to control viruses (Cushnie and Lamb, 2005). Danihelová and colleagues (2013) suggested that the acetylation of Quercetin probably favors absorção through the cell membrane. In this regard, the acetylation of Quercetin could improve its anti-RSV activity in the virucidal protocol, besides inhibiting the viral replication in cells infected with multiplicity of infection (MOI)≥0.1.

The aim of our study was to evaluate the anti-RSV activity of Quercetin and whether the chemical modification of Quercetin molecule to Quercetin pentaacetate-QPA could improve its anti-RSV activity.

Materials and methods

Reagents

The acetylated derivative (3,3',4',5,7- Pentaacetoxyflavone) was provided by Dr. Luis Otávio Regasini (Institute of Biosciences, Letters and Exact Sciences, IBILCE, UNESP, Brazil). Briefly, the ~~original~~ quercetin molecule was subjected to acetylation reactions using a solution of acetic anhydride and pyridine (1:1, 200 mL) at 80 °C for 5 days. The reaction was monitored by analytical thin layer chromatography (ADCC). The crude product was subjected to normal phase column chromatography (CC-FN; 3cmx13cm), eluted with Hex: AcOEt (80 → 0% Hexane), yielding 40 fractions of 35ml each. The fractions were pooled according to their chromatographic profile delineated by CCDA using Hex: AcOEt (1:1) as the mobile phase. The major product was pooled and named as QPA. Quercetin 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxychromen-4-one, Sigma-Aldrich, St. Louis, MO, USA) and QPA were diluted in dimethyl sulfoxide (DMSO; Sigma-Aldrich), protected from light, and stored in a freezer at -80°C for up to one month.

Cell Culture

Oropharynx carcinoma (Hep-2; Cell Bank of Rio de Janeiro) were grown at 37°C, 5% CO₂ in DMEM-F12 medium (Sigma) supplemented with 10% fetal bovine serum (FBS) (Vitrocell, Campinas, SP, Brasil) complemented with antibiotics and antimycotics (Gibco, Life Technologies, Gaithersburg, MD, USA). HEp-2 cells are derived from human respiratory tract and permissive to HRSV infection and replication (Brücková et al., 1978; Srinivasakumar et al., 1991).

Viral Stock

HRSV virus strain Long was provided by Dr. Eurico de Arruda Neto (Universidade de São Paulo, Brazil) and amplified in Hep-2 cells. Viral stocks were prepared in phenol red-free DMEM-F12 medium plus 2% FBS. Viral fractions were titrated and kept at -80°C freezer. The viral titer was established by plate reduction assay (McKimm-Breschkin, 2004) and confirmed by TCID₅₀ (Rasmussen et al., 2011).

Cytotoxicity

Cytotoxicity was evaluated by the colorimetric method of MTT (3-(4,5-dimethylthiazol- 2-yl)2,5-Diphenyl Tetrazolium Bromide) (Sigma). Briefly HEp-2 cells (5x10⁴/well) received Quercetin diluted in DMEM-F12 2% FBS. After 3 days of incubation, MTT was added and the absorbance was measured spectrophotometrically in a plate reader (FC™ Microplate Multiskan Photometer, Thermo) at 560nm. Cells cultured on medium alone were considered as 100% viable. CC₅₀ values were obtained by fitting a fourparameter nonlinear sigmoidal regression curve using GraphPad Prism6 software.

Evaluation of the anti-hRSV action of Quercetin and Quercetin Pentaacetate

Virucidal assay

HEp-2 cells were plated in 96-well plates (5×10^4 /well) and maintained for 24 hours in DMEM F12 culture medium supplemented with 10% FBS. The assay was performed as described by Petricevich and Mendonca (2003), with some modifications. The suspensions of hRSV (MOIs 0.1, 0.5 and 1.0) were incubated for 1 hour at 4 °C in the presence of different concentrations of Quercetin and QPA (1.5-24 μ M). At the end of the incubation, Quercetin and QPA-treated viral aliquots were added to the HEp-2 cell monolayers and then incubated for a period of 2 hours at 37 °C. The inoculum was replaced with DMEM-F12 2% (FBS) medium with final volume of 100 μ L. The plate was incubated for four days at 37 °C and 5% CO₂, and on the fourth day, MTT was added and the absorbance was measured spectrophotometrically in a plate reader (FC™ Microplate Multiskan Photometer, Thermo) at 560nm.

Post-treatment assay

Hep-2 cells were infected with RSV (MOI 1, 0.5 and 0.1) and incubated for 2 hours at 37 °C. The viral suspension was then replaced by different concentrations of Quercetin and QPA (1.5-24 μ M) in a final volume of 100 μ L. The plate was incubated for four days at 37 °C and 5% CO₂ and on the fourth day MTT was added and the absorbance was measured spectrophotometrically in a plate reader (FC™ Microplate Multiskan Photometer, Thermo) at 560nm.

Evaluation of possible mechanisms involved in anti-hRSV action

Adsorption Assay

The adsorption assay was performed as described by Kimura et al. (2000). hRSV suspension (MOI 0.5) was incubated for 1 hour at 4 °C or 37 °C in the presence of different concentrations of QPA (1.5-12 μ M). HEp-2 cells were precooled at 4 °C for 30 minutes before receiving MOI 0.5 containing viral suspension and QPA at different concentrations. Under these conditions, cells, virus and QPA were maintained for 2 h before being washed with physiological sodium chloride solution (0.15M NaCl) to remove the non-adsorbed viral particles. The plate was incubated for four days at 37°C and 5% CO₂ and on the fourth day MTT was added and the absorbance was measured spectrophotometrically in a plate reader (FC™ Microplate Multiskan Photometer, Thermo) at 560nm.

Time of Addition Assay

HEp-2 cells in the presence of F12 DMEM culture medium supplemented with 10% FBS were plated in a 96-well plate (5×10^4 /well) and incubated for 24 hours to reach confluence of approximately 80%. Hep-2 cells were infected with hRSV solution (MOI 0.5) for 1h, after which the inoculum was

replaced with DMEM-F12 2% (FBS) medium with final volume of 50 μ L in all wells. At time 0, the viral suspension with MOI 0.5 was added, followed by QPA addition. In the following intervals, the different concentrations of QPA were added over 48 hours, respecting the ranges: 3h, 6h, 12h, 24h, 36h and 48h post-infection. At the end of 48 hours, the plate was incubated at 37 °C, and on the fourth day, MTT was added and the absorbance was measured spectrophotometrically in a plate reader (FC™ Microplate Multiskan Photometer, Thermo) at 560nm.

Plate Assay

HEp-2 cells, grown in 24 well plates (3x10⁵/well), were treated and infected according to the procedures described above and the assay was performed according to the method described by McKimm-Breschkin, 2004. For visualization of the lysis plates, the cell monolayer received medium overlay (culture medium containing 0.6% agarose). Under these conditions, the cells were maintained at 37 °C 5% CO₂ for a period of 6 days or until the visualization of lysis plates with the aid of an optical microscope. Cells were formalized with 1% formaldehyde diluted in 0.15M saline. The lysis plates were evidenced after discarding of the overlay medium and addition of 0.05% neutral red solution which remained for 1 day at room temperature.

Molecular Modeling of the hRSV-F Glycoprotein

The Protein Data Bank (PDB) contains more than 20 crystal structures of hRSV-F Glycoprotein, in complex with antibodies or inhibitors. These structures lack the three helices forming the coiled-coil in the C-terminal region and most of the mucin-like regions. To obtain the complete protein structure, we performed a modeling procedure by submitting the corresponding sequence of hRSV-F Glycoprotein (UniProtKB: P03420, residues 1–574) to the I-TASSER web server and assigning the PDB crystal structure (ID: 4JHW) as a template. The I-TASSER server combines threading and *ab-initio* procedures and has been evaluated as the best threading modeling program in recent community-wide Critical Assessment of protein Structure Prediction (CASP) experiments (Yang, J. et. al., 2015), providing an on-line workbench for high-resolution modeling of protein structures and functions (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>). I-TASSER generates five models with different C-scores and structures. A C-score represents a confidence score for estimating the quality of models predicted by I-TASSER, and is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. The C-score is typically in the range of -5 to 2 and a higher value signifies a 3D model with a high confidence and vice versa (Yang, J. et. al., 2015). All 3D models produced by I-TASSER resemble the 4JHW crystal structure, and the highest C-score (equal to 0.08) was chosen to perform the final modeling. The modeled structure was then imported to UCSF Chimera, triplicated, and superimposed through the MatchMaker tool for biological assembly of 4JHW structure to obtain the final model.

Model Preparation for Molecular Dynamics (MD) Simulation

The system topology and coordinates, used as input for the GPU-enabled version of *pmemd*, the MD engine included in the AMBER-12 package, were obtained through the Leap module of AmberTools12 ([Case, D. et al., 2012](#)), assigning parameters for structures through AMBER-FF14SB force-field ([Hornak, V. et al., 2006](#)). Single-point calculations were performed in Gaussian-09 at HF-SCF/6-31G* level on the optimized structures to derive electrostatic potential and restrained electrostatic potential (RESP) partial charges were obtained from electrostatic potential using antechamber module in AMBER-12. The general AMBER force field (gaff) was used to parametrize the inhibitors using antechamber. The complexes were immersed in a rectangular box filled with TIP3P water molecules ([Pettersen, E.F. et al., 2004](#)), imposing a minimum distance between the solute and box of 14Å. Charges were neutralized by adding Cl⁻ counter-ions to solvated systems in favorable positions.

MD Protocol

A minimization run (5000 steps) was performed using the steepest descent, imposing harmonic constraint of 50 kcal·mol⁻¹·Å⁻² to remove unfavorable interactions introduced by the modelling procedure. The system was gradually heated in the isovolumic-isothermal (NVT) ensemble from 0 to 300 K over a period of 500 ps using Langevin thermostat ([Izaguirre, J.A. et al., 2001](#)) [with a coupling coefficient of 1.0 ps and a weak constraint of 15 kcal·mol⁻¹·Å⁻²](#). In the last equilibration step, the system was subjected to an equilibrium simulation for 1 ns, removing all constraints. The optimized system was then simulated using the isobaric-isothermal (NPT) ensemble for 50 ns, using periodic boundary conditions, a 2.0 fs time-step, a cut-off of 9 Å for evaluation of short-range non-bonded interactions, and the PME method for long-range electrostatic interactions. SHAKE algorithm was used to constrain covalent bonds involving hydrogen atoms. Temperature was fixed at 300K using Langevin dynamics, while pressure was held constant at 1 atm through the Langevin piston method. The atomic positions were saved every 500 steps (1.0 ps) for analyses. The simulations were entirely performed using 2 NVIDIA TESLA C2075 GPUs.

MD Simulations

The distance and hydrogen bonds between protein and ligands were evaluated using *g_dist* and *g_hbond* modules of the GROMACS 4.6.7 MD package ([Abraham, M. et al., 2014](#)). Contacts were evaluated through a personalized tool based on the python pmx package, considering all protein atoms interacting with ligands in a 4.0 Å radius. Interaction energies were evaluated through the MM-PBSA (Molecular Mechanics-Poisson-Boltzmann Surface Area) module of the Amber12 package using the MM-GBSA approach ([Weiser, J.; Shenkin, P. S.; Still, W. C., 1999](#)). Structural images were obtained using VMD ([Humphrey, W.; Dalke, A.; Schulten, K., 1996](#)) and UCSF Chimera programs ([Pettersen, E.F. et al., 2004](#)), while plots were obtained through the Grace plotting tool (<http://plasma-gate.weizmann.ac.il/Grace/>).

Molecular Docking

Protein-ligand docking was executed with the Auto-Dock/Vina 1.1.2 program, using the AutoDock/Vina PyMOL plugin (Trott, O. and Olson, A. J., 2010). Quercetin Pentaacetate and JNJ-2408068 SDF files, downloaded from the PubChem compound database, were converted into PDB files and filled with hydrogens using the Open-Babel program (O'Boyle, N.M. et al., 2011). The program SYBYL 6.0 was used to minimize the structures of compounds. Docking runs were performed using the protein centroid clusters as receptors obtained via clustering a previously performed molecular dynamics trajectory. Docking simulation was performed using the genetic algorithm with local gradient optimization. For each docking simulation, a box of 75x75x75 (grid points in XYZ axes) was centered over the binding site corresponding to empty cavities generated by the Phe488 residues. The affinity of docked compounds was expressed as binding energy (kcal/mol). Each docking simulation run took approximately 10 min on a dedicated Intel-i7 CPU workstation.

