



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

Dissertação de Mestrado

HAROLDO DE LIMA PIMENTEL CRAVO

**Modelagem Molecular da Interação entre a Proteína
de Fusão do Vírus Sincicial Respiratório Humano e
Inibidores da Ação Viral**

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

São José do Rio Preto
2012



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

Dissertação de Mestrado

HAROLDO DE LIMA PIMENTEL CRAVO

**Modelagem Molecular da Interação entre a Proteína
de Fusão do Vírus Sincicial Respiratório Humano e
Inibidores da Ação Viral**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Biofísica Molecular da Universidade Estadual Paulista “Júlio de Mesquita Filho” - Unesp - Campus de São José do Rio Preto como parte dos requisitos para obtenção do título de Mestre.

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

São José do Rio Preto
2012



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

Dissertação de Mestrado

HAROLDO DE LIMA PIMENTEL CRAVO

**Modelagem Molecular da Interação entre a Proteína
de Fusão do Vírus Sincicial Respiratório Humano e
Inibidores da Ação Viral**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Biofísica Molecular da Universidade Estadual Paulista “Júlio de Mesquita Filho” - Unesp - Campus de São José do Rio Preto como parte dos requisitos para obtenção do título de Mestre.

Apoio financeiro: Este projeto recebeu auxílio financeiro das agencias FAPESP processo: 2010/50444-4/ CNPq processo: 306342/2010-3, Prope e da CAPES.

São José do Rio Preto
2012

HAROLDO DE LIMA PIMENTEL CRAVO

MODELAGEM MOLECULAR DA INTERAÇÃO ENTRE A PROTEÍNA DE FUSÃO DO VÍRUS SINCICIAL RESPIRATÓRIO HUMANO E INIBIDORES DA AÇÃO VIRAL

Dissertação apresentada para obtenção do título de Mestre em Biofísica Molecular, junto ao Programa de Pós-Graduação em Biofísica Molecular do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

BANCA EXAMINADORA

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

Prof^a. Dr^a. Karina Alves de Toledo
Professora Doutora
UNESP – Assis

Prof. Dr. José Roberto Ruggiero
Professor Doutor
UNESP – São José do Rio Preto

São José do Rio Preto, 27 de Janeiro de 2012

À minha família e amigos, o alicerce de todas as minhas construções.

AGRADECIMENTOS

Agradeço:

Aos meus pais e irmãos, por me ensinarem, pela simples convivência, como caminhar pelo bem de toda sociedade.

A Professora Doutora Fátima Pereira de Souza, pela paciência, dedicação e orientação.

Aos participantes da banca, o Professor Doutor José Roberto Ruggiero, professor, amigo e colega de coral e a Professora Doutora Karina Alves de Toledo, pela contribuição neste trabalho.

Ao Instituto de Biociências, Letras e Ciências Exatas – IBILCE/UNESP, Campus de São José do Rio Preto, por ser a casa do início da construção da minha vida acadêmica.

Ao Departamento de Física, pelo apoio e confiança, em especial aos servidores Ilva, Paulo, Barbosa, Marcelino e Tiago, que sempre oferecem total suporte às nossas necessidades.

As grandes amigas Mirian Elisa, Aline e Jéssica que proporcionaram atenção e novos fôlegos para chegar até o objetivo final desta jornada.

Aos amigos Vinicius, Débora, Ícaro, Rafael Musa, Luciane e Angélica, pela companhia e coleguismo nas matérias do curso e nos congressos.

As queridas amigas e parceiras Mariana e Gabriela, pelas ideias, discussões, conselhos e apoio para chegarmos nos melhores resultados.

Ao pessoal do laboratório e do grupo de Biologia Molecular, Deriane, Luciana e Thiago pelo trabalho em grupo.

Ao professor doutor Elso Drigo Filho pela longa amizade, repassando sua experiência como uma luz nos momentos de dificuldades.

Aos professores doutores Marcelo Andres Fossey e Marinonio Lopes Cornélio pelo interesse e apoio incondicional no fornecimento das melhores condições para o desenvolvimento deste trabalho.

A professora Zuleica de Carvalho Moreira, regente do Coral IBILCE e ao Coral IBILCE, pelos momentos de descontração e canto, que espantavam todas as angústias e males.

À CAPES, FAPESP, CNPQ e Prope, pela bolsa e apoio financeiro, essenciais para o melhor desenvolvimento deste trabalho.

Ao programa de Biofísica Molecular, Seção de Pós-Graduação e a Pró-Reitoria de Pós Graduação pela ajuda e esclarecimentos prestados ao longo deste trabalho.

"Não existe instruções para o uso da vida, cada um deve fazer sua própria escrita!"

RESUMO

O Vírus Sincicial Respiratório Humano (hRSV) foi identificado em 1957 e mesmo após vários anos de investigação, nenhuma vacina foi desenvolvida. Acredita-se que a chave de inibição da ação viral são suas glicoproteínas de membrana, em especial a proteína de fusão (F), que com auxílio da proteína de ligação (G), é responsável pela instalação do hRSV na célula hospedeira. Há evidências experimentais de que compostos como flavonóides e glicosaminoglicanos podem diminuir a infecção viral, sendo então a proteína F um bom alvo para a ação destes compostos. O presente estudo utilizou de ferramentas de bioinformática para verificar as possíveis regiões de interação da proteína F com a Heparina Sulfatada e Flavonóides. Os programas de bioinformática foram utilizados para: modelagem dos compostos, caracterização e previsão da estrutura secundária da proteína, modelagem da estrutura terciária e *docking* molecular entre o modelo da proteína F e as estruturas tridimensionais dos Flavonóides e da Heparina Sulfatada. Modelos válidos foram obtidos para as estruturas tridimensionais dos flavonóides e para o modelo completo da proteína F. As características da proteína incluem um alto nível de conservação na sequência de aminoácidos e, especialmente, em seus sítios de ligação. O *docking* da proteína com a Heparina, e o virtual *screening* da biblioteca de Flavonóides e a estrutura da proteína, resultaram em sítios de interação com grande potencial de inibição, uma vez que concordam com evidências experimentais descritos na literatura. A Heparina liga-se ao sítio de clivagem II, importante região para obtenção da atividade de fusão da proteína. Os Flavonóides podem se ligar a região hidrofóbica que desestabiliza a formação do feixe de hélices, estrutura importante para o mecanismo de fusão da proteína. Ambas as regiões possuem grande relevância nas mudanças conformacionais que a proteína F sofre para desempenhar sua função. O resultado do *docking* mostrou que as interações moleculares são prováveis de ocorrer e indicou quais compostos são melhores candidatos a um possível inibidor. Essas evidências sugerem que ambos os compostos podem ser utilizados em estudos experimentais, para um melhor esclarecimento de suas ações sobre a infectividade e ação do vírus.

Palavras chaves: Vírus Sincicial Respiratório Humano, *Docking* Molecular, Modelagem de Proteínas, Proteína de Fusão.

ABSTRACT

Human Respiratory Syncytial Virus (hRSV) was identified in 1957 and even after several years of research, no vaccine has been developed yet. It is believed that the key to the inhibition of viral action is its membrane glycoproteins, including the Fusion Protein (F), responsible for the installation of the hRSV in the host cell. There are evidences that compounds such as flavonoids and glycosaminoglycans can decrease the viral infection, and F protein can be a good target for the action of these compounds. The present study checked the possible sites of interaction between F protein and heparin and flavonoids, using computational tools. Bioinformatics programs were used for: modeling compounds, characterization and prediction of protein secondary structure, tertiary structure modeling and the docking between the protein model and the structures of flavonoids and sulfated heparin. Valid models were obtained for flavonoids structures and the complete model of F protein. The characteristics of the protein include a high level of conservation in amino acid sequence and especially in its binding sites. The heparin docking and virtual screening of flavonoids resulted in interaction sites with great potential for inhibition, since they agree with other studies and experimental evidence of F protein inhibition. This study shows that compounds such as sulfated heparin and flavonoids interact in important sites of F protein. Heparin binds to the cleavage site II and flavonoids can bind to the hydrophobic site that destabilizes the formation of the six-helix-bundle region. Both regions are important for conformational changes that F protein undergoes to get its fusion activity. Docking showed that molecular interactions are likely to occur and selected the best candidates for a possible inhibitor. These evidences show that both compounds can foster further studies to discover effective inhibitors for fusion protein.

Key words: Respiratory Syncytial Virus, Molecular Docking, Protein Modeling, Fusion Protein.

Lista de Figuras

Figura 1 – Representação esquemática do Vírus Sincicial Respiratório (RSV)	19
Figura 2 – Esquema da organização genômica do RNA do RSV	19
Figura 3 – Representação da unidade dissacarídica da Heparina	24
Figura 4 – Estrutura geral dos Flavonóides	26
Figura 5 – Alinhamento das sequências selecionadas do GenBank.....	40
Figura 6 – Potencial região de interação da proteína F com a Heparina	42
Figura 7 – Esquema do genoma da proteína F do vírus hRSV	43
Figura 8 – Topologia da proteína F	47
Figura 9 – Modelo da proteína F	48
Figura 10 – Gráfico de Ramachandran para o modelo final da proteína F.....	50
Figura 11 – Comparação das estruturas modeladas dos flavonóides com as estruturas experimentais.....	52
Figura 12 – Modelo de interação entre a Heparina e a proteína de fusão.....	53
Figura 13 – Predição de sítios de interação da proteína F	54
Gráfico 14 – Histograma da quantidade de flavonóides por energia de interação	55
Figura 15 – Modelo da interação do flavonóide com menor energia de interação e a proteína de fusão.....	56
Figura 16 – Modelo a interação dos flavonóides no local da interrupção da atividade da proteína de fusão F.....	57
Figura 17 – Modelo da Proteína de Fusão por homologia	69
Figura 18 – Modelo da Proteína de Fusão por <i>Ab Initio</i>	70

Lista de Tabelas

Tabela 1 – Quadro de proporcionalidade das Estruturas Secundárias da Proteína F, de acordo com a análise do programa PREDICTPROTEIN.	44
Tabela 2 – Sítios de fosforilação da Proteína F separados pelo tipo de enzima que executa o processo, segundo a análise do programa PREDICTPROTEIN.....	46
Tabela 3 – Quadro de proporcionalidade das Estruturas Secundárias da Proteína F, de acordo com a análise do programa PDBSum.	49
Tabela 4 – Valores do RMSD da estrutura experimental de alguns flavonóides e a estrutura modelada pelo programa FROG2.....	51

SUMÁRIO

RESUMO	7
ABSTRACT	8
1. INTRODUÇÃO	13
1.1. Um breve histórico	13
1.2. hRSV: Epidemiologia, Patogênese e Profilaxia	14
1.3. Caracterização do Vírus Sincicial Respiratório Humano (hRSV)	18
1.4. Proteína F	21
1.5. Glicosaminoglicanos Sulfatados	23
1.5.1. características da Heparina Sulfatada	24
1.6. Flavonóides	25
1.7. Bioinformática	27
2. OBJETIVOS	29
3. METODOLOGIA	30
3.1. Construção do Banco de Dados de Sequências Primárias da Proteína de F.....	30
3.2. Análises das sequências	30
3.3. Modelagem da Proteína de Fusão	31
3.4. Otimização e Validação dos modelos.....	31
3.5. Busca por sítios de interação.....	34
3.5.1 Sítio de interação da Proteína de Fusão e a Heparina Sulfatada.....	34
3.5.2. Sítio de interação da Proteína de Fusão e os Flavonóides.....	35
3.6. <i>Docking</i> Molecular	36
3.6.1. <i>Docking</i> Molecular da Proteína de Fusão e a Heparina Sulfatada.....	36
3.6.2. Modelagem tridimensional dos Flavonóides.....	37
3.6.3. <i>Virtual Screening: Docking</i> Cego e <i>Docking</i> Dirigido	38
4. RESULTADOS E DISCUSSÕES	40
4.1. Análises da Estrutura Primária da Proteína de Fusão.....	40
4.1.1. Alinhamento das sequências e busca por motivos sequenciais de interação com Heparina Sulfatada.....	40
4.1.2. Predição da Estrutura Secundária, modificações pós-traducionais e características da proteína de Fusão	44
4.2. Análise, otimização e validação do modelo tridimensional da proteína F	47
4.3. Análise das Estruturas Tridimensionais dos Flavonóides	51
4.4. <i>Docking</i> Molecular do Modelo da Proteína F com a Heparina.....	52
4.5. <i>Docking</i> Molecular entre a Proteína F e Flavonóides	53
4.5.1. Determinação de Sítios de Interação na Proteína de Fusão	53
4.5.2. Interação entre o Modelo da Proteína de Fusão e Flavonóides: <i>Docking</i> Cego	54

4.5.3. Interação entre o Modelo da Proteína de Fusão e Flavonóides: <i>Docking</i> Dirigido	56
5. CONCLUSÕES	59
6. REFERÊNCIAS BIBLIOGRÁFICAS	60
7. ANEXOS	69
ANEXO A - Modelo da Proteína de Fusão gerado por homologia pelo programa MODELLER	69
ANEXO B - Modelo da Proteína de Fusão gerado por Ab Initio pelo programa ROSETTA	70

1 - Introdução

O vírus sincicial respiratório humano (hRSV, do inglês *human respiratory syncytial virus*) é um dos principais patógenos respiratórios, responsável pelas altas taxas de hospitalização de recém-nascidos, morbidade em crianças e adultos e excessiva mortalidade em idosos. Sua frequência aumenta todos os anos durante os meses do inverno, requisitando assim uma maior atenção das autoridades de saúde (AVEDAÑO et al., 2002). Os sintomas causados pelo vírus podem ser leves, como tosse e febre, mas ainda podem ser mais severos, como bronqueolite, bronquite e pneumonia viral. O hRSV já é um dos maiores problemas de saúde pública do mundo, uma vez que tem como consequência anual milhões de mortes e de gastos financeiros em cuidados com pacientes hospitalizados (WEISMAN & GROOTHIUS, 2000).

O vírus tem sua epidemiologia e patofisiologia bem descritas, contudo, os esforços para o desenvolvimento de fármacos eficientes ou de uma vacina efetiva ainda não foram bem sucedidos (BLACK, 2003). Embora existam alguns tratamentos contra a ação do hRSV, sua variabilidade antigênica e as características do sistema imune do hospedeiro dificultam a ação destes tratamentos (PARVEEN et al., 2006). As formas de combate da infecção viral causada pelo hRSV são pouco eficientes e de alto custo, tornando desta forma o uso inviável, principalmente nos sistemas públicos de saúde.

1.1 - Um breve histórico

Apesar de ter sido isolado há mais de 50 anos, acredita-se que os primeiros relatos de doenças provocadas pelo vírus datam de 160 anos antes de sua identificação (GROOTHIUS, HOOPES & HEMMING, 2010). Em Outubro de 1955, pesquisadores do *Walter Reed Army do Institute of Research* isolaram um vírus de secreções nasais de chimpanzés que apresentavam sintomas de doenças respiratórias. O agente isolado foi nomeado *Chimpanzee Coryza Agent* (CCA) e posteriormente, também foi identificado nos pesquisadores que mantiveram contato com os chimpanzés, mostrando que o vírus podia infectar humanos (WRIGHT & PIEDIMONTE, 2011).

No ano seguinte (1956), Chanock isolou o CCA de duas crianças, uma com bronqueolite e outra com pneumonia. O vírus foi replicado em cultura de células e percebeu-se que estas formavam um conjunto fusionado citoplasmático multinucleado, caracterizando desta forma um sincício. Baseado nestes resultados, Chanock propôs uma nova nomenclatura para o vírus, nomeando-o como vírus sincicial respiratório (WRIGHT & PIEDIMONTE, 2011).

Nos anos posteriores, Chanock e colaboradores continuaram seus estudos sobre a doença descrevendo a sua epidemiologia, tipos de manifestações clínicas e formas de recorrências da infecção. Durante este mesmo período foi testada uma vacina do vírus inativado. Porém, crianças que receberam a vacina desenvolveram a doença na mesma proporção das que não tinham sido vacinadas. Além disso, 80% das crianças vacinadas foram hospitalizadas com bronquite ou pneumonia severa, enquanto que apenas 5% das crianças não vacinadas apresentaram este quadro. Este resultado demonstrou uma significativa preocupação com o posterior desenvolvimento de novas vacinas e até os dias atuais, ainda não existe vacina disponível (WRIGHT & PIEDIMONTE, 2011).

Sabe-se hoje que o vírus sincicial respiratório humano é o principal causador de infecções respiratórias agudas (IRAs), que são responsáveis por uma grande demanda de serviços de saúde pública. As IRAs representam um problema de saúde prioritário a nível mundial, uma vez que age diretamente na morbidade e mortalidade infantil (AVEDAÑO, 2002). No Brasil, elas representam mais de 50% dos atendimentos em centros de saúde e cerca de um quinto das internações hospitalares de crianças no Sistema Único de Saúde (SUS), gerando altos custos (FAÇANHA & PINHEIRO, 2003). Os vírus são os principais agentes etiológicos das IRAs em crianças e são responsáveis por 50% a 90% destas infecções (MANOHA et al., 2008).

1.2 – hRSV: Epidemiologia, Patogênese e Profilaxia

Estudos realizados em 2010 pela Organização Mundial de Saúde (OMS) indicam que o hRSV é responsável em todo o mundo por mais de 60% das doenças do trato respiratório inferior em crianças, sendo 80% destas menores de um ano (WRIGHT & PIEDIMONTE, 2011). Aproximadamente de 0,5% a 2,0% destas crianças são hospitalizadas e dentre elas, de 50% a 90% tiveram bronquite e de 5% a 40% tiveram pneumonia (PAES et al., 2011).

Somente nos Estados Unidos, onde a taxa de infecção é baixa em relação a outros locais do mundo, o custo médio de hospitalizações de crianças infectadas pelo hRSV chega a 394 milhões de Dólares anualmente. Mais de meio milhão de crianças com idade abaixo dos 5 anos são tratadas na emergência de hospitais e mais de um milhão e meio de crianças passam pelos consultórios médicos com sintomas causados pelo vírus. Assim, estima-se que o custo econômico anual com hospitalizações ocasionadas pelo vírus chega a 500 milhões de Dólares, ou aproximadamente 800 milhões de Reais (WELLIVER, 2010).

Das crianças que contraem o vírus pela primeira vez, aproximadamente 40% desenvolve algum tipo de doença no trato respiratório inferior e pelo menos 1% é

hospitalizada por no mínimo 3 dias. Destas, 11% necessitam de tratamento intensivo e 8% são entubadas e respiram por ventilação mecânica (GREEN, 1989). Vários fatores estão relacionados ao risco de uma infecção mais severa em crianças. Dentre eles, os mais importantes são: o nascimento prematuro, crianças com doença congênita do coração e crianças imunocomprometidas, o que inclui as que tratam de câncer por quimioterapia, crianças recém transplantadas e as infectadas com o vírus HIV (WELLIVER, 2010).

A taxa de morbidade e mortalidade em crianças destes grupos são maiores devido a complicações causadas pelo vírus, com taxas de mortalidade de 1% a 4%. Em um estudo conduzido entre 1999 e 2007 com pacientes da Unidade de Tratamento Intensivo (UTI) de um hospital do Reino Unido, 8,6% foram a óbito devido ao vírus, caracterizando, assim, a infecção pelo hRSV como a maior causa de mortalidade infantil (PAES et al., 2011). Essa taxa é ainda maior em países em desenvolvimento, um estudo de 2005 mostrou uma estimativa de 66.000 a 199.000 mortes de crianças menores de 5 anos (NAIR et al., 2010). Em todo o mundo, a infecção pelo hRSV resulta em mais de um milhão de mortes de crianças por ano, o que representa um valor dez vezes maior que a taxa de mortalidade pelo vírus influenza (WRIGHT & PIEDIMONTE, 2011).

Fatores externos também se mostraram características importantes para o risco de uma infecção grave. O sexo da criança é um deles em que, apesar da taxa de meninas e meninos que são acometidos pelo vírus ser igual, a taxa de hospitalização de meninos é aproximadamente duas vezes maior que de meninas, indicando que a doença é mais grave para garotos. Crianças com famílias de baixa renda tendem a sofrer infecções mais severas, normalmente porque frequentam creches e são infectadas nos primeiros anos de vida. Crianças fumantes passivas e que estão expostas à poluição do ar também representam um grupo que sofre com infecções mais graves causadas pelo vírus (BLACK, 2003).

Além de crianças, foi reconhecido que o vírus possui uma significativa causa de morbidade e mortalidade em adultos e idosos, com um impacto que se aproxima ao do vírus da gripe. A média anual de incidência do hRSV na população adulta é de 5,5%, valor duas vezes maior que o da influenza. O vírus acomete de 3% a 7% de adultos saudáveis, enquanto que em grupos de alto risco a incidência aumenta para 10% (FALSEY et al., 2005). O hRSV é reconhecido como o maior causador de pneumonia e doenças pulmonares crônicas, apesar de normalmente ser confundido com uma gripe comum, já que ambos apresentam gravidade e mortalidade parecidas (DOWEL et al., 1996).

Alguns estudos já sugerem que o impacto clínico do hRSV em certas populações de adultos pode se assemelhar a pandemia provocada pelo vírus influenza. Os que aparecem

como grupos de risco são adultos com doenças cardiopulmonares, idosos frágeis, que vivem em asilos ou sozinhos e pessoas severamente imunocomprometidas, como pacientes com câncer, recém transplantados e pessoas infectadas pelo vírus HIV. Entre as pessoas destes grupos, o hRSV figura como o principal agente causador de pneumonia, com uma taxa de 10% a 20% por ano e representa 2% a 5% da taxa de mortalidade (FALSEY, 2000).

O período de infecção do vírus apresenta uma incidência sazonal, sendo este o principal fator que dificulta um estudo global destas infecções. Em locais de clima temperado, como o hemisfério norte, surtos anuais ocorrem no inverno ou no início da primavera. Em regiões com o clima tropical ou subtropical, como no Brasil, foi identificado um padrão bianual dos picos de epidemia do hRSV. Estas epidemias aparecem de forma regular, mas com diferentes padrões de sazonalidade. Áreas tropicais acima da linha do equador parecem estar associadas com a diminuição da temperatura e o aumento da umidade, enquanto que as áreas ao sul do equador têm uma sazonalidade associada com a diminuição da temperatura e da umidade (STENSBALLE et al., 2003).

Apesar de ser este o padrão identificado em países tropicais, a sazonalidade sofre influências da densidade demográfica, do ambiente de estudo e do método de detecção do vírus (STENSBALLE et al., 2003). No caso do Brasil, devido a sua grande extensão territorial, seu clima apresenta uma enorme variedade, o que reflete diretamente nos períodos de maior risco de infecção (CINTRA et al., 2001). Enquanto a região norte é marcada pelas chuvas, a variação da temperatura possui um papel importante nas regiões sul e sudeste. Desta forma Fortaleza, que possui um clima tipicamente tropical, apresentou epidemias do vírus distribuídas ao longo do ano, com prevalência nos meses de Abril, Maio e Junho (MOURA et al., 2001).

Na cidade de São Paulo os picos da epidemia foram observados no outono e no inverno, com incidência maior no mês de Maio. No entanto, verificou-se que a sazonalidade do vírus não está diretamente associada ao período de chuvas, nem ao período de meses frios, característicos destas estações. Na verdade os picos de infecção coincidiram com o aumento dos níveis de poluição do ar observados na cidade de São Paulo, o que sugeriu uma maior relação entre as condições ambientais e as taxas de infecções ocasionadas pelo vírus (VIEIRA et al., 2001).

A disseminação do hRSV ocorre pelo contato com secreções nasais de pessoas infectadas ou superfícies contaminadas. O vírus permanece viável em superfícies duras, como bancadas, por até 6 horas, em luvas de borracha por 90 minutos, e na pele por 20 minutos (WRIGHT & PIEDIMONTE, 2011). A permanência do vírus em qualquer superfície durante

horas explica a elevada taxa de infecções por hRSV, principalmente em crianças que frequentam creches, enfermarias pediátricas e berçários (LOURENÇÃO et al., 2005). O período de incubação do vírus é por volta de 2 a 8 dias, porém, a infecção viral pode ter um período de transmissão de até 3 semanas. Em pacientes imunocomprometidos a eliminação do vírus é significativamente prolongada e pode continuar por vários meses (WRIGHT & PIEDIMONTE, 2011).