Statistical analysis

The experimental data were processed using the software Microsoft Excel and Graphpad Prism 6 and evaluated by analysis of variance (one-way ANOVA), followed by the Bonferroni test. A significance level of $p \leq 0.05$ was adopted.

RESULTS AND DISCUSSION

Cellular cytotoxicity of Quercetin and QPA in HEp-2 cells

Initially, the cytotoxic concentration was determined for 50% of Hep-2 cells (CC_{50}). The range of Quercetin and QPA concentrations for CC_{50} determination was based on previous studies with Quercetin (Choi et al., 2009; Johari et al., 2012). HEp-2 cells (5×10^4) were incubated with Quercetin and QPA at concentrations (2-1024 μ M) and the viability of these cells was analysed after 3 days of compound addition. Cells not incubated with Quercetin and QPA were assumed to be 100% viability. The CC_{50} obtained for QPA compound was 38.8 μ M (Supplemental Figure). This value was higher than that determined for Quercetin (11.4 μ M) in Hep-2 cells (Supplemental Figure) and higher than the CC_{50} determined for human cervical adenocarcinoma (HeLa) epithelial cells, 29.6 μ M (Danihelová et al., 2013). To our knowledge, no study has been conducted regarding the cytotoxic concentration of QPA on Hep-2 cells, emphasizing the importance of this study.

Quercetin presents anti-RSV activity

In addition to the studies initiated by Kaul et al. (1985), we demonstrated that quercetin has a virucidal activity (16-30%) for MOI 0.1 and lack of protection for infected cells at concentrations of 2.5–10 μ M. For MOIs -0.5 and 1.0 Quercetin did not inhibit RSV infection. Our data corroborate those of Kaul et al., who reported that Quercetin (12.5 μ M) protected up to 38% of cells in a virucidal protocol.

The minimum viral load in our study (MOI 0.1) was 142-fold higher than that of Kaul *et al.* All results from the pretreatment and post-treatment protocols were negative. None of the tested experimental conditions showed statistically significant cellular protection (Supplemental Figure). This effect is similar to that described previously ([Rasmussen, L. et. al., 2011](#)).

QPA presents anti-hRSV activity

The treatment of the viral particle with QPA ($\leq 12\mu\text{M}$) resulted in high cellular protection cell protection of about 100% for MOI 0.1 at all concentrations of QPA (1.5-24 μM). As for morphological analysis, the cells of viral control presented elongated morphology, numerical reduction and high syncytium formation. For the concentrations of QPA that were effective in the protection of cellular damage (1.5-12 μM), a great decrease in the number of syncytia in relation to viral control was observed ().

The viral inoculum (MOI) was increased in order to establish a threshold in the anti-hRSV activity of QPA. We observed that the rise of the viral inoculum resulted in reduction of the anti-hRSV activity of QPA. Thus, 5x and 10x larger inoculum (MOI 0.5 and 1.0), treated with concentrations $\leq 12\mu\text{M}$ resulted in cellular protection values of 50-75% or of maximum 50%, respectively. The concentration of 24 μM QPA was not considered effective for MOIs of 0.5 and 1.0, since it reduced the number of viable cells. Data analysis suggests that the presence of 24 μM QPA in association-with the growth in viral load (0.5 and 1.0), acted synergistically regarding the individual cytotoxic effect. This hypothesis would explain, at least partially, the reduction in the number of viable cells. The morphology of the cells and the formation or absence of syncytia in the virucidal treatment for MOIs 0.5 and 1.0 were also analyzed microscopically after three days of culture. Morphological analysis of cells showed that in all concentrations a large reduction in the number of syncytia was observed and, in consequence, there was a rise in the number of viable cells.

In the post-treatment assay the results showed that the presence of QPA (1.5-24 μM), after the viral inoculum of MOI 0.1, provided cellular protection (25% to 100%) (In the cytomorphological analysis (), uninfected cells (cell control) presented confluence, no syncytium formation and absence of dead cells. Infected cells (viral control) have elongated morphology, reduced number of cells and the presence of syncytia. For MOI 0.1, concentrations of 1.5-6 μM induced a slight reduction in syncytium formation compared to viral control. Only the concentration of 12 μM drastically reduced the formation of syncytia. The increase in viral inoculum (MOI 0.5) resulted in cellular protection only for the concentration of 24 μM (~ 25%) ($p < 0.001$), was 24 μM (~ 25%) (). The rise of the viral inoculum for MOI 1.0 (2-fold higher) culminated in a marked-diminution-of the anti-RSV effects of QPA, similarly to what was obtained with MOI 0.5. Only the concentration of 24 μM QPA led to a slight maintenance of cell viability (~ 25%) with statistically significant cellular protection ($p < 0.01$), when compared to viral control (0%) (). The protective effect in the post-treatment assays was not observed for higher inocula (MOIs of 0.5 or 1.0), with no protection > 50%. The lack of cellular protection in MOI 0.5 and 1.0 may indicate a limited after treatment activity of QPA in situations where viral load is high. As for the cell

morphological analysis the infections with MOI 0.5 and 1.0 showed similarities (). There was syncytium formation and cell death at concentrations of 1.5-6 μM , which reduced at the concentration of 12 μM .

The results indicate that QPA has anti-RSV activity with direct effect on the viral particles, suggesting an interaction between the compound and the proteins present on the viral surface, namely G, SH or F. According to the literature, the proteins G and SH are not fully required for RSV infection, since G and SH protein-deficient mutants shows that F protein is the only transmembrane protein essential for viral infection in vitro (Bukreyev et al., 1997; et al., 2000, Karron et al., 1997 Collins and Graham, 2008, Bohmwald et al., 2016). In order to investigate by which virucidal mechanism QPA was inhibiting the cytopathic effect of hRSV, adhesion and time-of-addition assays were performed. Initially, the ability of QPA to inhibit virus adhesion to the cell surface was evaluated. Viral incubation with QPA compound at 4 °C resulted in mild cell protection (~ 25%) at the concentration of 1.5 μM . There was no statistically significant difference in that concentration when compared to the viral control (0%). In the other concentrations, no cellular protection was observed.

Although the standard adsorption protocol specifies that the incubation of the viral particle with compounds must be performed at 4 °C (Kimura et al., 2000), it is known that temperature can act as a facilitator of chemical reactions, besides being related to the degree of molecules agitation, what facilitates a greater number of shocks among them and, thus, increase the number of effective interactions (Lavorenti, 2002). The adhesion assays were therefore repeated and the incubation of the virus with QPA was performed at a temperature of 37 °C. The remaining steps were not modified. Data show that all the concentrations (1.5-12 μM) were effective in reducing 50% of RSV cytopathic effect. It can be inferred that the temperature under which the virus was incubated with QPA led to an improvement of anti-RSV activity. We suggest that the inhibition of viral adsorption to the cell surface is due to the interaction of QPA with hRSV-F.

Other steps in the viral cycle can be assessed by time-of-addition assays. They consist on the addition of the tested compounds to the cell culture at different times of the viral infection. As shown herein, there was cell protection of 50% at 0 hour at 3 μM with a statistically significant difference $p < 0.001$ in relation to the viral control (0%). The effect disappears in the following hours (3-12 h) and returns after 24 h of infection (~ 25%) and maintaining this effect in the intervals of 36 h and 48 h. In the second condition (Figure 6B), the addition of QPA to the infected cells in the periods of (0-48 h) presented protection (~ 25% to ~ 50%) at 6 μM with statistically significant differences when compared to in relation to the viral control (0%) in all intervals.

Regarding the concentration of 12 μM (), the cell protection was about (50%) at hour 0, which was reduced at the intervals of 3 h, 6 h and 12 h. After 24 h of infection, the addition of QPA to the infected cells showed a slight increase in cell viability (~ 25%), but with significant results only at the 36 h and 48 h intervals. According to Douglas and Panis (2003), who evaluated the inhibitory activity of triphenol compound VP-14637 (ViroPharma) against hRSV in the same protocol, the compound acted as a virus fusion inhibitor because there was a large increase in cell viability when the compound was added at 0 h, along with the viral inoculum. Our results resemble those from Douglas and Panis (2003), since cellular protection (50%) appeared at time 0, what confirmed virucidal assay, as well as the

hypothesis that there is interaction between the molecule (QPA) and F protein, leading to interference in virus replication.

The results obtained with the MTT salt protocol in the virucidal and adhesion assays were confirmed by counting the number of plaques in plaque reduction assay. Initially, lysis plate counts were performed visually with the aid of a negatoscope. The 24-wells plate was placed on the negatoscope, and the count of the number of lysis plates from each well was performed for analysis of the results. Through visual counting, it was possible to observe that in wells treated with QPA, there was a gradual reduction in the number of lysis plaques when compared to the viral control group (virus without QPA). The counting was then performed for both protocols (virucidal and adhesion). However, there were difficulties in counting since the process is laborious and subjective. The Laboratory of Bioinformatics-Unesp Campus of Assis, developed an algorithm based on image processing and pattern recognition using MATLAB® software (with the aim of automating the identification and counting of lysis plates, minimally depending on human intervention). MATrix LABoratory). The automated counting has guaranteed us a standardized, accurate and reproducible result. Statistical analysis of the number of lysis plates showed that there were significant differences in all the concentrations tested (1.5 - 12µM) when compared to the viral control, confirming the previously obtained data.

Quercetin virucidal effect suggests that it can interact with RSV surface proteins. The role of Fusion protein (RSV F) in the RSV infection has been reported and its structure documented in specialized database. RSV F is expressed as a single precursor composed by 574 aminoacids, F0, which oligomerizes in the endoplasmic reticulum and is proteolytically processed at two sites in each monomer, resulting in an oligomer of two disulfide-linked fragments: F2 and F1 (Morton, C.J. et al., 2003). RSV F is a type I fusion protein (Walsh, E.E., Hruska, J., 2016; McLellan, J.S. Chen, M., Leung, S. et.al, 2013), whose mature state exists as a homotrimer (association of F1-F2 units), folded into a compact prefusion structure (Battles, M.B., Langedijk, J.P., Furmanova-Hollenstein, P. et al., 2016; McLellan, J.S. Chen, M., Leung, S. et.al, 2013). There are two states of conformation of RSV F: a metastable state adopted before virus-cell interaction (prefusion) and a stable state that occurs after merging of virus and cell membranes (postfusion) (McLellan, J.S. Chen, M., Leung, S. et.al, 2013). RSV F is essential for infection, as mutant RSV lacking F protein (RSV ΔF) are not able to infect cells on their own but rather requires a helper virus to gain entry into (Mastrangelo, P., Hegele, R.G., 2013; Batonic, M., Oomens, A.G., Wertz, G.W., 2008). Also, RSV F expression on the surface of the cell can cause fusion with neighboring cells, especially in vitro, leading to the formation of multinucleated syncytia (McLellan, J.S. Chen, M., Leung, S. et.al, 2013). Furthermore, RSV F is one of the targets for pharmacological anti-RSV intervention because it is essential for entry, as it facilitates pH-independent fusion of the viral membrane with the host-cell plasma membrane, leading to infection of the host cell (Walsh, E.E., Hruska, J., 1983, McLellan, J.S. Chen, M., Leung, S. et.al, 2013).

Recently, the major target for intervention has been RSV F prefusion conformation. Several regions of RSV F have been targeted as pharmacological intervention candidates, as antigenic sites (López, J.A., Bustos, R., Örvell, C. et al., 1998; Zimmer, G., Trotz, I., Herrler, G., 2001), N-glycosylation sites (Zimmer, G., Trotz, I., Herrler, G., 2001), cysteine residues (Day, N.D., Branigan, P.J., Liu, C. et

al., 2006), as well as different RSV F residues which can interact with peptides and small-molecule RSV entry inhibitors (McLellan, J.S. Chen, M., Leung, S. et.al, 2013; Douglas, J.L., Panis, M.L., Ho, E. et al., 2005, Sun, Z., Pan, Y., Jiang, S., Lu, L., 2013).