O primeiro passo para a infecção é a entrada do vírus na célula, seguida da replicação do material genético no citoplasma, com liberação do vírus por brotamento e conseqüente fusão com as células adjacentes. O vírus penetra o corpo humano por meio das mucosas dos olhos, boca ou nariz e em seguida atinge a mucosa respiratória (DUDAS & KARRON, 1998). A infecção ocorre tanto no trato respiratório superior quanto no inferior. Os sintomas das infecções do trato respiratório superior são caracterizados por tosse, dificuldade respiratória, coriza, dispnéia e febre. Infecções que atingem o trato respiratório inferior apresentam geralmente quadro clínico mais grave, sendo a pneumonia e a bronquiolite doenças comumente associadas (WU et al., 2007). A primeira infecção causada pelo vírus normalmente é mais grave e ocorre durante os dois primeiros anos de vida, contudo, a resposta imune no hospedeiro não é eficiente, sendo normal a recidiva da infecção com os sintomas atenuados (BROOKS et al., 2000).

No vírus são reconhecidos subgrupos antigênicos que podem ser definidos pela sua reação com anticorpos monoclonais. Cerca de 100% das crianças em idade pré-escolar já foram infectadas e muitas delas sofrem mais de uma infecção pelo hRSV ao longo da vida. Com isso, podemos dizer que a infecção pelo hRSV induz a uma memória imunológica fraca, pois sua recidiva é freqüente. Apesar de possuírem anticorpos circulantes contra o vírus, idosos também se infectam especialmente se estão imunodeprimidos (AVEDAÑO et al., 2002).

Mesmo com mais de 50 anos do descobrimento do vírus, as várias tentativas de identificar uma terapia que melhore a situação clínica do paciente e as conseqüências da infecção ainda são limitadas. Acredita-se que o desenvolvimento de uma vacina de combate ao vírus, tanto para crianças como para adultos, seja a melhor forma de imunização da população (WRIGHT & PIEDIMONTE, 2011). No entanto, todas as tentativas de desenvolver uma vacina segura e efetiva para o uso humano estão longe do sucesso. Vacinas de formas inativas do vírus já foram testadas e falharam em proteger crianças contra a doença. Vacinas com o vírus atenuado ainda são desenvolvidas, mas induzem a uma fraca resposta

imune ou chegam a reverter o quadro de virulência em crianças pequenas (OLSZEWSKA & OPENSHAW, 2009).

Atualmente, Palivizumab é um dos medicamentos utilizados para prevenção da doença. No entanto, o produto possui controvérsias com relação ao seu custo e benefício. Palivizumab é um anticorpo monoclonal que atua contra a proteína de fusão (F). O produto é administrado de forma intravenosa e é desenvolvido para proporcionar imunidade passiva contra o hRSV, prevenindo ou reduzindo a severidade de uma infecção pelo vírus. Palivizumab não tem nenhuma reação adversa séria e não há evidências de que intensifica os sintomas da doença (WANG et al., 2008). Contudo, devido ao alto custo do remédio (cerca de 288 Dólares por quilo de massa corporal) é recomendada a administração apenas à pacientes de alto risco, como recém transplantados ou crianças com doenças cardíacas e pulmonares (HU & ROBINSON, 2010).

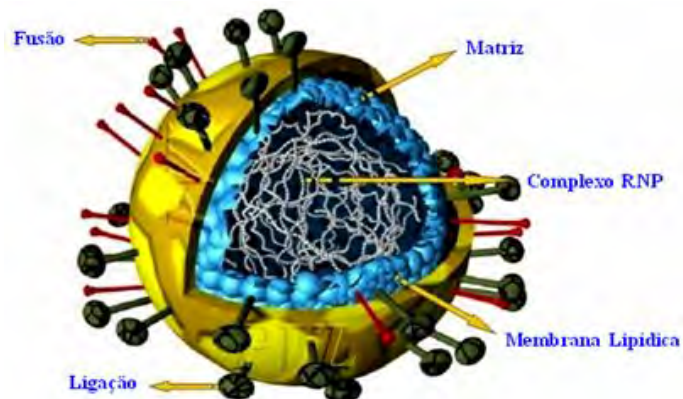
Uma das primeiras intervenções medicamentosas contra o vírus foi o uso da Ribavirina. A Ribavirina é usada pra tratar casos de infecções graves, apesar das limitadas evidências de benefício. A droga atua reduzindo a síntese do RNA mensageiro e da polimerase viral. A forma de administração é em aerosol para crianças com bronquiolite severa e pacientes imunocomprometidos. O seu custo também é bastante alto e já existem relatos de mutantes virais resistentes a sua ação, o que tem feito o seu uso cada vez menor (WELLIVER, 2010). Sendo assim, existe uma clara e urgente necessidade do desenvolvimento, tanto de vacina eficiente, quanto de novas drogas antivirais (OLSZEWSKA & OPENSHAW, 2009).

Estudos recentes para a neutralização do vírus concentram-se nas interações de suas proteínas de com substâncias que impedem a infecção da célula hospedeira. Conhecer a forma de interação destas proteínas com substâncias de ação inibidora é importante para sugerir um mecanismo efetivo de interrupção do processo de infecção da célula pelo vírus.

1.3 - Caracterização do Vírus Sincicial Respiratório Humano (hRSV)

O hRSV é um vírus envelopado pertencente ao gênero *Pneumovirus*, da subfamília *Pneumovirinae*, da família *Paramyxoviridae*, da ordem *Mononegavirales*, na qual também estão inclusos vírus como o que causa o sarampo e o vírus influenza (MORTON, 2003). O vírus possui um tamanho médio de 120 a 300 nm, simetria helicoidal (Figura 1), é sensível ao calor (56° C em 30 minutos), ao pH ácido (aproximadamente 3,0) e a solventes lipídicos (SANTOS et al., 2002).

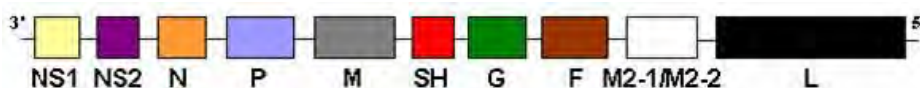
Figura 1 - Representação esquemática do Vírus Sincicial Respiratório (RSV).



No esquema estão representadas as proteínas de fusão F, de ligação G e de matriz M; o complexo ribonucleoprotéico e o envelope composto por bicamada lipídica. Fonte: <http://pathology5.pathology.jhmi.edu/micro/Image23.jpg>

O genoma do hRSV (Figura 2) é formado por RNA de fita simples e polaridade negativa, não segmentada, composta por aproximadamente 15.200 nucleotídeos e 10 genes, os quais codificam 11 proteínas. A ordem de transcrição dos genes é: NS1 (139aa e 15kDa) e NS2 (124aa e 14kDa), que correspondem às proteínas não estruturais; gene da nucleoproteína N (391aa e 44kDa); da fosfoproteína P (241aa e 34kDa); da matriz glicosilada M (256aa e 28kDa); a pequena proteína hidrofóbica SH (65aa e 7,5kDa); o gene da glicoproteína de ligação G (298aa e 32kDa); gene da proteína de fusão F (574aa e 70kDa) e o gene da proteína polimerase L (2.165aa e 200kDa). As proteínas G, F e SH constituem o envoltório do vírion e são responsáveis pela indução dos anticorpos neutralizantes (KOSAKA et al., 2004).

Figura 2 – Esquema da organização genômica do RNA do RSV.



Os genes NS1 e NS2 codificam proteínas não estruturais (NS1 e NS2); N, P e L codificam proteínas associadas ao nucleocapsídeo viral; M, M2.1 e M2.2 codificam proteínas de matriz; G, F e SH codificam proteínas transmembrânicas de superfície responsáveis pela ligação, fusão e adsorção viral à célula hospedeira, respectivamente. Fonte: Modificada a partir de www.micro.msb.le.ac.uk.

O vírion apresenta um envelope lipídico que envolve o nucleocapsídeo helicoidal associado às proteínas N, P e L, necessárias para direcionar a replicação viral (FALSEY & WALSH, 2000). A proteína N se liga fortemente ao RNA genômico e antígenômico na formação do nucleocapsídeo. A proteína P atua como chaperonina para a forma solúvel da proteína N, enquanto que a proteína L é a polimerase viral, fazendo parte do complexo RNA polimerase

(SILVA et al., 2009). Tem sido verificado que a proteína M2-1 (194aa e 22kDa) tem sua função associada à transcrição ou elongação e provavelmente está associada ao nucleocapsídeo viral. A proteína M2-2 (93aa e 12kDa) está relacionada à regulação da síntese do RNA do vírus (TENG & COLLINS, 1998).

O envelope é derivado da membrana da célula hospedeira e contém projeções que correspondem as glicoproteínas transmembrânicas de ligação G e de fusão F, que estão envolvidas no processo de ligação e adsorção viral. A proteína SH não parece ter um papel importante na replicação ou patogênese do vírus, uma vez que sua deleção pouco diminui a infectividade do vírus *in vitro* e em chimpanzés. Sua função ainda não foi bem especificada, mas acredita-se que tenha um papel auxiliar no processo de fusão do vírus (KOSAKA et al., 2004), pois quando expressa em bactérias, altera a permeabilidade da membrana do organismo (COLLINS & GRAHAM, 2008).

A proteína M é pequena, não glicosilada, interna a parte viral, tendo como função promover a associação do nucleocapsídeo com o envelope viral nascente (SILVA et al., 2009). Pouco se conhece sobre a função e a conformação das proteínas não-estruturais NS1 e NS2, mas sabe-se que elas podem interagir com o citoesqueleto de células infectadas e que atuam como antagonistas de interferons (IFN). O hRSV é um dos vírus mais eficientes no bloqueio da síntese de IFN do tipo I em indivíduos infectados. Estas proteínas inibem a indução de IFN-alfa e beta, bloqueando a ativação do IFN por meio do fator regulatório 3 (EVANS et al., 1996, COLLINS & GRAHAM, 2008).

O processo infeccioso se inicia com a ligação do vírus à célula através da proteína G. O envelope viral se funde com a membrana plasmática da célula hospedeira por meio da proteína F, sendo este um passo necessário para a transferência da ribonucleoproteína viral para o interior da célula. Após a penetração, o nucleocapsídeo do vírus é lançado no citoplasma celular, onde se inicia a transcrição de seu material genético. O RNA viral é usado como modelo para síntese do RNA mensageiro (mRNA). O RNA mensageiro, por sua vez, é usado como modelo para tradução das proteínas virais e transcrição do RNA do vírion. (OGRA, 2004). Após a síntese de suas proteínas e montagem dos vírions, as partículas virais são liberadas das células infectadas por brotamento, causando a ruptura e fusão com as células adjacentes (HACKING et al., 2002).

Os isolados do hRSV podem ser classificados em dois grupos, A e B, em que cada um deles possui subgrupos distintos. A classificação é baseada nas diferenças antigênicas e genômicas encontradas nas proteínas virais, especialmente na proteína G. A proteína F, pelo contrário, é altamente conservada, mostrando alta identidade entre os dois grupos do vírus.

Este fato é comprovado pela resposta semelhante do sistema imune à ação da proteína F, tanto do grupo A quanto do grupo B, enquanto que a resposta à proteína G é altamente específica ao grupo ou mesmo aos seus subgrupos (FALSEY & WALSH, 2000).

Os grupos A e B do vírus geralmente circulam simultaneamente, embora o vírus do grupo A seja mais predominante que o do grupo B. No Brasil, estudos mostraram que os grupos virais do hRSV circularam simultaneamente em diferentes anos nas regiões norte, sudeste e sul do país, apesar de nenhum tipo de padrão epidemiológico ter sido estabelecido (SIQUEIRA et al., 1991, MELLO & SILVA, 1992, STRALIOTTO et al., 1994).

Alguns estudos indicam diferenças no quadro clínico das crianças infectadas por cada um dos grupos virais. Foi observado que a maioria das crianças hospitalizadas infectadas pelo vírus do tipo A apresentavam maior grau de severidade da doença em relação àquelas infectadas pelo do tipo B (SULLENDER, 2000). Os mecanismos pelos quais o vírus do tipo A causa infecções mais graves é desconhecida, porém acredita-se que seja pelas diferenças funcionais das proteínas virais ou do seu genoma. Após a observação das diferenças relacionadas à replicação viral dos dois grupos, notou-se que a replicação do vírus do tipo A é muito maior que as do tipo B, tanto em células em cultura como em teste com animais. Estes resultados indicaram a hipótese de que o grau de patogenicidade de cada grupo viral está relacionado com o nível de replicação (WALSH et al., 1997).