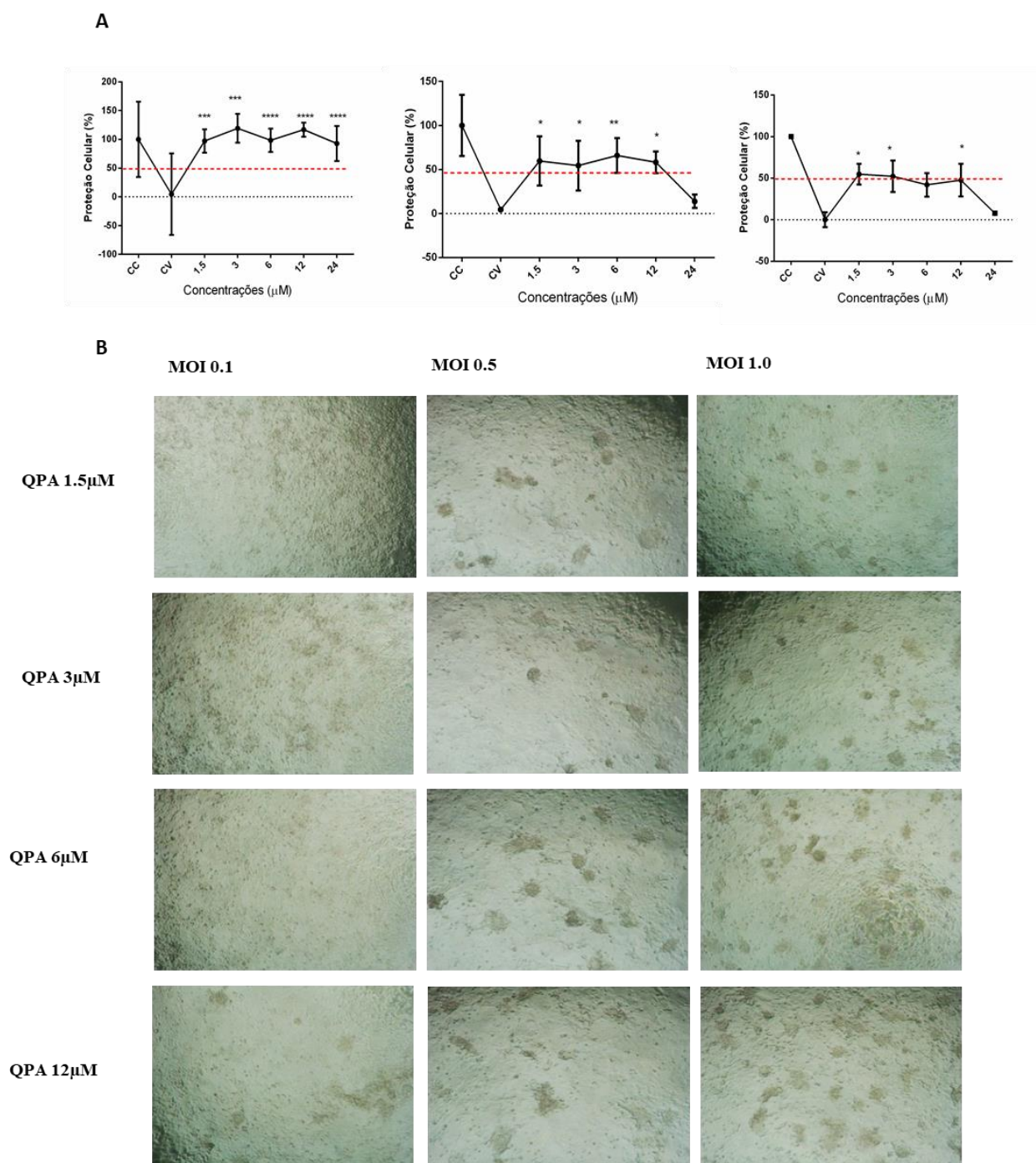
Battles *et al.* (2016) have recently showed a mechanism in which possible inhibitory-virus entry small molecule candidates targeting F protein, bind with high affinity to a three-fold-symmetric pocket within the central cavity of prefusion RSV F, thereby stabilizing the metastable prefusion conformation of RSV F by restricting the movement of two structurally labile regions (Figura x). Among these candidates, JNJ-24008068/R170591 binds to that three-fold-symmetric pocket, interacting through aromatic stacking with phenylalanine (F) (residues 488- chain A, 488 - chain B, 488 - chain C) and through electrostatic interactions with the negatively charged pocket formed by aspartate (D) and glutamate (E) (residues 486, 487, 489 – chain A; 486, 487, 489 – chain B; 486, 487, 489 – chain C) (McLellan, J.S. Chen, M., Leung, S. et.al, 2013). The interaction between the modelled RSV F (RSVmF) and quercetin or JNJ-24008068/R170591 inhibitor as a control have been investigated by docking analyses and classical MD simulations. The docking experiments indicate that both molecules completely fill the previously described cavity and are characterized by the same interaction energy (around -6.5 kcal/mol). In order to overcome docking limitations and explore the conformational panorama sampled by the ligands, each of the two best complexes has been simulated through MD for a total time of 20 ns.

Trajectory analysis indicated a diverse behavior for ligands, as shown in Figure 3, which shows the time-dependent evolution of the distance between the center of mass of the binder and the center of mass of the central cavity. Quercetin shows a tendency to escape from that cavity, while the inhibitor and the QPA remain there, due to the constant hydrophobic interaction with the phenylalanines (residues 488, A, B and C chains). Hydrogen bonds indicated that quercetin only establishes a hydrogen bond with the protein, whereas the inhibitor establishes three hydrogen bonds with two chains, one of which remained for more than 75% of the simulation time (Table 1). It was not possible to verify hydrogen bonds between QPA and protein, which suggests other types of interactions as Van der Waals forces.

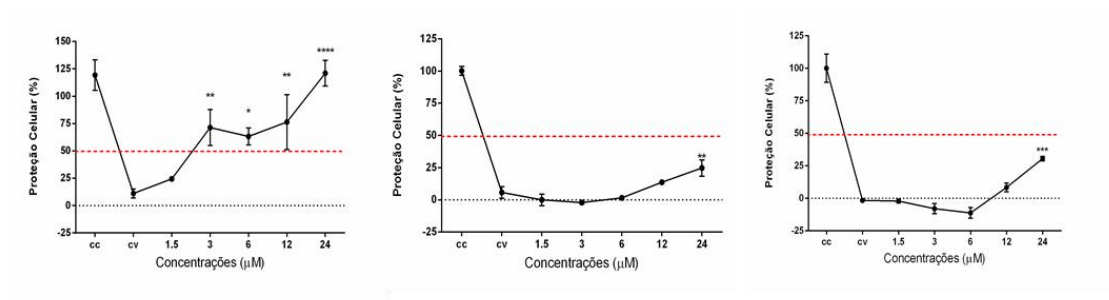
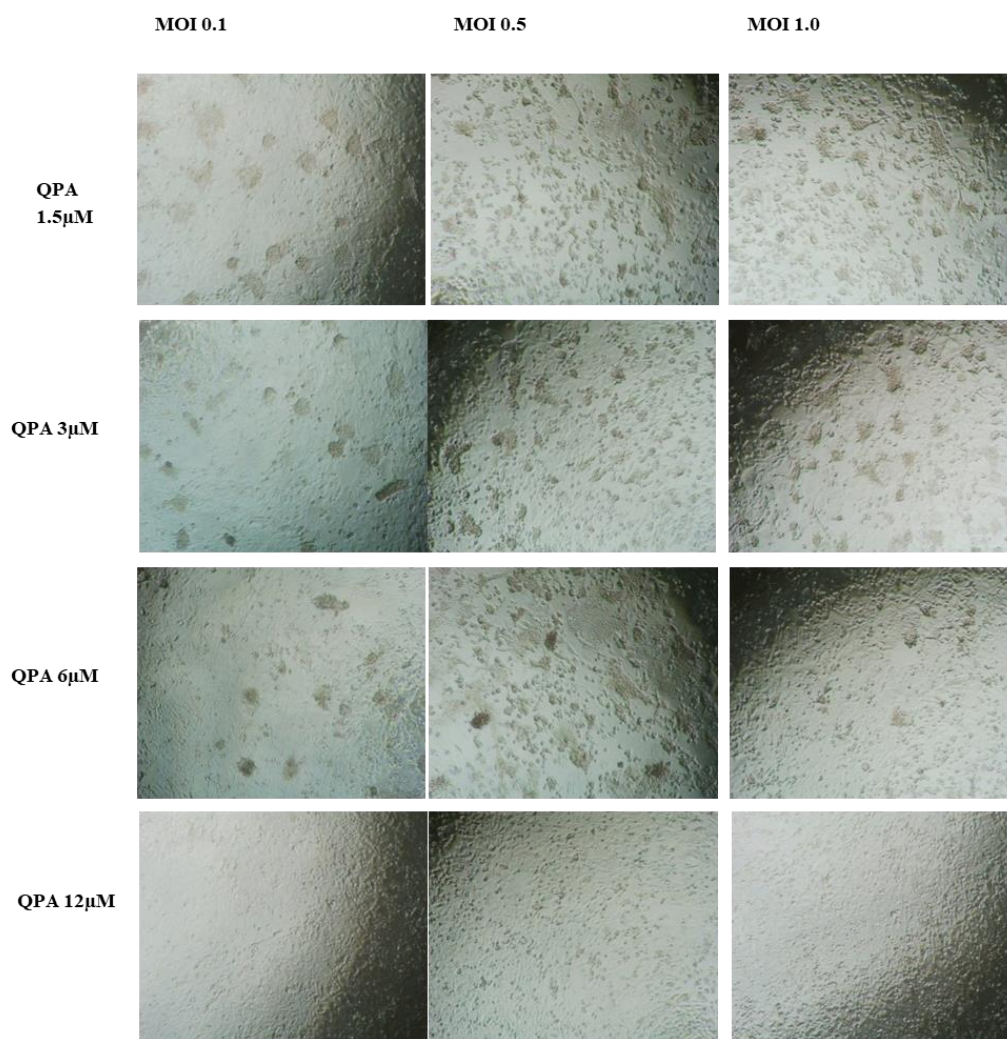
The MM-GBSA method was employed to calculate the free energy difference between the protein and the ligands as a final confirmation. Basically, the total free energy of Gibbs obtained by this method represents a measure of how favorable or also spontaneous the formation of a complex or a reaction can be. The more negative the value, the more spontaneous and favorable the reaction / complex formation is. The results indicated that RSVmF in complex with quercetin had a ΔG value of -4.92, whereas for the QPA a ΔG of -14.22 was obtained and for the inhibitor a ΔG of -22.78, as shown in Table 2 RSVmF in complex with quercetin presented a lower value than the complex with the inhibitor. However, the QPA derivative, although still lower in value than the inhibitor, is already much closer in energetic terms. These results, together with the experimental results, show that, although quercetin and QPA have lower energy values, they are still able to induce a virucidal effect, which is more promising for QPA and such a mechanism has been investigated.

Conclusion

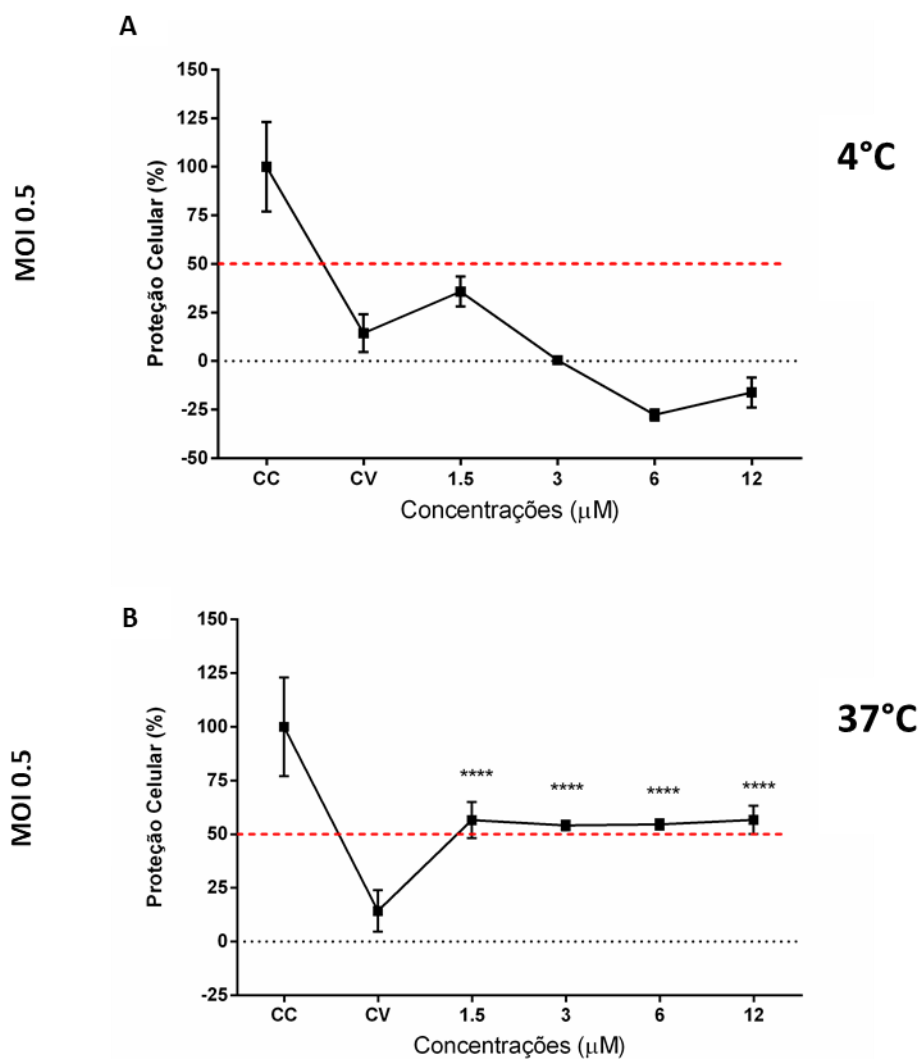
We demonstrate *in vitro* and *in silico* that QPA compound derived from acetylation of Quercetin reduces cytotoxic effects and increases its virucidal activity. QPA has a promising anti-hRSV activity by virucidal mechanism, as it reduces cell damage and syncytium formation. The interactions between hRSV-F and QPA is at least partially responsible for such effects.



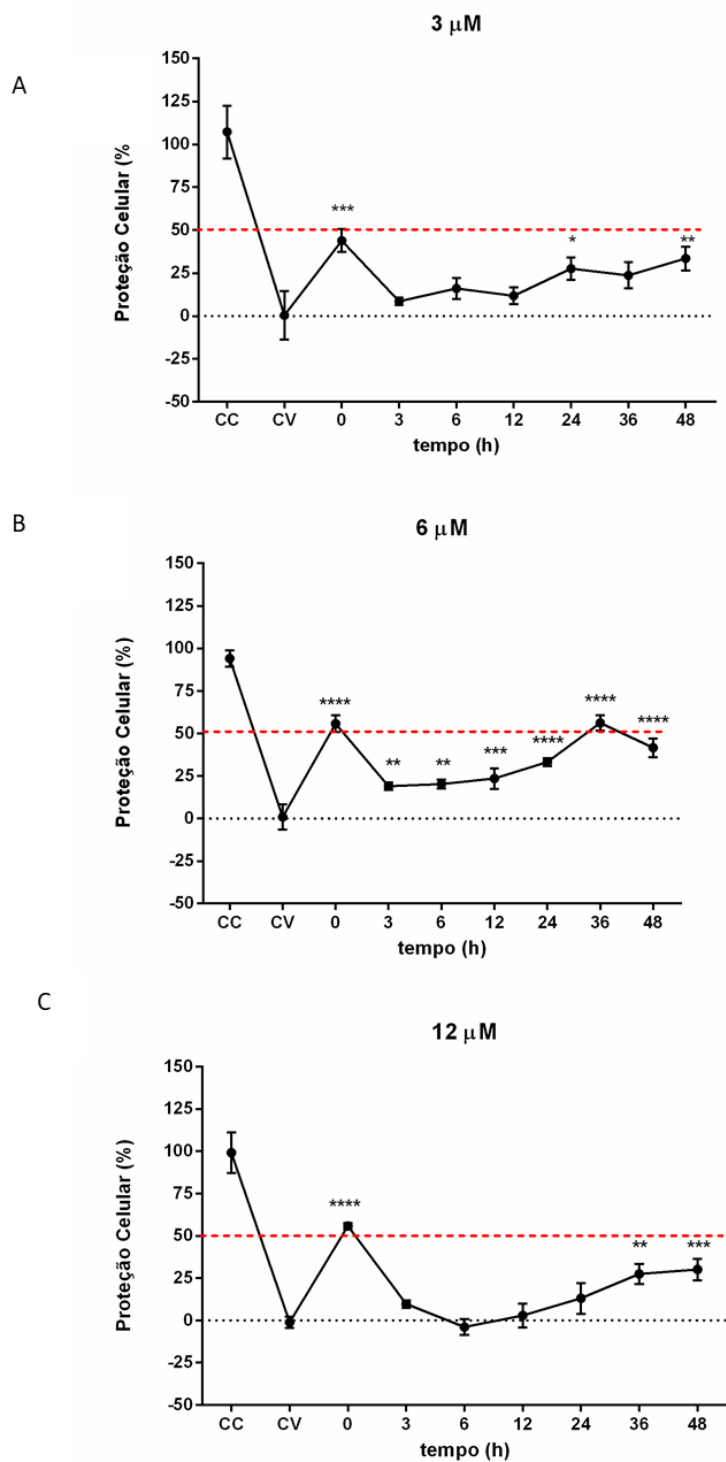
(A) Avaliação da atividade virucida do composto QPA nas concentrações de 1.5-24µM em diferentes MOIs. Os dados representam a média de um experimento realizado em triplicata. Os dados foram analisados com auxílio do software Graphpad Prism 6 e os valores expressam média $\pm$ desvio padrão (One-way ANOVA, seguida pelo teste de Bonferroni, $*p \leq 0.05$). (B) Morfologia de células HEp-2 infectadas por hRSV (MOI 0.1, 0.5 e 1.0) com diferentes tratamentos de QPA (1.5-12 µM). As imagens foram avaliadas e documentadas em microscópio de luz branca (4X).

A**B**

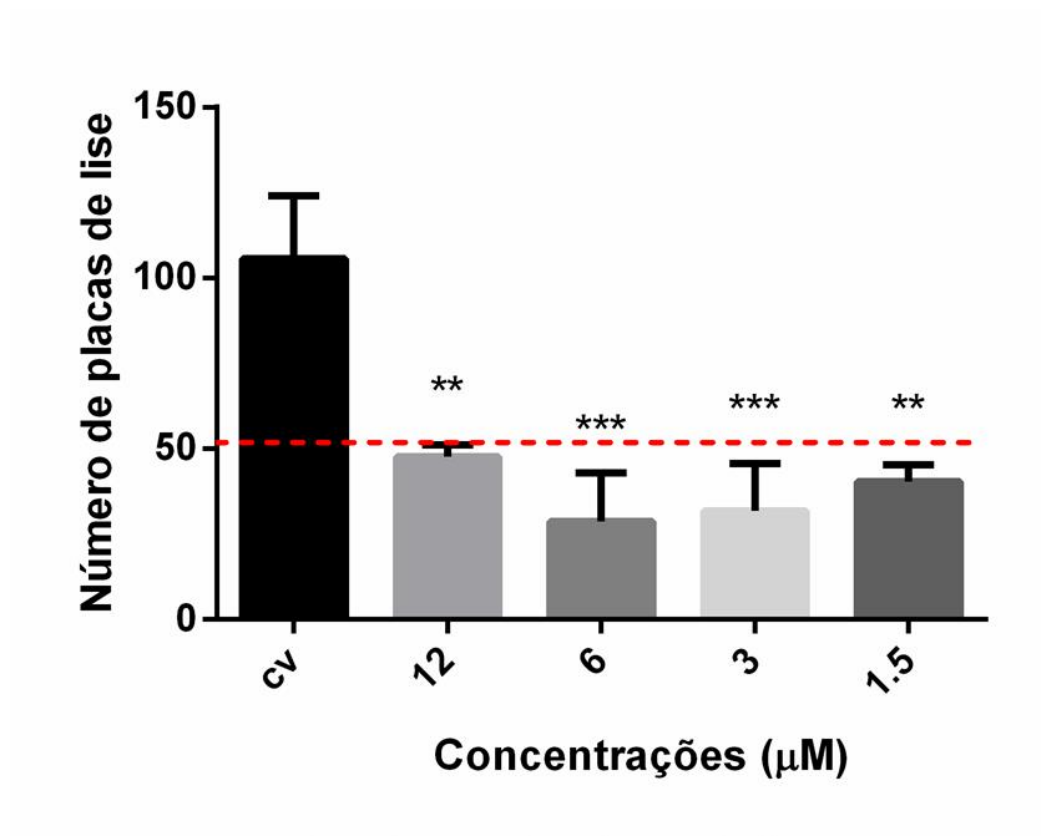
(A) Avaliação da atividade de Pós-tratamento do composto QPA nas concentrações de 1.5-24 μM em diferentes MOIs. Os dados representam a média de um experimento realizado em triplicata. Os dados foram analisados com auxílio do software Graphpad Prism 6 e os valores expressam média $\pm$ desvio padrão (One-way ANOVA, seguida pelo teste de Bonferroni, $*p \leq 0.05$). (B) Morfologia de células HEP-2 infectadas por hRSV (MOI 0.1, 0.5 e 1.0) com diferentes tratamentos de QPA (1.5-12 μM). As imagens foram avaliadas e documentadas em microscópio de luz branca (4X).



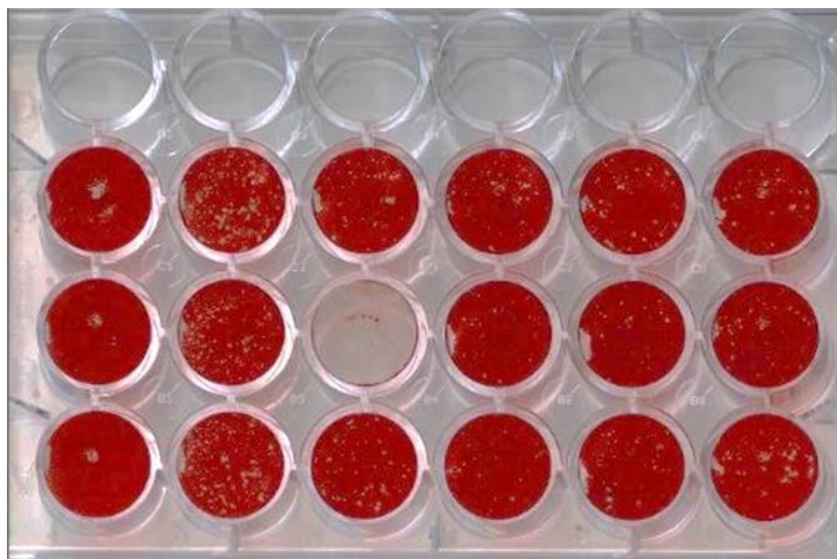
Efeito do ensaio de adesão do hRSV pela avaliação da % de proteção celular da QPA. Os dados representam a média de um experimento realizado em triplicata. Os dados foram analisados com auxílio do software Graphpad Prism 6 e os valores expressam média $\pm$ desvio padrão (One-way ANOVA, seguida pelo teste de Bonferroni, $*p \leq 0.05$).

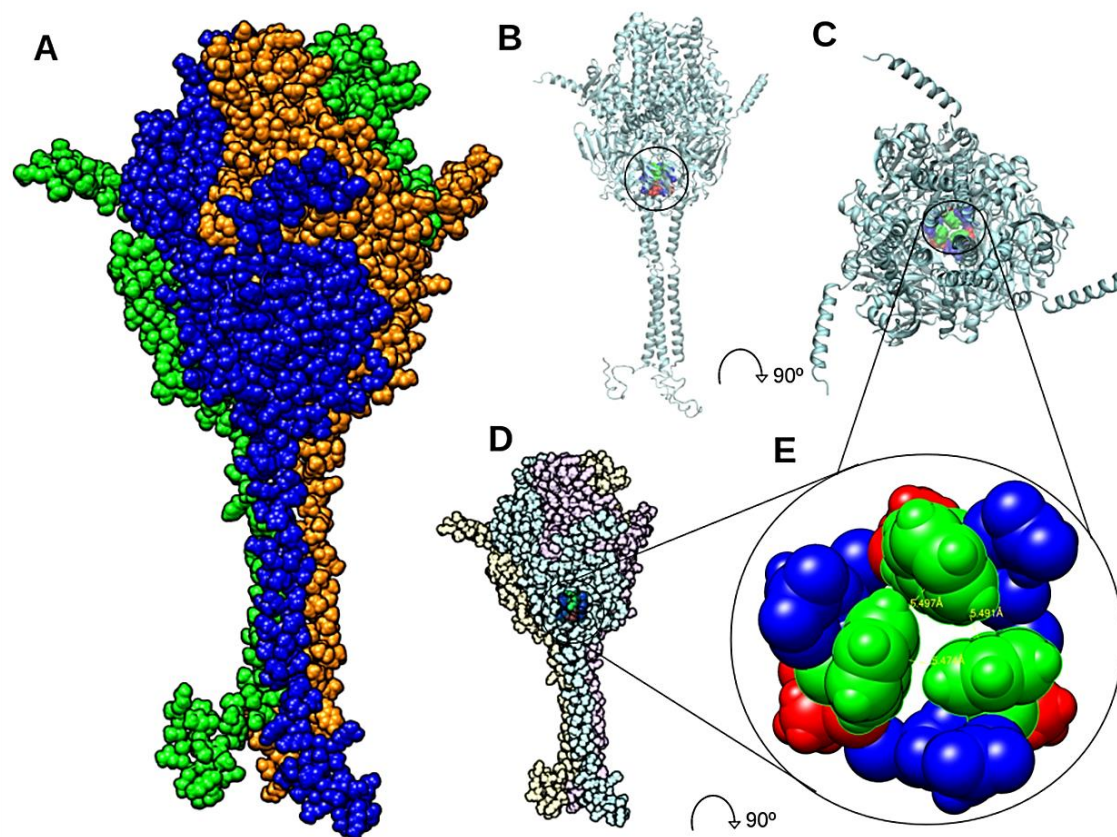


Avaliação do ensaio de Tempo de Adição do composto QPA nas concentrações de 3 μ M, 6 μ M, 12 μ M em diferentes intervalos de tempo (0-48hs). Os dados representam a média de um experimento realizado em triplicata. Os dados foram analisados com auxílio do software Graphpad Prism 6 e os valores expressam média $\pm$ desvio padrão (One-way ANOVA, seguida pelo teste de Bonferroni, * $p \leq 0.05$).