1.4 - Proteína F

A proteína de fusão é uma das proteínas de membrana que apresentam importante papel para replicação viral, além de possuir muitas características semelhantes a proteínas homólogas de outros *Paramyxovirus* (MATHEWS et al., 2000). A principal função desta proteína é mediar o processo de adsorção e propagação viral, por intermédio da fusão da membrana da célula hospedeira com o envelope viral, sendo também responsável pela formação do sincício, intervindo na fusão da membrana das células infectadas com a membrana das células adjacentes (GONZALEZ-REYES et al., 2001).

A proteína F é classificada como uma glicoproteína do tipo I por ser uma proteína de envelope que contém *coiled-coils*, assim, como as demais proteínas de fusão de outros vírus (DIMITROV, 2004). Ela compartilha de diversas características estruturais de proteínas de fusão de outros *Paramyxovirus*, principalmente com relação a domínios hidrofóbicos, resíduos de cisteínas e ativação proteolítica, que por sua vez resulta na exposição do peptídeo hidrofóbico de fusão (ZIMMER et al., 2002).

Alguns estudos evidenciaram que, diferentemente de todos os outros membros da subfamília *Pneumovirinae*, nem a proteína G ou a proteína SH são estritamente necessárias para a infecção do hRSV. Este fato foi confirmado pelo estudo de isolados recombinantes do hRSV em que os genes da proteína G e SH foram deletados. Nas células infectadas pelo vírus recombinante observaram que a formação dos sincícios multinucleados ainda ocorria, porém, ficou evidente que a proteína G reforça ainda mais o poder de fusão da proteína F (KARRON et al., 1997). Por essa razão acredita-se que a proteína F possa combinar a função de fusão com a de ligação viral à membrana da célula, sendo suficiente para mediar a infecção celular pelo vírus (SCHLENDER et al., 2003).

Outra característica comum da proteína F do hRSV com a de outros vírus membros da sua família, é a síntese da proteína como um precursor inativo, denominado F0 que necessita ser clivado para desempenhar sua função. Este precursor é sintetizado com 574 aminoácidos e peso molecular de 67 kDa. A clivagem é feita por proteases do tipo furina durante o transporte para a superfície da célula, resultando em dois polipeptídios, F1 e F2, ligados por uma ponte dissulfeto, sendo o polipeptídio F2 proveniente da extremidade N-terminal e o polipeptídio F1 proveniente da extremidade C-terminal. Como um dos resultados da clivagem, tem-se um pequeno peptídeo composto de 27 aminoácidos (pep27), localizado entre os dois sítios de clivagem. Todos os três possuem cadeias de oligossacarídeos N-ligados (GONZALES-REYES et al., 2001).

O processamento proteolítico ocorre em duas seqüências consenso distintas: sítio I, com seqüência RAR/KR, após o resíduo 109 e o sítio II, com seqüência KKRKRR, localizado após o resíduo 136. A clivagem em ambos os sítios é necessária para aquisição do potencial de fusão de membrana da proteína F, sendo que a clivagem completa de ambos os sítios, resulta em modificações conformacionais na molécula (RUIZ-ARGUELLO et al., 2002).

A modificação pós-traducional e conformacional expõe o peptídeo de fusão hidrofóbico na extremidade N-terminal de um fragmento ancorado na membrana. A maioria das proteínas de fusão, incluindo a proteína F do hRSV, possui um motivo de clivagem do tipo RX(K ou R)R, logo após seu peptídeo de fusão. Esta é uma seqüência reconhecida por furina endoproteases do Complexo de Golgi (KLENK & GARTEN, 1994). Estas modificações já foram demonstradas por estudos em microscopia eletrônica, comprovando que a mudança conformacional da proteína ocorre durante o ataque e subsequente processo de fusão, representando um passo importante para a finalização do processo de infecção (GONZÁLES-REYES et al., 2001).

Existem três seqüências hidrofóbicas no polipeptídeo F. A primeira é o peptídeo sinal, localizada na porção N-terminal da subunidade F2, que age no transporte para o retículo endoplasmático e é removida da glicoproteína madura pela ação de proteases (RUIZ-ARGUELLO et al., 2002). A segunda, composta de aproximadamente 20 resíduos, está localizada na região N-terminal da subunidade F1 e é denominado peptídeo de fusão, por ser responsável pela inserção da proteína F na membrana da célula hospedeira, provocando sua desestabilização com a conseqüente fusão da membrana do vírus com a membrana da célula hospedeira (SCHIMITT et al., 2004). A terceira seqüência hidrofóbica é uma região transmembrânica, localizada próxima a região C-terminal da subunidade F1 (GONZÁLES-REYES et al., 2001).

Adjacente ao peptídeo de fusão e a região transmembrânica estão localizadas duas seqüências “*heptad-repeat*” A e B (HRA e HRB). Os peptídeos HRA e HRB estão separados por um domínio de aproximadamente 270 resíduos, e formam um complexo trimérico em solução (feixe de alfa-hélices ou *6 helix bundle*) (ZHAO et al., 2000, NI et al., 2005). A informação cristalográfica destes complexos revelou um núcleo central formado por três alfa-hélices na região da HRA que interage com três alfa-hélices da HRB, acomodados em uma depressão da HRA trimérica (GONZALES-REYES et al., 2001).

Há evidências de que interações da proteína F com proteoglicanos heparan-sulfato (HSPG) facilitam a ligação e a entrada do vírus na célula hospedeira, o que mostra que os glicosaminoglicanos (GAGs) de superfície da célula podem representar um tipo de receptor para o hRSV (SCHIMITT et al., 2004). Após a ligação com o receptor, a proteína sofre mudanças conformacionais que direcionam o processo de entrada do vírus na célula. Assim, há pelo menos três estados conformacionais para a proteína F: um estado nativo antes da fusão, formado antes da clivagem; um estado intermediário em “*pré-hairpin*”, formado após o processo de clivagem proteolítica, e, o estado pós-fusão, em que duas regiões “*heptad-repeat*” formam um conjunto de seis alfa-hélices (ZHU et al., 2003).

1.5 – Glicosaminoglicanos Sulfatados

Os glicosaminoglicanos (GAGs) são polissacarídeos longos formados por unidades repetidas de dissacarídeos, normalmente um ácido hexurônico e uma hexosamina. Com exceção da hialuronana, todas as GAGs são sulfatadas e sintetizadas como modificações pós-traducionais das proteínas, chamadas proteoglicanas. Esta família de proteínas é grande e heterogênea no que diz respeito à estrutura e função. Muitas células, incluindo todas as células dos vertebrados, sintetizam proteoglicanos. A maioria é secretada na matriz

extracelular onde elas são os principais constituintes da cartilagem, do tecido conjuntivo frouxo e das membranas basais (POLLARD & EARNSHAW, 2002).

A principal função atribuída aos glicosaminoglicanos é relacionada à sua capacidade de interação com outras moléculas. Várias macromoléculas ligam-se aos glicosaminoglicanos, sendo a maioria dos ligantes proteínas ou proteínas conjugadas. A heterogeneidade das GAGs em termos de seu padrão de sulfatação e da sua estrutura química, facilita a ligação de uma variedade de proteínas. Essas ligações foram postuladas em vários contextos funcionais, como na organização estrutural da matriz extracelular e no controle da regulação do comportamento e do metabolismo celular (RAMAN et al., 2005).

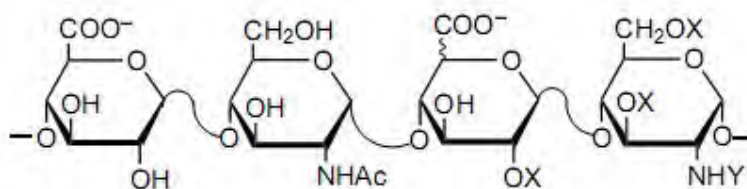
1.5.1 – Características da Heparina Sulfatada

Os polissacarídeos sulfatados e os polissacarídeos modificados solúveis possuem uma grande aplicação no campo biomédico e farmacêutico. Entre os polissacarídeos modificados com utilizações farmacêuticas, a heparina sulfatada tem sido uma das mais estudadas.

Estruturalmente, a heparina é um glicosaminoglicano de peso molecular variável, composta de unidades de Ácido Idurônico e Glucosamina, que se alternam entre ligações glicosídicas (Figura 3). Possui uma conformação linear, não ramificada, altamente sulfatada, e seu peso médio é aproximadamente 15 kDa. Uma unidade dissacarídica de heparina apresenta 2,5 grupos sulfatados para 1 grupo carboxílico, sendo que estes dois grupos desempenham um papel importante na ligação com ligantes carregados, como proteínas e íons bivalentes (VOET, 1995).

Sob o ponto de vista químico, a heparina é um mucopolissacarídeo sulfatado com grande quantidade de cargas elétricas, especialmente cargas negativas. É distinguida de outros polissacarídeos por ser extremamente ácida, devido à presença dos radicais sulfatados (SO_3^-) e do grupo carboxílico (COO^-), que garantem também a carga líquida negativa da molécula (CASU, 1989).

Figura 3 - Representação da unidade dissacarídica da Heparina



Fonte: CASU, 1989

A heparina é solúvel em água e é metabolizada pelo fígado, possuindo meia vida entre 90 minutos e 2 horas, pois pode ser afetada por alterações do fluxo sanguíneo hepático, metabolismo e hipotermia. Descoberta acidentalmente em 1916 pelo estudante de medicina Jay Mclean, esse glicosaminoglicano é utilizado há quase 70 anos por suas propriedades anticoagulantes e antitrombóticas (BARROSO, 2002).

A molécula de heparina utilizada na área farmacêutica é extraída de tecidos altamente vascularizados, como a mucosa intestinal de suínos ou de células pulmonares de bovinos. Durante o isolamento, as cadeias glicosaminoglicanas se degradam levemente produzindo uma mistura heterogênea de fragmentos com massa molecular de 3 a 30 kDa (BARROWCLIFFE, 1989).

A interação da heparina com a antitrombina desempenha um papel importante no mecanismo de regulação da coagulação. Estudos da estrutura cristalográfica do complexo antitrombina-heparina indicam que algumas regiões de contato dessa proteína com a heparina apresentam um grande número de resíduos básicos. Estes resultados indicam que a interação eletrostática entre grupos básicos da proteína e grupos ácidos do polissacarídeo desempenha um papel fundamental no reconhecimento molecular entre elas. De fato, os aminoácidos básicos formam pares iônicos com as cargas negativas dos grupos sulfatados e carboxílicos que estão distribuídos de forma definida ao longo da cadeia das GAGs. Além disso, interações não eletrostáticas como pontes de hidrogênio e interações hidrofóbicas contribuem para a estabilidade da ligação (FOSSEY, 2003).

1.6 - Flavonóides

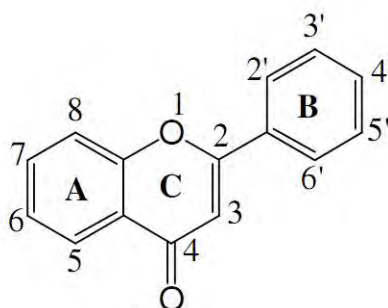
Um dos maiores grupos de polifenóis que ocorrem naturalmente são os flavonóides, responsáveis pelas cores de muitas frutas e flores. Já foram identificados cerca de 4000 compostos que são classificados de acordo com sua estrutura química. Eles foram descobertos juntamente com a vitamina C em 1928, por Albert Szent-Gyorgyi, que os havia chamado inicialmente de vitamina P (BYLKA & MATLAWSKA, 2004).

Diversos estudos demonstraram que os flavonóides possuem uma gama de efeitos biológicos, como atividades antibacterianas, antivirais, antiinflamatórias, antialérgicas e ações vasodilatadoras (COWAN, 1999). Além disso, ainda apresentam capacidade antioxidativa (uma das atividades mais estudadas até hoje), ação contra o desenvolvimento de tumores, antiulcerogênica e antiplaquetária. Muitas das ações biológicas dos flavonóides são atribuídas às suas propriedades antioxidantes, entretanto os mecanismos pelos quais estes compostos exercem seu papel benéfico ainda não foram bem esclarecidos (LOPES et al., 2003).

Os flavonóides podem ocorrer como agliconas (flavonóides sem açúcar ligado), glicosídeos e derivados metilados. Estruturalmente os flavonóides são compostos aromáticos de baixo peso molecular, constituídos basicamente por 15 átomos de carbono que ocorrem em unidades na forma C6-C3: dois anéis de 6 carbonos ligados a uma unidade de 3 carbonos que pode ou não ser parte de um terceiro anel. Por conveniência, os anéis são nomeados A, B e C (Figura 4) (TAPAS, SAKARKAR & KAKDE, 2008).