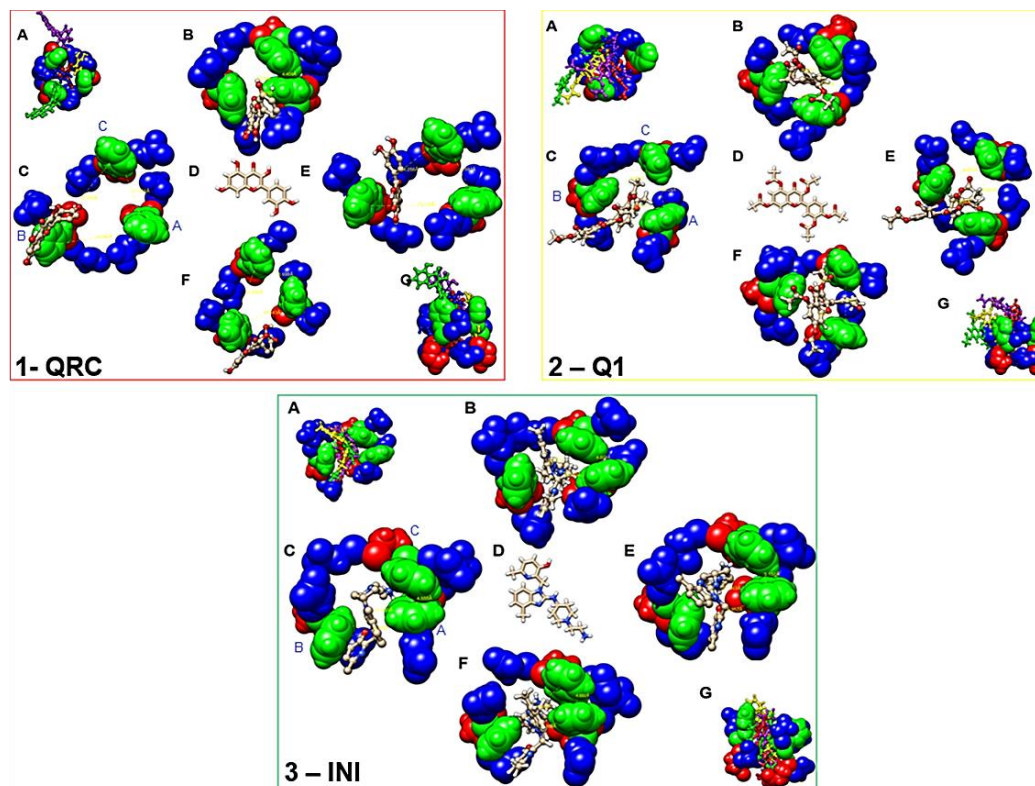


Contagem do número de Placas de lise no ensaio Virucida. A contagem foi realizada através do software MATLAB® (MATrix LABoratory).



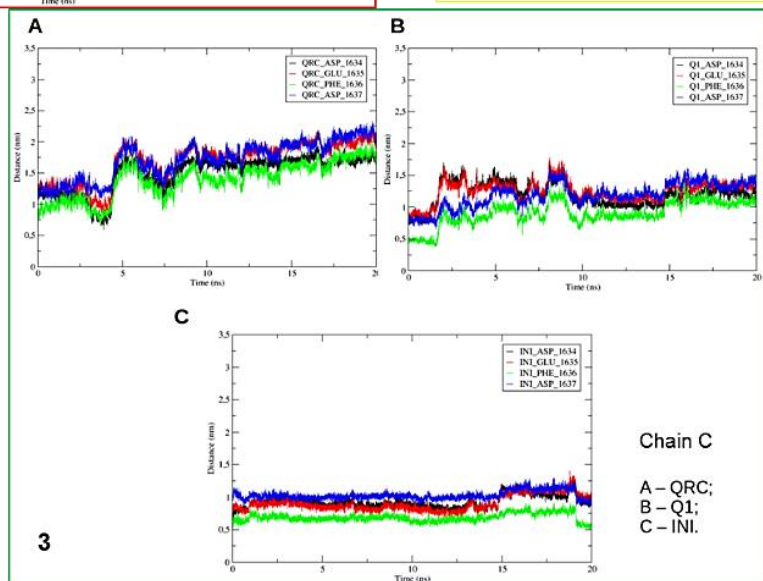
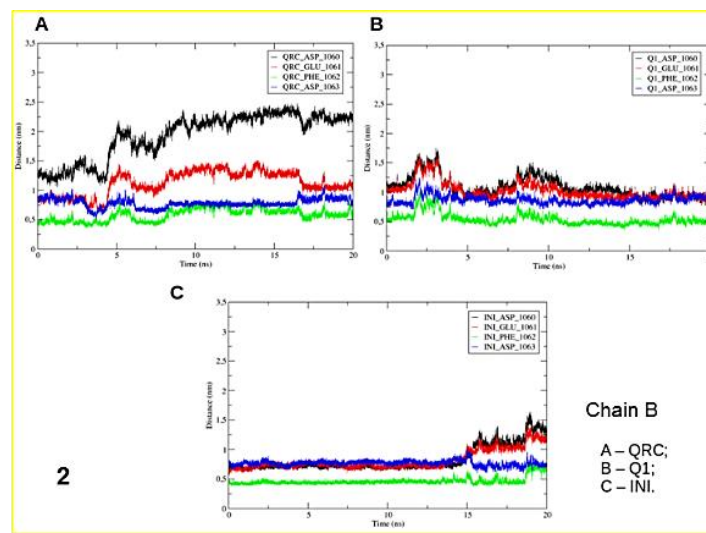
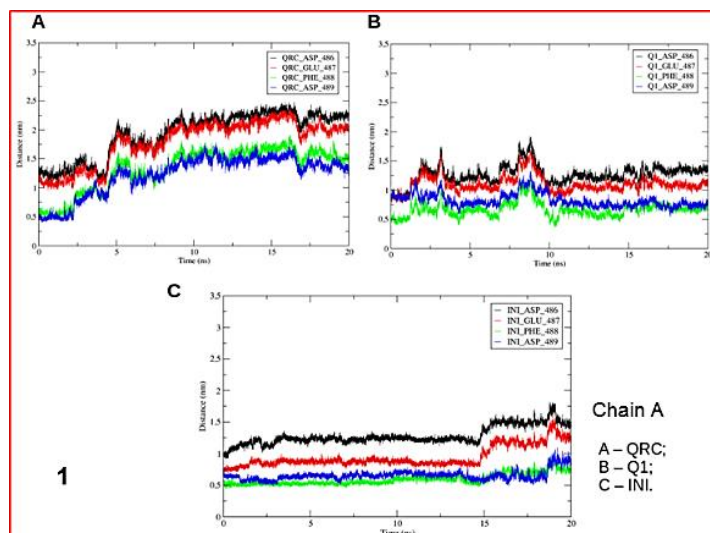


A) The modeled structure of RSV-F trimer (RSVmF), in spheres representation. B-D) Different angle and representation views of RSVmF, and the trissimetric highlighted pocket. (B) shows a frontal view of the protein in new cartoon representation. (C) corresponds to its top view. (D) exhibits a lateral view, in spheres representation. (E) emphasizes the negatively charged trissimetric pocket, isolated from RSVmF. The cavity (E) is composed of residues of phenylalanine (in green), aspartate (in red) and glutamate (in blue). The distances between the phenylalanines are displayed, and are equivalent to 5.474 Å between chain A (right below) and B (left below), 5.497 Å between B and C chains (above, in the middle), and 5.491 Å between C and A chains. VMD and Chimera were used for image edition.

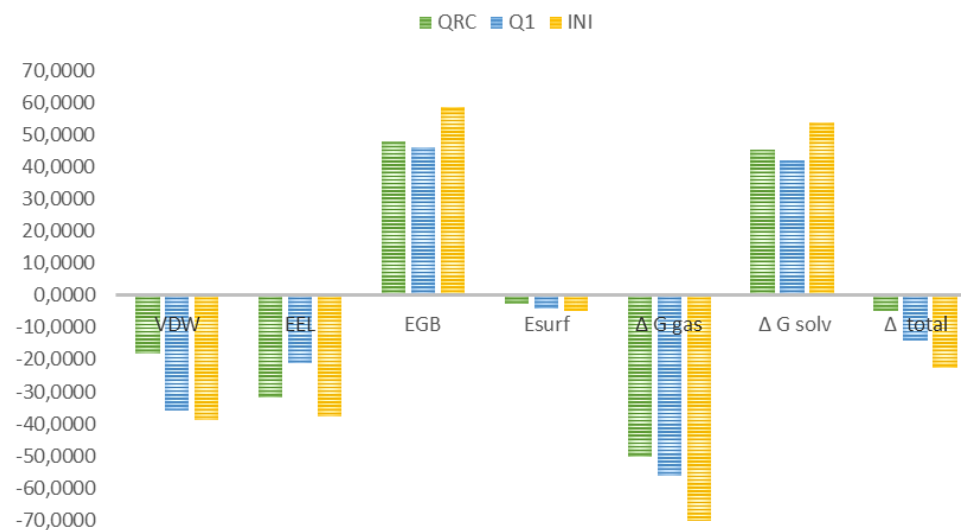


Molecular Dynamics analysis displays the behavior of 1-querctetin, 2-Q1, and 3-JNJ2408068 (reference inhibitor) along a trajectory simulation of 20 ns. The figure shows the three ligands and their most representative conformations within the central cavity of RSVmF. The cavity is a trissimetric pocket, composed of residues of phenylalanine (in green), aspartate (in red) and glutamate (in blue). For all the three ligands, both in (A) and (G), the cavity was taken as "static", without trajectory modifications, only to emphasize the movement of the ligand - under a superior and frontal view, respectively. All the first frames ($t = 1\text{ns}$) appear in (B). (C) shows the protein chains evinced by blue letters. In (D), the two-dimensional (2D) representation of each ligand is shown. The conformations exhibited in (C), (E) and (F) are related to cluster analysis of the trajectory. Thus, they represent the first three clusters. Therefore, (C) depicts the first most representative conformation, (E), the second, and (F), the third.

QRC = querctetin, Q1= querctetin pentacetate, INI= JNJ2408068.



Distance between the trissimetric pocket within the central cavity of RSVmF and quercetin, Q1 and JNJ-2408068 (reference inhibitor), along a trajectory simulation of 20 ns. In this analysis the chains were not renumbered. Therefore, residues 486-489 in chain A are equivalent to residues 1060-1063 in chain B and to 1634-1637 in chain C. They follow the same order: ASP-GLU-PHE-ASP. Chains A, B and C are represented in (1), (2) and (3), respectively. Quercetin distance is shown in (A), Q1 in (B), and JNJ-2408068 in (C). QRC=quercetin, Q1=quercetin pentacetate, INI= JNJ2408068.



The graph corresponds to the average energy interaction between RSVmF-ligands, obtained by the difference (complex - receptor - ligand). The types of energy involved in the interaction of RSVmF-Quercetin, RSVmF-Q1 and RSVmF-JNJ – 2408068 are shown.