Estes compostos podem sofrer diversas modificações estruturais como hidroxilação, metilação, acilação e glicosilação, o que explica a sua grande diversidade. A maioria dos flavonóides ocorre naturalmente associados com açúcares sob forma conjugada (TAPAS, SAKARKAR & KAKDE, 2008). Nas plantas, os flavonóides estão presentes nas formas O-glicosiladas ou C-glicosiladas. A estrutura O-glicosilada possui a ligação de açúcar substituindo um grupo -OH nas posições 3 e 7, enquanto que o açúcar da forma C-glicosilada normalmente é ligado aos carbonos 6 ou 8 (SANDHAR et al., 2011). Os flavonóides são derivados de benzo-gama-pirona de origem vegetal e por isso facilmente pode haver interconversões entre eles (KOES et al., 2004).

Figura 4 - Estrutura geral dos Flavonóides.



Na figura, podemos ver os 15 carbonos além da formação ligada dos anéis aromáticos. Fonte: (TAPAS, SAKARKAR & KAKDE, 2008).

Os tipos de flavonóides se diferem pela posição de grupos laterais ligados, por esta razão eles podem ser distribuídos em várias classes que são determinadas de acordo com suas estruturas moleculares e seus substituintes. Algumas das principais classes são: flavonas, flavanonas, flavanas, flavonóis, isoflavonóides e biflavonas (NIJVELDT et al., 2001). Uma variação no anel C fornece a divisão em subclasses (SANDHAR et al., 2011).

A atividade antiviral dos flavonóides tem sido reconhecida desde os anos 40, mas é recente a intensificação de trabalhos sobre esta sua atividade. A ação antiviral de diversos flavonóides foi evidenciada para pelo menos 11 tipos de vírus. Devido à propagação do vírus

HIV em todo mundo, grande parte dos relatos sobre a atividade antiviral dos flavonóides focam a inibição deste vírus. Muitos trabalhos já constataram a atividade anti-HIV dos flavonóides, sendo pelo menos cerca de 28 compostos com ação inibitória sobre o vírus (TAPAS, SAKARKAR & KAKDE, 2008).

Os flavonóides possuem propriedades antivirais em condições *in vitro* e *in vivo*. Verificou-se que estes compostos reduzem a infectividade tanto de vírus envelopados como não envelopados, incluindo os de fita simples de RNA e os de fita dupla de DNA. Estudos recentes indicam que os flavonóides podem interferir na ligação do vírus a membrana celular, no processo de adsorção, na replicação e tradução das proteínas virais na célula do hospedeiro, além de intervir no processo de formação das glicoproteínas do envelope viral. Entretanto os mecanismos pelos quais os flavonóides agem para inibir a atividade viral, ainda não foram completamente elucidados (ANDRES et al., 2009).

Em um estudo que analisou a ação dos flavonóides quercetina, hesperetina e catequina sobre diversos vírus, concluiu-se que estes são efetivos contra: herpes simplex, poliovírus do tipo 1, hRSV e parainfluenza do tipo 3. A hesperetina reduziu a replicação intracelular dos 4 vírus, enquanto que a catequina inibiu o potencial de infecção dos vírus hRSV e HSV-1, porém, não a suas replicações intracelulares. Mais de um estudo concluiu que derivados da classe de flavonas possuem ação inibidora sobre o hRSV. Acredita-se que pequenas diferenças em suas estruturas são fatores determinantes para suas atividades antivirais (COWAN, 1999).

1.7 - Bioinformática

Atualmente a bioinformática é uma das ciências imprescindíveis para a manipulação dos dados biológicos. Estes dados são relativamente complexos quando comparados aos provenientes de outras áreas científicas, dada a diversidade e o inter-relacionamento. Toda informação genômica só é possível de ser organizada, analisada e interpretada com o apoio da informática. Por essa razão, a bioinformática pode ser definida como a modalidade que abrange todos os aspectos de aquisição, processamento, armazenamento, distribuição, análise e interpretação da informação biológica (SANTOS & ORTEGA, 2003).

O termo bioinformática referia-se inicialmente a análise de seqüências de proteínas e do DNA (TRIFONOV, 2000), mas nos últimos anos o termo foi generalizado de forma a incluir todos os aspectos da biologia molecular, desde a análise seqüencial de aminoácidos até análise de estrutura de proteínas, predição e modelagem de processos celulares, passando por análise de textos e armazenamento de dados (CORNE & MARTIN, 2001). Assim, a

bioinformática pode ser vista como uma fusão entre a biologia molecular e ciência da computação, destinada a organizar e interpretar informações biológicas em larga escala.

O papel primordial da pesquisa por ferramentas de bioinformática é o desenvolvimento de drogas de efeitos medicinais. No entanto, para que se possa chegar a este objetivo final, uma série de passos intermediários são necessários. Estes passos envolvem a organização dos dados de forma a permitir o acesso de outros pesquisadores às informações existentes, viabilizar a apresentação de novos dados, desenvolver ferramentas e recursos que tornem possível a análise dessas informações e interpretar os resultados das análises computacionais de forma a torná-los significativos no campo da biologia (LUSCOMBE et al., 2001).

As várias subáreas de estudo em bioinformática requerem as mais variadas técnicas computacionais. Uma vez que a seqüência do DNA de um organismo é conhecida, as proteínas codificadas por essa seqüência podem ser previstas, sendo um dos principais objetivos de uma análise computacional de seqüências primárias, a previsão estrutural e funcional dessa proteína. Outro aspecto importante para se estudar, é se possível, determinar como uma dessas proteínas pode interagir com outras proteínas ou até mesmo, fármacos e outros ligantes (SANTOS & ORTEGA, 2003).

A bioinformática apresenta-se assim, como uma ferramenta capaz de direcionar os estudos em biologia molecular, de forma a mapear e indicar uma direção para um melhor estudo genômico experimental.

2 – Objetivos

Através de diferentes ferramentas em bioinformática, este trabalho teve por objetivos:

- a) Buscar possíveis sítios sequenciais de interação e características importantes que indiquem interações da proteína F com compostos inibidores da atividade viral;
- b) Obter um modelo tridimensional válido da proteína F do hRSV e busca de motivos estruturais de interação com outras substâncias;
- c) Estudar através da técnica de *docking* molecular o modelo de interação da proteína F do hRSV com a Heparina Sulfatada e com Flavonóides, dadas as evidências experimentais destas interações e suas propriedades antivirais;
- d) Sugerir um provável inibidor contra a ação viral, para futuros ensaios experimentais em laboratório.

3 - Metodologia

3.1 – Construção do Banco de Dados de Sequências Primárias da Proteína F

Foram selecionadas 43 seqüências da proteína de Fusão do hRSV da base de dados internacional GenBank, que é mantida pelo *National Center of Biotechnology Information* (NCBI). Genbank é um banco de dados global que contém publicamente seqüências de nucleotídeos e de aminoácidos de mais de 300.000 organismos, obtidas pela submissão individual de laboratórios e projetos de sequenciamento em larga escala (BENSON et al., 1996). Esta característica fornece ao trabalho uma variabilidade entre as seqüências escolhidas, facilitando a análise de possíveis diferenças entre seqüências de várias regiões do mundo. O GenBank é acessado por meio da *homepage* do NCBI (www.ncbi.nlm.nih.gov). Todas as seqüências foram salvas em formato *fasta* e catalogadas por meio de seus números de acesso do NCBI.

3.2 – Análises das seqüências

Com o auxílio do programa BIOEDIT (HALL, 1999), realizou-se o alinhamento global de todas as seqüências selecionadas com o programa ClustalW (LARKIN et al., 2007), para posterior análise de grau de identidade e de possíveis mutações entre as seqüências selecionadas da proteína de fusão. O programa BIOEDIT é um editor de seqüências biológicas e destina-se a fornecer funções básicas para edição de seqüências de proteínas e nucleotídeos, bem como alinhamento destas seqüências e análise. O programa é apenas para plataforma Windows XP ou versões anteriores e pode ser obtido pelo endereço <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>.

Após o alinhamento e análise do grau de identidade entre as seqüências da proteína, as seqüências foram submetidas a programas de predição de estrutura secundária e de características da proteína como: sítios de glicosilação, sítios de fosforilação, sítios de sulfatação, ponto isoelétrico, predição da seqüência do peptídeo sinal, topologia da proteína, (indicando regiões transmembrânicas) e seqüências de localização celular. Os programas utilizados foram os *webservers* PREDICTPROTEIN (ROST, YACHDAV & LIU, 2004), PSIPRED (MCGUFFIN, BRYSON & JONES, 2000) e PDBSum (LASKOWSKI et al., 1997).

PREDICTPROTEIN é um serviço *on line* para análises de seqüências primárias por meio da predição de aspectos estruturais e funcionais da proteína. O servidor pode ser acessado através do endereço www.predictprotein.org. O usuário submete a seqüência de

aminoácidos ou um alinhamento e o programa resulta em múltiplos alinhamentos, faz busca de motivos de interação proteína-proteína ou até mesmo, interação proteína-DNA; procura por regiões de baixa complexidade, regiões sem estrutura regular, regiões de acessibilidade do solvente, regiões globulares, regiões em *coiled-coil* e pontes dissulfeto. Além disso, ele prevê a probabilidade de existir sítios de modificação pós-traducionais, como sítios de glicosilação, fosforilação e sulfatação (ROST, YACHDAV & LIU, 2004).

PSIPRED é um servidor de predição da estrutura secundária de uma proteína baseado em métodos de redes neurais. O servidor é mantido pelo *UCL Bioinformatics Group* e pode ser acessado através do endereço <http://bioinf.cs.ucl.ac.uk/psipred/>. Além da estrutura secundária, outras informações podem ser retiradas do programa, como a topologia da proteína e prováveis regiões transmembrânicas (BUCHAN et al., 2010). O programa PDBSum fornece um resumo de informações sobre o modelo estrutural da proteína. As características que o programa analisa incluem diagrama da topologia, interação proteína-proteína e domínios estruturais. O programa é mantido por uma parceria entre diversos centros de pesquisas do mundo, mas é hospedado no servidor do *European Bioinformatic Institute* (EBI) e é acessado pelo endereço <http://www.ebi.ac.uk/pdbsum> (LASKOWSKI et al., 2009).

3.3 – Modelagem da Proteína de Fusão

A determinação da estrutura terciária de uma proteína é um dos maiores desafios para biologia estrutural computacional da atualidade (WU, SKOLNICK & ZHANG, 2007). Ao mesmo tempo, conhecer a estrutura tridimensional de uma proteína é essencial para a compreensão da sua função (ZHANG, 2008). Uma vez que poucas informações estruturais experimentais da proteína de fusão são encontradas, utilizamos a técnica de modelagem de proteínas para um melhor estudo das características da proteína F.

Foram considerados os três métodos de modelagem disponíveis:

- I) Modelagem Comparativa (por homologia);
- II) Modelagem por *Threading*;
- III) Modelagem *Ab Initio*;

Tanto para o método I quanto para o método II, foi preciso procurar por moldes (*template*) da proteína F, que seriam usados como informação experimental para a construção do modelo. A busca foi feita pelo programa BLAST (ALTSCHUL et al., 1997), acessível pelo endereço <http://blast.ncbi.nlm.nih.gov/Blast.cgi>, que é mantido pelo servidor do NCBI. Com a seleção dos moldes, os programas MODELLER (<http://salilab.org/modeller/>) e I-TASSER

(<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) foram utilizados para a modelagem da proteína F pelos métodos I e II, respectivamente.

A modelagem por homologia baseia-se nos padrões gerais a nível molecular, em que a homologia entre seqüências de aminoácidos implica em semelhança estrutural e funcional. As proteínas homólogas apresentam normalmente regiões internas conservadas, principalmente constituídas de elementos de estruturas secundárias como alfa-hélices e folhas-beta. Porém, um dos impedimentos dessa técnica são as diferenças estruturais que ocorrem nas regiões externas das proteínas homólogas, principalmente as que são constituídas de alças que se ligam a elementos da estrutura secundária (SANTOS-FILHO & ALENCASTRO, 2003).