VDW = Van der Waals interactions between protein-ligand;

EEL = electrostatic interactions between protein-ligand;

EGB = the electrostatic contribution to the free energy of solvation calculated by GB;

Esurf = the energy of the surface area - SA (energy of nonpolar solvation),

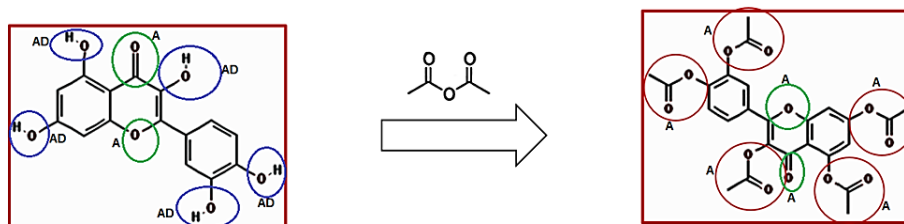
ΔG_{gas} = Gibbs free energy in the gas phase;

ΔG_{solv} = EGB + Esurf;

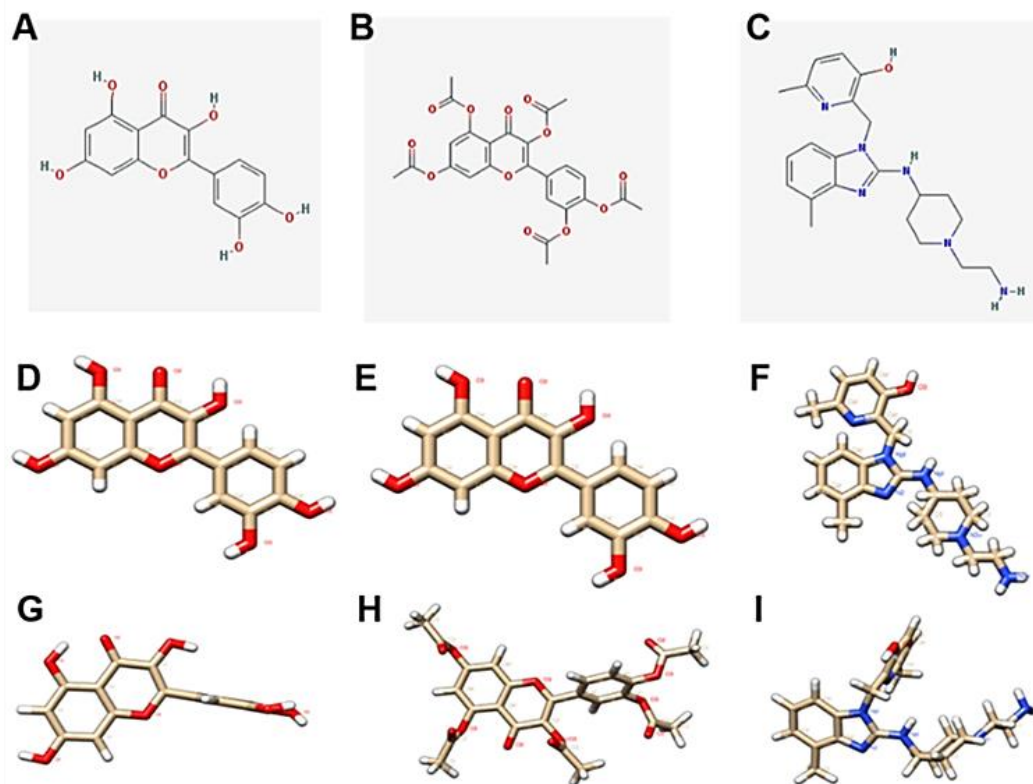
Δ_{total} = calculation result involving all of the above terms

Unit of all terms above = kcal / mol.

QRC=quercetin, Q1=quercetin pentacetate, INI= JNJ2408068

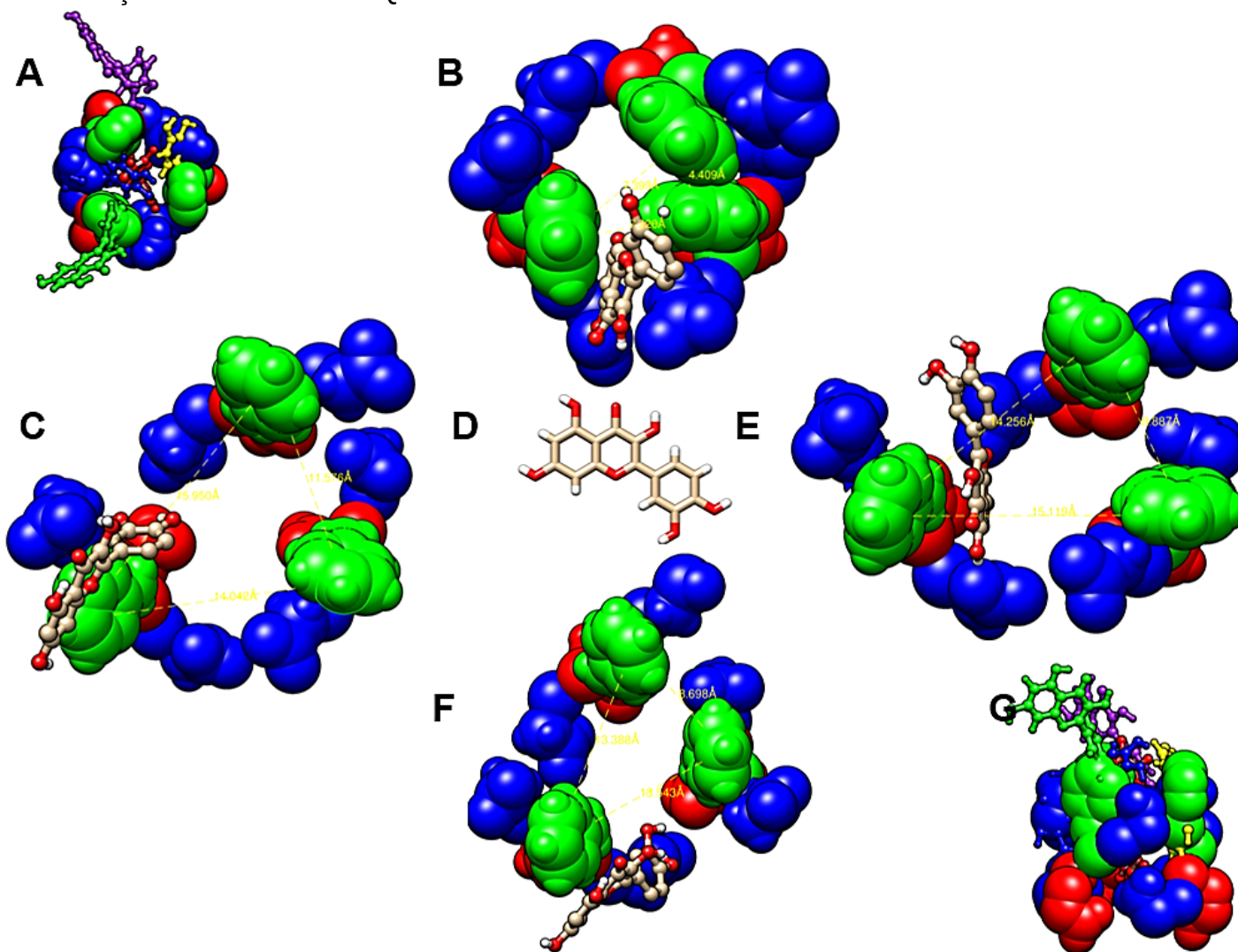


The figure illustrates the acetylation of quercetin, which is then converted into quercetin pentacetate (Q1).
A = acceptor of hydrogens; AD = acceptor/donor of hydrogens.

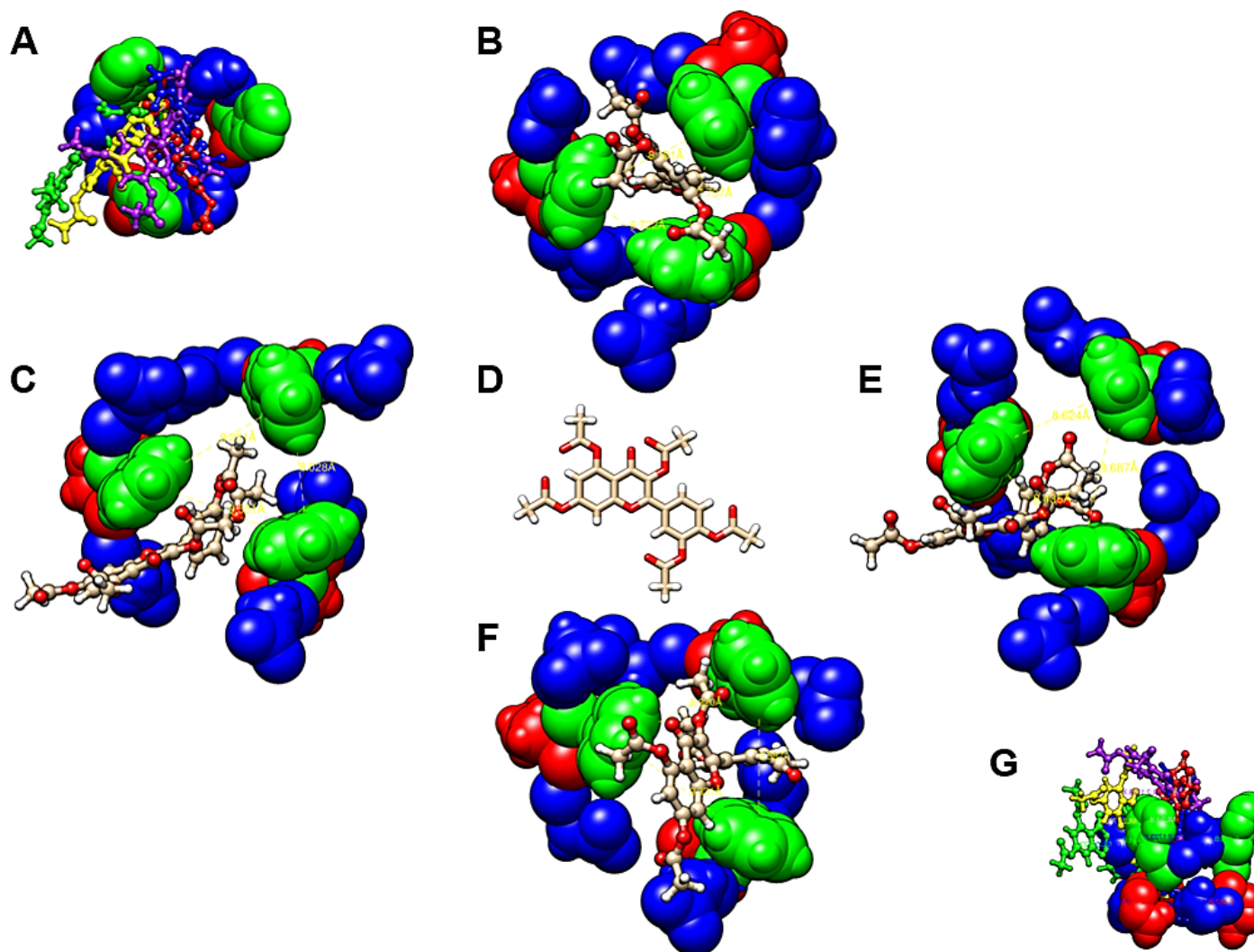


The figure illustrates the structures of quercetin, quercetin pentacetate (Q1) and JNJ2408068 (reference inhibitor). (A), (B), (C), (D), (E) and (F) show the 2D image of these ligands, whereas (G), (H) and (I) exhibit their 3D structure. Quercetin is represented in column (A, D and G). Q1 appears in (B, E, and H), and JNJ-2408068 is depicted in (C, F and I). All files were extracted from Pubchem® database, and the image edition was performed in Chimera. Nitrogen atoms appear in blue, Oxygens in red, Hydrogens in white and Carbons in light beige. (A), (D) e (G) = quercetin, (B), (E) e (H) = Q1; (C), (F) e (I) = JNJ2408068.