O método de *threading* é baseado no princípio de que o número dos diferentes enovelamentos das proteínas é bem mais limitado que o número de seqüências existentes. O número de enovelamentos tem sido estimado pelo agrupamento em família das estruturas depositadas no *Protein Data Bank* (PDB). O método de *threading*, desta forma, tem por objetivo encaixar a seqüência alvo de uma proteína a uma biblioteca de estruturas com enovelamentos conhecidos. O método ainda permite integrar limitações na proteína alvo, como pontes dissulfeto ou restrições de distância (TOOMULA, KUMAR & KUMAR A, 2011).

Para o método III foi utilizado o programa ROSETTA (RHOL et al., 2004), desenvolvido por um grupo da Universidade de Washington. Neste método, submete-se a seqüência de aminoácidos ao programa para os cálculos posteriores. Assim, o programa calcula uma estrutura, implementando um modelo de enovelamento em fragmentos pequenos de proteínas, que se alternam entre diferentes conformações locais, copiados de segmentos de proteínas conhecidas, não necessariamente homólogas. A probabilidade de assumir uma conformação particular é baseada na similaridade da seqüência local e da estrutura secundária predita da seqüência alvo (GADJA, PAWLOWSKI & BUJNICKI, 2011). O algoritmo do programa trabalha com funções de energia potencial para calcular interações com e entre macromoléculas, bem como otimização do método pela determinação da estrutura de mais baixa energia para uma seqüência de aminoácidos (KAUFMANN et al., 2010).

3.4 – Otimização e Validação dos modelos

Os modelos gerados pelos três métodos foram inicialmente avaliados para a escolha do melhor modelo. Os métodos de validação incluem análise pelos programas PROCHECK (LASKOWSKI, 1993) e PROQServer (WALLNER & ELOFSSON, 2003). O programa PROCHECK faz uma avaliação estereoquímica do modelo, ou seja, analisa comprimentos de

ligações, ângulos diedrais e distância entre átomos não ligados. O programa gera um gráfico de Ramachandran que indica resíduos que estão em posições energeticamente favoráveis e desfavoráveis, utilizando como parâmetro proteínas que já tiveram sua estrutura resolvida experimentalmente (LASKOWSKI, 1993). O programa PROQServer utiliza redes neurais para avaliar a qualidade dos modelos de proteínas, através de características como contato entre resíduos, contatos entre átomos e superfície de acessibilidade do solvente. A qualidade do modelo é determinada usando parâmetros de medidas baseados nos métodos LGscore (CRISTOBAL et al., 2001) e MaxSub (SIEW et al., 2000). Estes são métodos específicos para avaliação da qualidade da estrutura de uma proteína, uma vez que métodos mais simples, como o desvio médio quadrático (RMSD), calculado sobre todos os átomos, mostram-se pobres como fatores de avaliação. O RMSD de regiões preditas erroneamente resulta em valores tão altos que dificulta afirmar se o modelo possui regiões que foram preditas de forma correta. Assim, é necessária uma forma de identificar regiões preditas corretamente e dar a elas uma pontuação que represente quanto elas se aproximam de regiões experimentais similares. Sendo assim, os métodos LGScore e MaxSub tem se mostrados bons métodos para cobrir essa lacuna (SIEW et al., 2000).

A otimização do melhor modelo foi feita pelos programas WinCoot (LOHKAMP et al., 2005) e Chiron (RAMACHANDRAN et al., 2011). WinCoot é um *software* gráfico molecular diferente de programas que apenas fazem a exibição gráfica das coordenadas de um arquivo PDB, como o Pymol por exemplo. O WinCoot permite a construção de modelos macromoleculares e tem a função de finalização e validação. Realizando cálculos específicos para a proteína de trabalho, é possível avaliar a simetria das coordenadas dos átomos do modelo, o que possibilita a correção de imperfeições encontradas. A manipulação do modelo é permitida em um espaço próprio para o refinamento, em que o programa permite a rotação ou translação das posições das coordenadas, ajustes de corpos rígidos, criação do efeito de solvatação, mutações, construção do gráfico de Ramachandran para validação, análise da simetria cristalográfica do modelo, entre outras funções (EMSLEY & COWTAN, 2004).

O programa Chiron é um servidor *on line* que faz a minimização de choques estéricos entre os átomos da estrutura modelada. Os choques estéricos são um dos principais problemas em modelos de baixa resolução e de modelos gerados por técnicas de modelagem. Eles acontecem devido à sobreposição não natural de dois átomos não ligados na estrutura da proteína. O programa Chiron apresenta um protocolo que de forma eficiente resolve vários destes choques com o mínimo de perturbação no esqueleto da proteína. Inicialmente o protocolo define choques estéricos como qualquer sobreposição atômica que resulte em uma

energia de repulsão de Van der Waals maior que $0.5kT$ (em que k é a constante de Boltzmann e T a temperatura), com exceção para átomos ligados, pontes dissulfeto e átomos do esqueleto da proteína. Utilizando o método de Dinâmica Molecular Discreta, usando campo de força CHARMM, o protocolo faz as alterações necessárias para, quando possível, diminuir os choques estéricos. A vantagem deste método está na eficiência e no pouco tempo computacional para obtenção dos resultados (RAMACHANDRAN et al., 2011).

Após a otimização do modelo, este novamente passou pelos métodos de validação dos programas PROCHECK e PROQ-Server para análise da melhoria e encerramento dos procedimentos de otimização.

3.5 – Busca por sítios de interação

3.5.1 – Sítio de interação da Proteína de Fusão e a Heparina Sulfatada

Um dos primeiros estudos voltados para a determinação de sítios de ligação entre proteínas e a Heparina foi realizado por Cardin e Weintraub em 1989. Por meio da comparação de sítios de ligação conhecidos da Heparina com várias proteínas, eles perceberam que estes domínios possuíam seqüências similares e motivos definidos, determinando então dois tipos de seqüências consenso, [-X-B-B-B-X-X-B-X-] e [-X-B-B-X-B-X-], em que B é um resíduo básico e X um resíduo hidropático. Esta região também foi encontrada em outras proteínas estudadas que apresentavam conhecidamente ligação com a Heparina.

Sítios de ligação com a Heparina comumente são observados na superfície externa das proteínas e correspondem a grupos com alta densidade de carga positiva. Estudos desenvolvidos por Capila e Linhardt (2002) mostraram os resíduos mais comuns nos domínios de ligação das proteínas com a Heparina. Peptídeos ricos em Arginina e Lisina, mas não Histidina, apresentam maior afinidade de ligação. A força relativa de ligação da Arginina é cerca de 2,5 vezes maior que a Lisina. Este fato é resultado da forte interação entre o grupo guanidina da Arginina e o grupo sulfato da Heparina.

Estudos identificaram glicosaminoglicanos (GAGs) na superfície celular, em particular a Heparina Sulfatada, como um dos fatores de maior importância para infecção do hRSV em células de cultura. Estudo experimental realizado por Rubio e colaboradores (2010) também verificou que a Heparina em concentrações acima de $3\mu\text{M}$, inibe a infecção viral. Mesmo na ausência das outras proteínas que mediam a interação (as proteínas G e SH), o vírus possui a capacidade de se ligar à célula e promover a infecção do hospedeiro, sugerindo

que a proteína F também realiza uma ligação com a Heparina Sulfatada presente na membrana celular (MATHEWS et al., 2000).

Os glicosaminoglicanos são encontrados na superfície de tipos celulares de mamíferos, em tecidos e na matriz extracelular. Vários vírus, incluindo o hSRV, têm utilizado os glicosaminoglicanos, em especial a Heparina, para a ligação celular e entrada do vírus à célula. O papel da Heparina durante a ligação com o vírus foi demonstrado por diferentes experimentos, tais como a remoção prévia da Heparina Sulfatada das células por tratamento enzimático, ou pré-incubação de partículas do hRSV com Heparina Sulfatada solúvel (MELERO et al., 2004). Estes experimentos mostraram uma significativa diminuição da infectividade do vírus, porém, ainda não foi esclarecido qual o papel da Heparina em casos clínicos isolados, assunto que requer maior atenção e estudos.

Desta forma, a busca por possíveis sítios de ligação a glicosaminoglicanos baseou-se na identificação de motivos quanto às propriedades físico-químicas das seqüências alinhadas pelo programa BIOEDIT. Devido ao importante papel desempenhado pelas interações eletrostáticas, procuraram-se motivos de aminoácidos básicos, principalmente aqueles que apresentavam Arginina e Lisina, já que estes resíduos estão presentes na maioria dos sítios de ligação conhecidos entre proteínas e Heparina, como descrito anteriormente.

3.5.2 – Sítio de interação da Proteína de Fusão e os Flavonóides

Diferentemente da Heparina, Flavonóides não possuem relatado na literatura algum tipo de padrão de interação com proteínas. Desta forma, para determinarmos possíveis sítios de interação na proteína de Fusão com os Flavonóides, foi utilizado programas de predição da interação de proteínas com pequenas moléculas, como são os Flavonóides usados neste estudo. A predição é feita sobre a estrutura tridimensional da proteína, desta forma, o modelo otimizado e validado da proteína F foi submetido à avaliação dos programas de predição.

Foram utilizados 3 programas para comparação dos resultados: fpocket (LE GUILLOUX et al., 2009), CASTp (BINKOWSKI, NAGHIBZADEH & LIANG, 2003) e Q-SiteFinder (LAURIE & JACKSON, 2005). Programas de predição de sítios de interação entre proteínas e pequenas moléculas baseiam-se na ideia de Emil Fischer (1894), em que a ligação em um substrato de uma enzima é como inserir uma chave em uma fechadura. A complementariedade da forma entre o ligante e a proteína é um importante fator na ligação, já que pequenas moléculas normalmente ligam-se a sítios côncavos na superfície das proteínas. No entanto, a interação proteína-ligante não se baseia somente na complementariedade da forma, mas também na complementariedade devido às propriedades físico-química. O

balanço entre as interações de Van der Waals, eletrostáticas, ligações de hidrogênio, interações hidrofóbicas, e interações com o solvente devem resultar em uma ligação energeticamente favorável (HENRICH et al., 2009).

Os resultados obtidos pelos programas foram comparados e analisados quanto a sua importância no que diz respeito à função e atividade da proteína.

3.6 – Docking Molecular

A predição da ligação de pequenas moléculas com proteínas é uma prática importante, uma vez que por meio dela, é possível realizar uma busca em uma biblioteca de moléculas com potencial terapêutico que possam ser usadas futuramente em pesquisa e desenvolvimento de novas drogas (TROTT & OLSON, 2009).

Novos ligantes a cada dia são descobertos utilizando o método computacional. Recentemente, estruturas conhecidas de ligantes complexados com seus receptores, têm sido corretamente preditas pelas técnicas computacionais quando é simulada a interação entre as duas estruturas independentes. Do ponto de vista químico, a busca computacional já identificou pelo menos 50 novos ligantes de estruturas conhecidas, ou mesmo em alguns casos, de estruturas modeladas computacionalmente. Sendo assim, os compostos de uma biblioteca utilizados nestes estudos ficam disponíveis para constatações experimentais (SHOICHET, 2004).

3.6.1 – Docking Molecular da Proteína de Fusão e a Heparina Sulfatada

O *docking* molecular é um procedimento computacional que tenta prever ligações não covalentes de uma macromolécula (receptor) e uma pequena molécula (ligante) de forma eficiente, a partir do estado não ligado das estruturas. Estas estruturas podem ser obtidas por simulações de dinâmica molecular ou técnicas de modelagem, dentre outros métodos que fornecem estruturas tridimensionais. O objetivo é prever a provável conformação e a afinidade da ligação (TROTT & OLSON, 2009).

O modelo da estrutura terciária da proteína de fusão foi utilizado para este procedimento. Inicialmente, preparou-se o arquivo por meio do programa AUTODOCKTOOLS (MORRIS et al., 2009). Foram retiradas possíveis moléculas de água, adicionaram-se átomos de hidrogênio não polares e calcularam-se as cargas da proteína. O arquivo final foi convertido do formato PDB para o formato PDBQT. O AUTODOCKTOOLS é uma interface gráfica que facilita a formatação dos arquivos das moléculas de entrada, com um conjunto de métodos que guiam o usuário na realização dos

procedimentos de protonação, cálculo de cargas e rotação de ligações no ligante e na proteína (MORRIS et al., 2009).