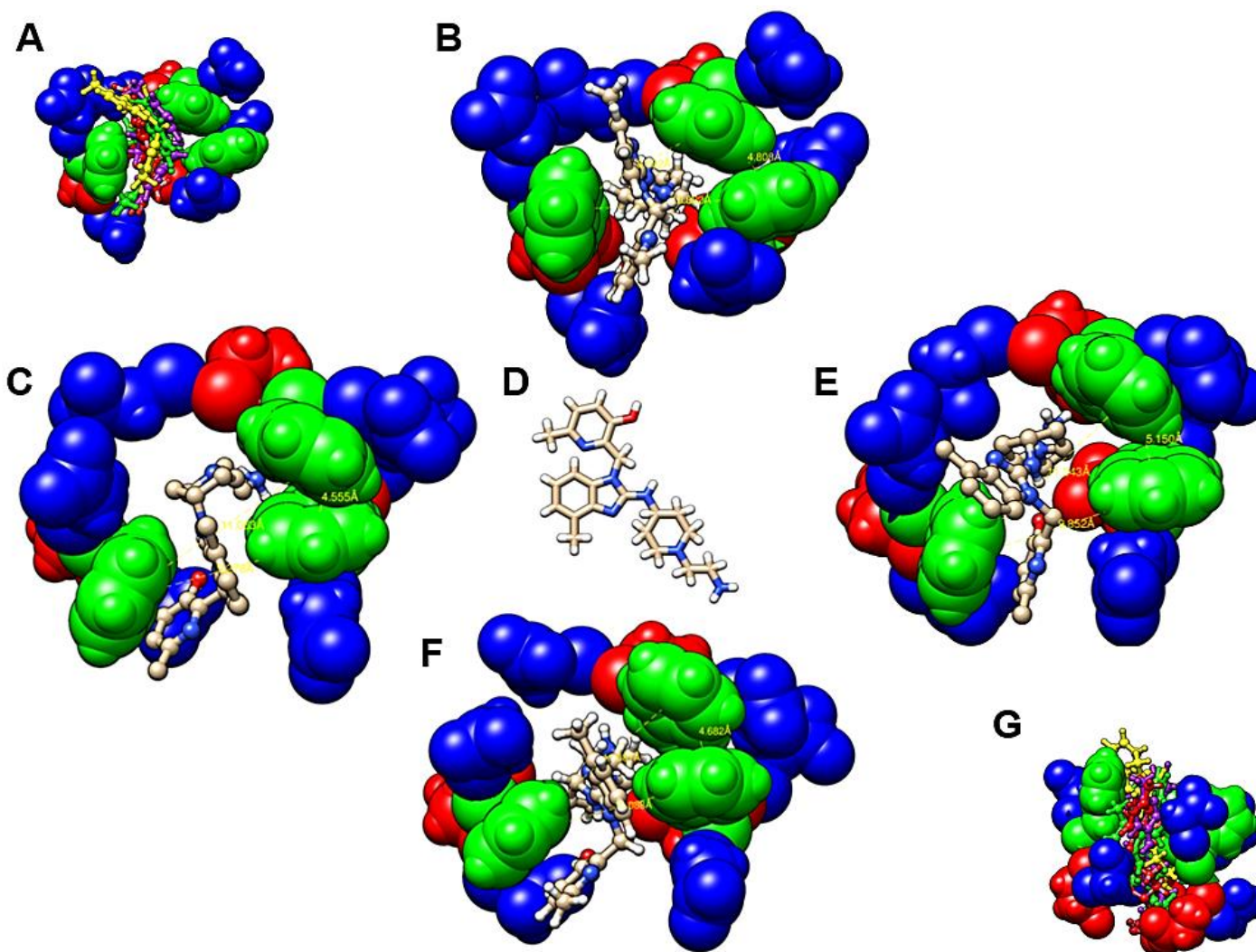
APÊNDICE II – Simulação de DM: RSVmF-QRC



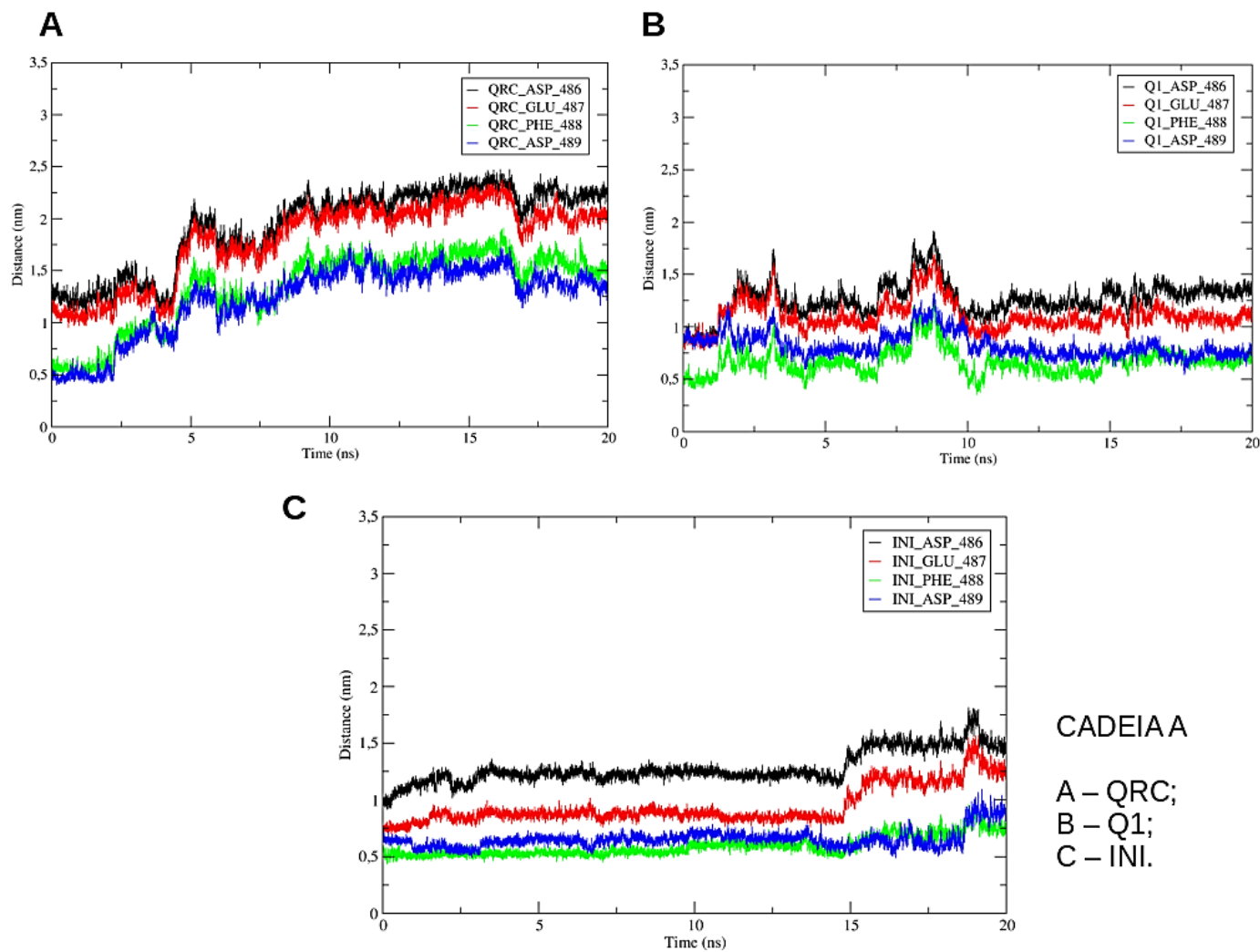
APÊNDICE III – Simulação de DM: RSVmF-Q1



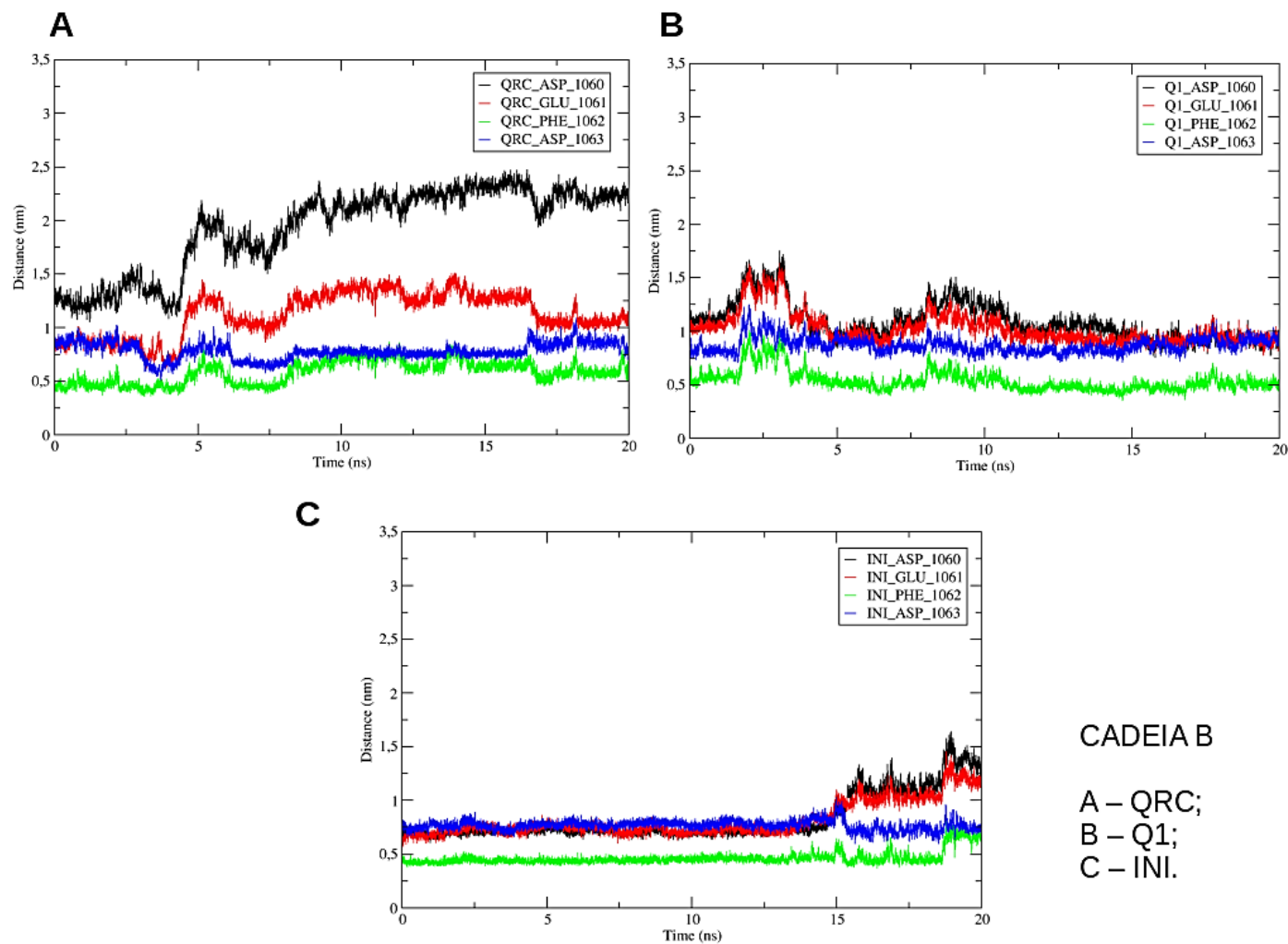
APÊNDICE IV – Simulação de DM: RSVmF-JNJ