Uma vez que a estrutura experimental da Heparina Sulfatada já foi resolvida, utilizou-se a estrutura disponível no banco de dados PDB (Protein Data Bank, acessível pelo endereço <http://www.rcsb.org/pdb/home/home.do>) de código 1HPN. A estrutura também foi preparada pelo programa AUTODOCKTOOLS. Foram calculadas as cargas do composto e estabelecidas as flexibilidades de suas ligações, definindo a forma como eles iriam se comportar durante a simulação. Uma vez que a Heparina é um composto extremamente flexível, todas as torsões foram permitidas. O docking foi realizado pelo programa AUTODOCKVINA, que apresenta um tempo computacional menor e um nível de precisão mais alto em seus resultados.

3.6.2 – Modelagem tridimensional dos Flavonóides

Os compostos utilizados neste estudo foram cedidos pelo Prof. Dr. Wagner Vilegas, docente do Instituto de Química da UNESP, campus de Araraquara. Estes flavonóides fazem parte de um estudo que avaliou a composição da flora do estado de São Paulo e avaliou pequenas moléculas com potencial farmacológico. A maioria destes compostos não possui sua estrutura tridimensional experimental e apenas alguns dos flavonóides agliconas (sem glicosilação) possuem estruturas resolvidas. Neste caso, uma vez que são pequenas moléculas, suas estruturas tridimensionais foram modeladas a partir de informações da sua estrutura 1D/2D.

Vários estudos já foram desenvolvidos para analisar e comparar estruturas modeladas, com estruturas de pequenas moléculas já determinadas experimentalmente (MITEVA, GUYON & TUFFÉRY, 2010). As estruturas tridimensionais de pequenas moléculas podem ser preditas com bastante precisão por métodos de mecânica quântica ou aproximações de mecânica molecular, porém estas são técnicas que demandam muito tempo computacional. Por outro lado, aproximações baseadas em regras simples ou que utilizam molde de estruturas já resolvidas, são técnicas que apresentam bons resultados em menos tempo e que podem ser usadas em grande escala (ANDRONICO et al., 2011). Desta forma, escolhemos o programa FROG2 para modelar as estruturas tridimensionais dos flavonóides.

O programa FROG2 é capaz de gerar uma estrutura única ou um conjunto de conformações tridimensionais de baixa a média energia, a partir da informação 1D/2D ou 3D da estrutura do ligante. O dados 1D/2D podem estar no formato SMILES (*Simplified Molecular Input Line Entry System*), 1D cansmiles (*Canonical SMILES*) ou em formato SDF

2D (*Structure Data File*). As estruturas geradas por este programa são totalmente ou parcialmente não ambíguas do ponto de vista estereoquímico, incluindo sítios quirais. As estruturas calculadas passam por um processo de otimização devido a dois estágios do algoritmo de Monte Carlo que estão incorporados ao programa. O primeiro passo evita conformações redundantes pela exploração da flexibilidade limitada dos ângulos diedrais, enquanto que o segundo passo considera pequenas rotações nas estruturas. Uma minimização de energia é aplicada às conformações usando campo de força AMMP, implementado pelo programa AMMOS (MITEVA, GUYON & TUFFÉRY, 2010).

Foram modelados 34 compostos, em que a partir de seus nomes, foi desenhada a sua fórmula estrutural. Posteriormente estas estruturas foram convertidas em formato SMILES, para servirem como dados de entrada para o programa FROG2. A validação dos modelos foi feita calculando o RMSD entre as poucas estruturas experimentais existentes, com aquelas modeladas pelo programa.

3.6.3 – *Virtual Screening: Docking Cego e Docking Dirigido*

O rápido avanço de informações na área genômica resultou em um aumento de potenciais alvos terapêuticos, bem como a demanda por métodos que possam identificar rapidamente possíveis candidatos para interação com estes alvos. *Virtual screening* é uma técnica computacional utilizada na pesquisa de novas drogas. Sua aplicação se dá principalmente quando existe uma biblioteca de compostos químicos e se quer avaliar quais deles interagem melhor com o receptor. Tradicionalmente, ferramentas computacionais são mais bem utilizadas para analisar moléculas individuais ou uma biblioteca de compostos, com o objetivo de eliminar estruturas não desejadas (SCHNEIDER & BOHM, 2002). Uma vez que possuímos 34 flavonóides, esta técnica foi aplicada.

No entanto, como não existe um sítio de interação importante conhecido entre a proteína de fusão do hRSV e algum flavonóide, aplicou-se duas técnicas de *docking*. Em um primeiro momento, foi realizado o *docking* cego (*blind docking*), em que o espaço de interação entre os compostos e a proteína corresponde à superfície total da proteína. Posteriormente, foi realizado o *docking* dirigido, em que a área de interação é delimitada a uma região com importância funcional para a proteína e assim seja avaliada a possibilidade de interação entre os compostos e esta região.

Da mesma forma que com o *docking* com a Heparina, os arquivos de todos os flavonóides foram preparados no programa AUTODOCKTOOLS. Todos os compostos foram protonados, tiveram suas cargas calculadas e flexibilidade definida. Os arquivos foram

convertidos do formato PDB para o formato PDBQT. Em todos os procedimentos, o *docking* foi realizado pelo programa AUTODOCKVINA.

provavelmente é devido a restrições estruturais ou funcionais com relação à substituição de aminoácidos (JOHNSON & COLLINS, 1988). Estas propriedades tornam a proteína de fusão um dos maiores alvos para o desenvolvimento de drogas antivirais (WELLIVER, 2010).

O grau de identidade entre as sequências é devido a alguns fatores que apontam para a conservação do mecanismo de fusão entre as proteínas F de vírus da família *Paramyxoviridae*. Foi descrito por Dutch (2010) que proteínas de fusão dos vírus dessa família possuem seus resíduos de cisteína, glicina e prolina em posições conservadas sugerindo uma estrutura tridimensional similar. As proteínas de fusão dos vírus da família *Paramyxoviridae* também possuem peptídeos de fusão e regiões *heptad repeat* conservadas. As regiões *heptad repeat* têm se mostrado regiões importantes no bloqueio da atividade da proteína de fusão de vários vírus (DUTCH, 2010).

A busca por sítios de interações dos tipos [-X-B-B-B-X-X-B-X-] e [-X-B-B-X-B-X-], entre a proteína de Fusão e a Heparina Sulfatada mostrou a presença deste motivo na região 130 a 137 da proteína, formada pelos resíduos [S-K-K-R-K-R-R-F]. Como pode ser observado na figura 6, este motivo está presente em todas as seqüências alinhadas, mostrando-se altamente conservado, sugerindo que esta é uma região de grande importância no processo de infecção viral. Este resultado concorda com o encontrado por Cardin e Weintraub (1989) que descreveram motivos seqüenciais de proteínas que comprovadamente interagem com a Heparina Sulfatada.

Figura 6 – Potencial região de interação da proteína F com a Heparina .

NP_044596.1	RELERFMNYTLNNTKNTNVT	SKKKRRFLG
AAC57037.1	RELERFMNYTLNNTKNTNVT	SKKKRRFLG
NP_056863.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
AAB82436.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
ACO83297.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
CAA81295.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
AAC57027.1	RELERFMNYTLNNTKNTNVT	SKKKRRFLG
AAX23994.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
AAA47410.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
AAB82446.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
AAB38520.1	RELERFMNYTLNNTKNTNVT	SKKKRRFLG
ABI35685.1	RELERFMNYTLNNTKNNNVT	SKKKRRFLG
AAB59858.1	RELERFMNYTLNNAKKTNVT	SKKKRRFLG
SwissProtP11	RELERFMNYTLNNTKNTNVT	SKKKRRFLG
SwissProtP13	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
SwissProtP12	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
BAA00240.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
BAA00105.1	RELERFMNYTLNNTKNTNVT	SKKKRRFLG
PRF226438	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
D00334.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
D00151.1	RELERFMNYTLNNTKNTNVT	SKKKRRFLG
Z26524.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
I25351.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
M22643.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
NC_001781.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
AF013254.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
ACY68435.1	RELERFMNYTLNNAKNTNVT	SKKKRRFLG
AA893666.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG
AAO72325.1	RELERFMNYTLNNTKNTNVT	SKKKRRFLG
ACO83302.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
AAA47410.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
AEC32087.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
AD295786.1	RELERFMNYTLNNTKKTNVT	SKKKRRFLG
NP_056863.1	REAPQYMNYTINTTKNLNVS	SKKKRRFLG

A região 130 a 137 está destacada mostrando o possível sítio de ligação com a Heparina Sulfatada conservado em todas as seqüências selecionadas. Fonte: CRAVO, H. L. P., 2012.

A região 130 a 137 é formada principalmente por resíduos de Arginina (R) e Lisina (K), o que é esperado para motivos consenso deste tipo. Estes resíduos são caracterizados por serem resíduos básicos, com carga líquida positiva em pH fisiológico. Desta forma, uma grande quantidade destes resíduos concentrados em uma região cria centros de alta densidade de carga positiva, que por sua vez, interagem eletrostaticamente com os grupos ácidos da Heparina, que possuem carga negativa. Estas características e indícios experimentais reforçam a idéia de este ser um peptídeo de interação com a Heparina (CARDIN & WEINTRAUB, 1989).

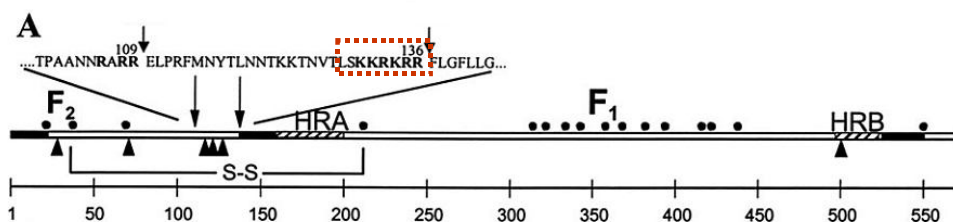
A Heparina é um polímero linear carregado que se comporta como um polieletrólito de carga negativa. Esta característica faz com que exista uma forte energia de repulsão agindo sobre sua estrutura molecular, como consequência da presença dos diversos grupos carregados negativamente. Essa energia favorece a ligação da Heparina a aminoácidos básicos, uma vez que a alta densidade de carga positiva estabiliza os efeitos de repulsão na molécula. Sendo assim, a ligação da Heparina com a região 130 a 137 da proteína F é favorecida, já que a interação da concentração das cargas positivas desta região da proteína

com os grupos negativos da Heparina irá minimizar os efeitos de repulsão sofridos pela molécula.

Outros estudos já evidenciaram experimentalmente este motivo como sendo potencial região inibidora da ação viral, como os estudos de Crim e colaboradores (2006). Eles sugeriram potenciais sítios de ligação entre a proteína F e a Heparina, observando regiões que abrangem o motivo consenso [S-K-K-R-K-R-R-F], como o verificado através da análise de bioinformática realizada neste estudo. Curiosamente eles notaram que o peptídeo contendo a região 130 a 137, quando inibido na presença de Heparina, também impedia o vírus de se ligar à membrana celular. Uma vez que a ligação do vírus com a célula hospedeira é feita pela proteína G do hRSV, não se espera que esta ação do vírus seja bloqueada por regiões da proteína F. Entretanto, sugerimos que existe uma ação interrelacionada entre as proteínas de membrana F e G, que podem atuar concomitantemente neste processo da infectividade viral.

A região 130 a 137, que compreende o possível sítio de ligação, está localizada entre os 27 resíduos de aminoácidos em que ocorre a clivagem da unidade precursora F0 (Figura 7). A proteína F é clivada por proteases celulares na rede trans-Golgi e dá origem ao heterodímero F2-F1, que são duas subunidades protéicas ligadas por pontes dissulfeto. A clivagem libera o chamado peptídeo de fusão, localizado logo após o sítio II de clivagem da proteína F (GONZALEZ-REYES et al., 2001), sendo que este sítio é justamente o motivo encontrado como possível região de ligação com a Heparina Sulfatada. Estas evidências sugerem uma explicação para ação inibidora da Heparina sobre o vírus.

Figura 7 - Esquema do genoma da proteína F do vírus hRSV.



As setas indicam o ponto de clivagem da proteína F0 e em destaque, temos a possível região de interação da proteína com a Heparina Sulfatada. (modificada de GONZALES-REYS et al., 2001).