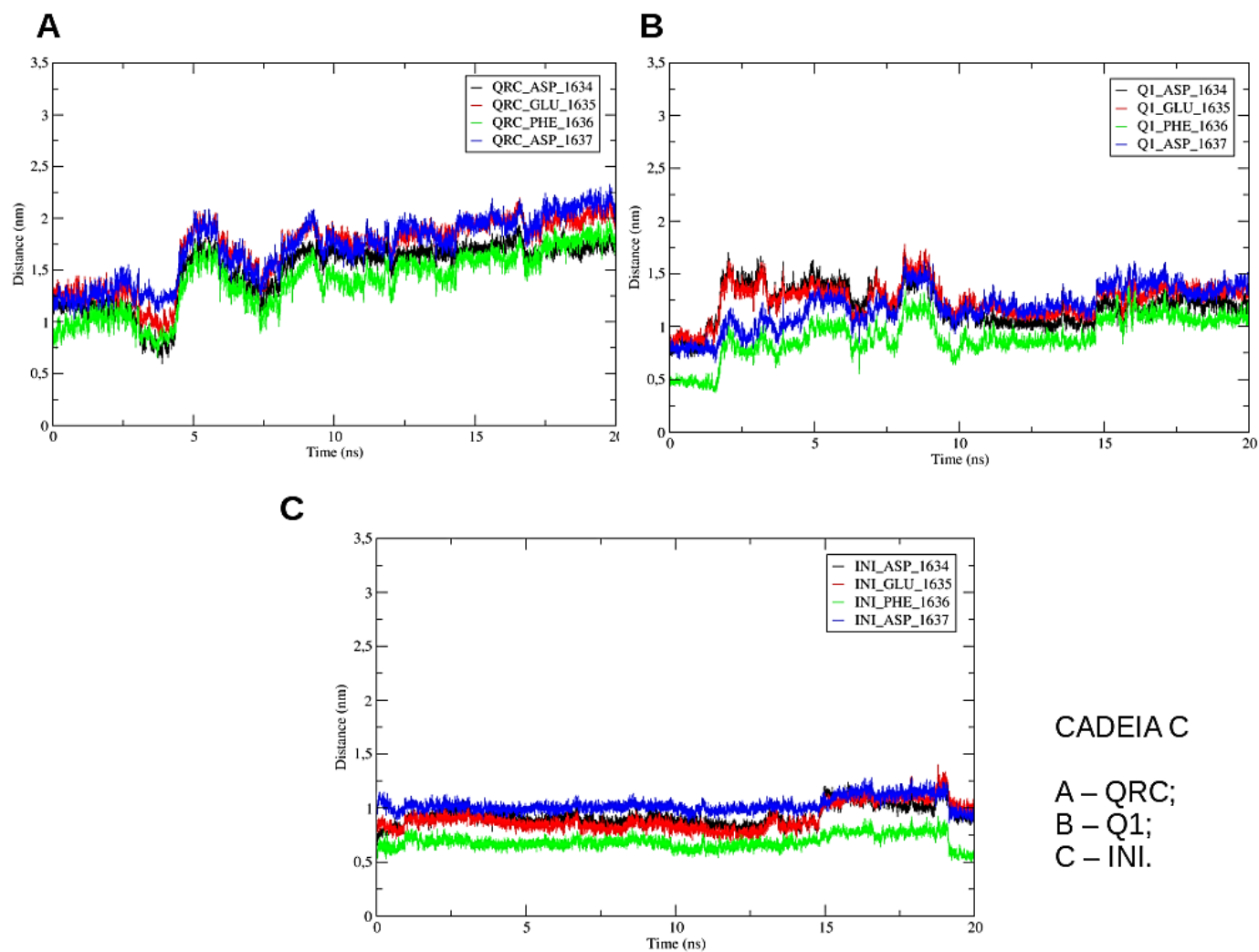
APÊNDICE V – Variação da distância RSVmF-ligantes – CADEIA A



APÊNDICE VI – Variação da distância RSVmF-ligantes – CADEIA B

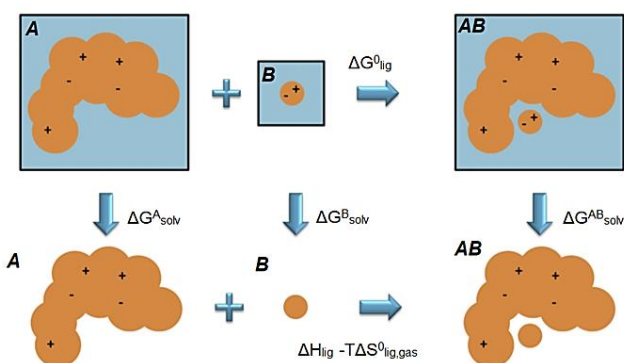


APÊNDICE VII – Variação da distância RSVmF-ligantes – CADEIA C

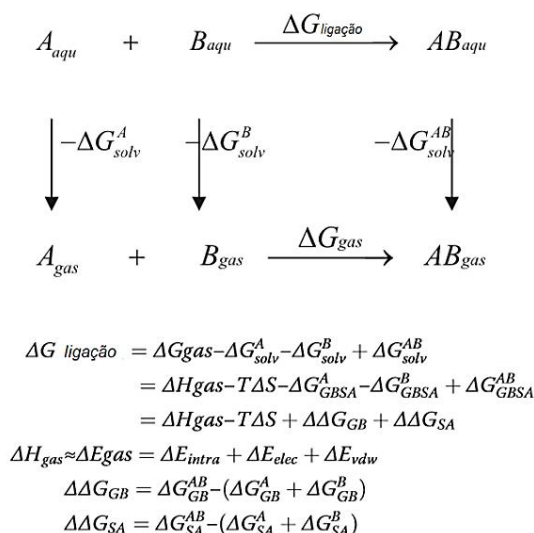


APÊNDICE VIII – Método MM-GBSA

1



2



O ciclo termodinâmico empregado no método de MM-GBSA (*Molecular Mechanics Generalized Born Surface Area* – Mecânica Molecular de Área de Superfície Generalizada de Born). O método MM-GBSA se utiliza de cálculos de aproximações de área e considera o soluto (proteína e ligante) como esferas (o termo *GB*), para os quais existe uma área acessível ao solvente (*SA*), e, por isso, a abreviação “*GBSA*”. O termo “*MM*” se refere à inclusão dos cálculos de interações intramoleculares, eletrostáticas (Coulomb) e de Van der Waals. O solvente, nessa abordagem, é considerado um “*continuum*”, o qual ocupa todo espaço disponível, isto é, o termo “*SA*” e apresenta uma constante dielétrica. Em (1), pode-se visualizar a representação de uma proteína (A) e de um ligante (B) hipotéticos, com suas conformações espaciais já aproximadas. Em um primeiro instante, em contato com solvente e sozinhos, para posteriormente serem combinados e formarem um complexo, imersos em solvente (parte superior). Ainda em 1, na parte inferior, exibem-se proteína e ligante sozinhos, e, em seguida, complexados, numa fase gasosa, isto é, sem solvente. Em (2), apresenta-se um resumo breve e simplificado dos cálculos realizados por essa abordagem. T= temperatura; S= entropia do soluto (proteína/ligante), calculada por Análise de Modos Normais, ΔG_{gas} é a energia de interação entre A e B na fase gasosa; ΔG_{solv}^A , ΔG_{solv}^B , ΔG_{solv}^{AB} são a energia livre de solvatação, calculado pelo método GBSA; $\Delta\Delta G_{GB}$ e $\Delta\Delta G_{SA}$ representam os termos eletrostáticos (interação entre o soluto o solvent *continuum*) e não-polares (forças de atração e repulsão de Vander Waals entre soluto e solvente), respectivamente. As energias de ligação, angular e de torção constituem o termo ΔE_{intra} do complexo (energia intramolecular), enquanto que ΔE_{elec} e ΔE_{vdw} correspondem às interações eletrostáticas e de Van der Waals entre proteína-ligante.

Fonte: 1 adaptado [327], 2 adaptado [328].

APÊNDICE IX – Propriedades físico-químicas dos ligantes quercetina, Q1 e JNJ-2408068

	<u>Quercetina</u>	<u>Q1</u>	<u>JNJ-2408068</u>
Propriedades computadas			
Nome Propriedade	Valor propriedade	Valor propriedade	Valor propriedade
Fórmula molecular	C ₁₅ H ₁₀ O ₇	C ₂₅ H ₂₀ O ₁₂	C ₂₂ H ₃₀ N ₆ O
Peso molecular	302,238 g/mol	512,423 g/mol	394,523 g/mol
Contagem doadores Hidrogênio	5	0	3
Contagem aceptores Hidrogênio	7	12	6
Contagem ligações rotativas	1	11	6
Complexidade	488	998	514
Área de superfície polar topológica	127 Å ²	158 Å ²	92,2 Å ²
Massa Monoisotópica	302,043 g/mol	512,095 g/mol	394,248 g/mol
Massa Exata	302,043 g/mol	512,095 g/mol	394,248 g/mol
XLogP3-AA	1,5	2,1	2,4
Carga Formal	0	0	0
Contagem de átomos pesados	22	37	29
Contagem Estereocentros Átomos Definidos	0	0	0
Contagem Estereocentros Átomos Indefinidos	0	0	0
Contagem Estereocentros Ligações Definidas	0	0	0
Contagem Estereocentros Ligações Indefinidas	0	0	0
Contagem Átomos Isotópicos	0	0	0
Contagem de unidade ligada por covalência	1	1	1

Tabela de propriedades físico-químicas da quercetina, do Q1 e do inibidor usado como controle (JNJ-2408068), sob número de identificação (*Compound Identification* – CID – no Pubchem) [5280343](#), [14005](#), [511013](#), respectivamente.

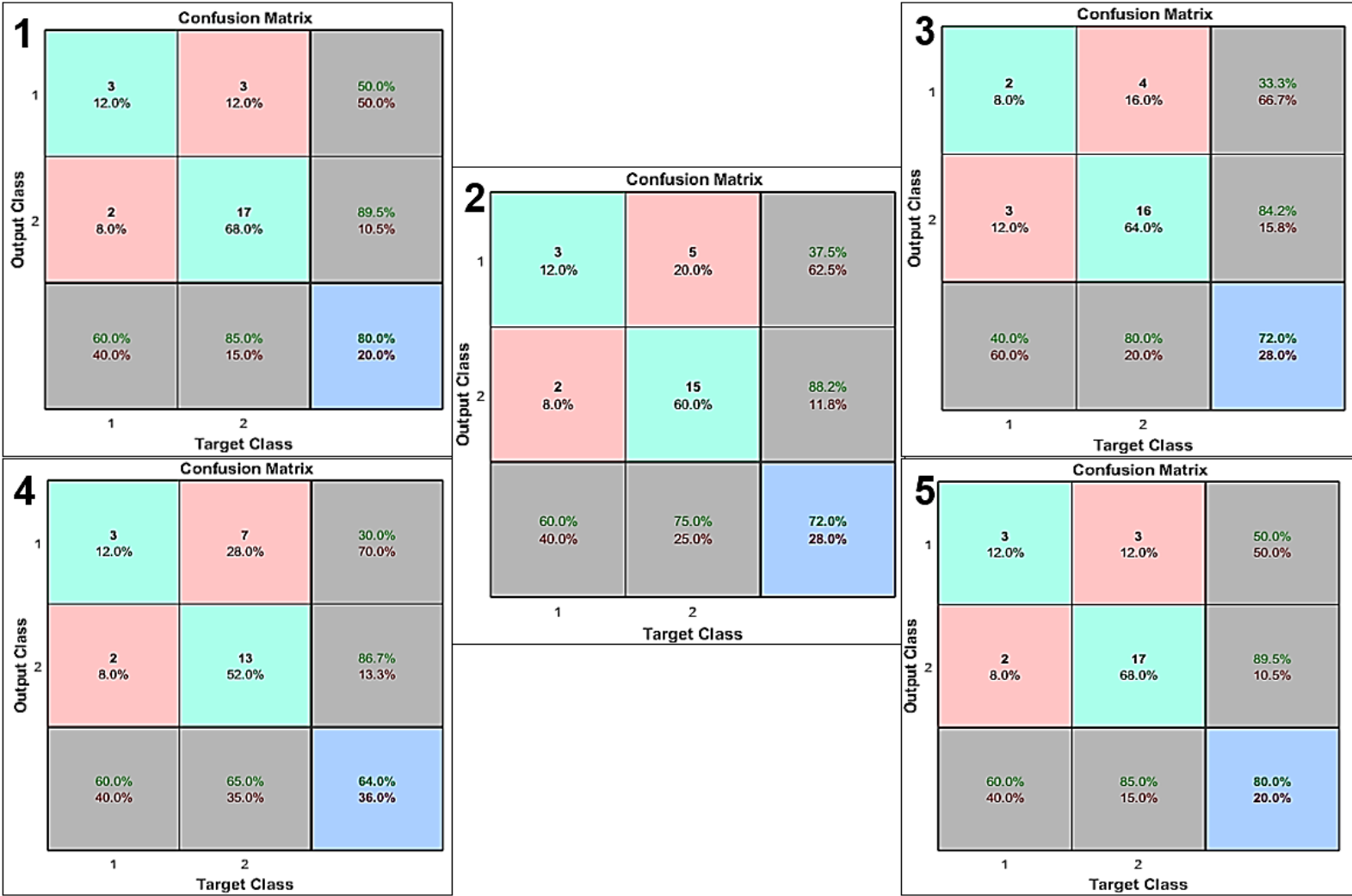
*Átomos pesados: número de átomos que não sejam hidrogênio na molécula.

* Massa exata ou massa monoisotópica é a massa calculada usando-se a massa do isótopo de ocorrência natural mais abundante, relativo à massa do C¹².

* XLogP3-AA: é o coeficiente de partição entre η -octanol e a água. Revela o caráter lipofílico da molécula, logo, é relacionada com a solubilidade. Assim, quanto maior o índice, maior a lipofilicidade/hidrofobicidade [\[325\]](#).

* Área de superfície topológica polar: É a soma da superfície de átomos polares, tais como Oxigênio, Nitrogênio e os Hidrogênios ligados a eles. Está relacionada com a permeabilidade da molécula [\[326\]](#).

APÊNDICE X – Primeira simulação ANN



APÊNDICE XI – Matrizes de confusão da simulação 2