Estudos *in vitro* mostram que, quando a clivagem dos monômeros está completa, a proteína F sofre uma mudança conformacional que expõe o peptídeo de fusão (MARTÍN et al., 2006). Acredita-se que este peptídeo inicie o processo de fusão pela interação direta com a

bicamada lipídica da membrana da célula hospedeira, sendo a clivagem da unidade precursora F0 essencial para ativar a capacidade de fusão da proteína F (GONZALES-REYS et al., 2001). A possível interação da Heparina com o sítio II de clivagem da unidade F0 impediria o reconhecimento deste sítio pela protease do tipo furina, impedindo a clivagem da unidade F0, evitando a infecção celular pelo vírus. Esta é uma possível explicação para a diminuição do número de células infectadas quando são tratadas com Heparina.

Este passo da infecção viral é tão importante que estudos já verificaram que a proteína F, além da sua funcionalidade fusão, também é responsável por facilitar a ligação do vírus à membrana celular (FELDMAN, AUDET & BEELER, 2000). Estudos realizados com mutantes do vírus hRSV com a proteína F como sua única proteína de superfície, mostraram que a ligação do vírus com a célula ocorre através das moléculas de Heparina Sulfatada da superfície celular e que esta interação possui um papel crucial na infectividade do vírus. Embora a natureza exata destas interações ainda não tenha sido totalmente esclarecida, é possível que a proteína F reforce o potencial de ligação do vírus, aumentando assim a probabilidade de interações subseqüentes necessárias para a infecção (FELDMAN et al., 2000).

4.1.2 – Predição da Estrutura Secundária, modificações pós-traducionais e características da proteína de Fusão

A predição da estrutura secundária da proteína foi realizada pelos programas PREDICTPROTEIN e PSIPRED. Os resultados de ambos os programas concordam entre si, indicando prevalência de alfa hélices, seguida de regiões em *coil* e menor prevalência de folhas beta. A porcentagem informada pelo programa PREDICTPROTEIN pode ser vista na tabela 1.

Tabela 1: Quadro de proporcionalidade das Estruturas Secundárias da Proteína F, de acordo com a análise do programa PREDICTPROTEIN.

Tipo de Estrutura	α-Hélice	Folha-β	Coil
Porcentagem na Proteína	39,55%	28,22%	32,23%

Com o objetivo de entender melhor quais os tipos de modificações que a proteína de fusão pode ter, foram analisados pelo programa PREDICTPROTEIN possíveis sítios de glicosilação e fosforilação. A N-glicosilação é uma das modificações pós-traducionais mais comuns, mas também uma das mais complexas, pois está envolvida no processo de

enovelamento, secreção de proteínas, tráfego intracelular e comunicação celular. A evidência da importância da N-glicosilação na regulação da estrutura e função de glicoproteínas é incontestável (SKROPETA, 2009).

A análise realizada por este programa resultou nos sítios de N-glicosilação: 27, 70, 116, 120, 126 e 500, em que os açúcares se ligam a um resíduo de Asparagina. As posições encontradas foram indicadas em todas as seqüências do nosso banco de dados. De acordo com o alinhamento realizado, estes sítios são conservados em todas as seqüências selecionadas e concordam com evidências experimentais. Os cinco primeiros sítios fazem parte da subunidade F2, enquanto que o último sítio fica localizado adjacente à região transmembrânica da proteína. Deste conjunto de seis sítios, apenas três deles são efetivamente glicosilados: 27, 70 e 500 (SKROPETA, 2009). Os sítios 120, 126 e 116 são sítios que não são encontrados em outras proteínas de fusão de outros vírus da ordem *Mononegavirales*. Eles são localizados adjacente ao peptídeo de fusão (Pep27), provavelmente para auxiliar a clivagem da proteína, formando um peptídeo maduro (ZIMMER, TROTZ & HERRIER, 2001). Poucos estudos relataram qual o papel da N-glicosilação para a proteína F do hRSV. Um estudo feito por McDonald e colaboradores (2006) mostrou que o sítio 500 tem um papel importante na mediação da atividade de fusão da proteína F, e a glicosilação dos sítios 27 e 70 pode ter um pequeno papel na habilidade de ligação da subunidade F2 (MCDONALD et al., 2006).

Nenhuma das N-ligações são necessárias para a clivagem da proteína F ocorrer. Quando se faz uma interconversão da Asparagina dos sítios 27, 70 e 500 em Glicinas, eliminando a adição do açúcar, em qualquer combinação, a síntese da proteína não é afetada de nenhuma forma. Entretanto, a atividade de fusão da proteína é influenciada de diferentes maneiras. Quando o açúcar do sítio 70 é eliminado, ocorre o aumento da atividade de fusão em mais de 40%. Contudo, o mesmo não ocorre se apenas o sítio 27 é eliminado, apesar da remoção dos dois sítios aumentar a atividade de fusão em 50%. Quando o sítio de glicosilação 500 é inibido, poucos sincícios são formados, bem como a atividade de fusão também é afetada. Esta atividade é reduzida a aproximadamente 10% com a remoção do sítio 500 ou removendo os pares de sítios 500 e 27, e 500 e 70. A eliminação dos três sítios abole definitivamente a atividade de fusão (ZIMMER, TROTZ & HERRIER, 2001).

A razão pela qual o sítio 500 é importante se deve a sua localização na região *heptad repeat B*, indicando que esta região precisa da maturação do glicano para a formação correta do feixe de hélices. Acredita-se que os outros sítios também possam ser necessários para proteção da proteína contra o reconhecimento de anticorpos, manter sua conformação correta

ou ainda, permitir a mudança conformacional pós processamento proteolítico. Funções similares a estas foram propostas para três sítios de glicosilação conservados do vírus influenza (OHUCHI et al., 1997).

O programa PREDICTPROTEIN ainda forneceu dados de sítios de fosforilação que estão dispostos na tabela 2.

Tabela 2: Sítios de fosforilação da Proteína F separados pelo tipo de enzima que executa o processo, segundo a análise do programa PREDICTPROTEIN.

Quinase C	Caseína Quinase II	Tirosina Quinase II
130		
174		
189	215	
337	219	
357	253	191
362	335	
397	398	
425		
443		

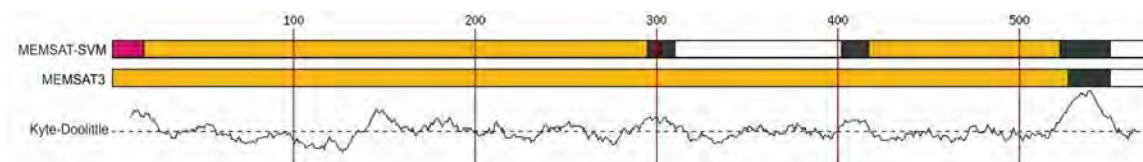
A conservação destes sítios, além de demonstrar sua relevância para ação da proteína, ainda reforça a idéia da importância da modificação conformacional sofrida pela proteína F após o processo de clivagem. Uma vez que este é um mecanismo fundamental para o processo de adsorção viral, a sua perpetuação é fundamental para a capacidade de reprodução do vírus.

As proteínas quinases são importantes para transdução do sinal. A existência destes sítios na proteína F sugere uma função de transdução de sinais que apresenta importante papel na infectividade e replicação do vírus respiratório. Além disso, sítios de fosforilação indicam possíveis pontos de modificações conformacionais pelas quais a proteína sofre ao ser fosforilada. A grande quantidade destes sítios indicada pelo programa pode direcionar novos estudos que visem compreender o real papel das modificações conformacionais da proteína F no processo de infecção viral.

A topologia da proteína de fusão foi obtida pelo programa PSIPRED. A figura 8 mostra regiões do domínio extracelular, a região do citoplasma e a região transmembrânica. A predição indica os aminoácidos de 527 a 550 como a região ancorada na membrana lipídica do vírus. Como em outros vírus da família *Paramyxoviridae*, a região transmembrânica é localizada na extremidade C-terminal da proteína e é adjacente ao peptídeo de fusão. Acredita-se que a presença de duas seqüências apolares próximas uma da outra permite que a proteína de fusão interaja simultaneamente com a membrana viral e a membrana celular, se

aproveitando da característica hidrofóbica destas regiões, para promover uma rápida e eficiente fusão das duas membranas (WHITE, 1990).

Figura 8 - Topologia da proteína F.



A previsão é realizada por dois métodos, representados pelas duas fitas da figura. A região rosa indica o peptídeo sinal (aminoácidos 1-18), as regiões em laranja representam o ectodomínio, as regiões brancas representam o domínio intracelular e as regiões transmembranas são representadas em preto. A topologia é mostrada pelo diagrama de Kyte-Doolittle, que mostra regiões hidrofóbicas (acima da linha pontilhada) e hidrofílicas (abaixo da linha pontilhada). Fonte: CRAVO, H. L. P., 2012.

4.2 – Análise, otimização e validação do modelo tridimensional da proteína F

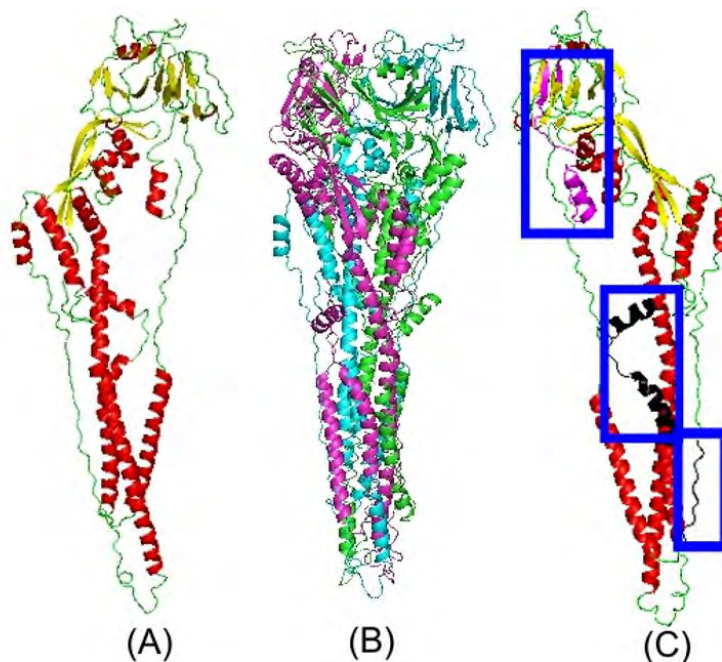
O modelo que melhor atendeu as expectativas para ser utilizado neste estudo foi o construído pelo método de *threading*, gerado pelo programa I-TASSER. O modelo produzido por modelagem comparativa, por meio do programa MODELLER, foi construído de forma incompleta, não apresentando estruturas modeladas das regiões N-terminal e C-terminal. Uma vez que este estudo procura por sítios de interação da proteína de fusão com possíveis compostos inibidores, entendemos que nenhuma região pode ser excluída. Para uma melhor análise dos resultados, é necessário que o modelo esteja completo e contenha todas as estruturas da forma mais similar possível a situação *in vivo*. Dessa forma, o modelo foi descartado.

O modelo construído pelo programa ROSETTA não condizia com as características de uma proteína de fusão, estando em total desacordo com o que já foi descrito na literatura sobre a estrutura destas proteínas. O programa e o método *Ab Initio* ainda apresentam limitações para os cálculos de estruturas tridimensionais de proteínas com muitos aminoácidos (proteínas com mais de 200 aminoácidos) (ZHANG, 2009). Uma vez que a proteína F possui 574 aminoácidos, era esperado que os modelos gerados por esta técnica apresentassem inconsistências. Pelos modelos não satisfazerem as características da proteína, este método também foi descartado. Os modelos descartados podem ser vistos nos apêndices A e B.

Quando modelos por homologia não podem ser construídos, a modelagem por *threading* é a técnica que fornece modelos mais confiáveis. Este tipo de método tem produzido previsões excepcionalmente boas e são muito utilizadas atualmente para previsão da estrutura de uma proteína. Em particular, seu baixo custo computacional torna este método muito vantajoso, quando se quer estudar o genoma inteiro de um organismo (GADJA, PAWLOWSKI & BUJNICKI, 2011).

Como pode ser visto na figura 9, o modelo escolhido apresenta características estruturais como os já descritos na literatura para proteínas de fusão dos *Paramyxovirus*. O monômero é dividido em partes chamadas cabeça, pescoço e caule. A cabeça e o pescoço são regiões compostas tanto por sequências da subunidade F1, quanto da subunidade F2. O caule é composto de estruturas *coiled-coil* e inclui a região C terminal. A proteína forma um trímero que visto por cima, tem forma triangular (MORRISON, 2003). O trímero foi formado com o programa WinCoot, utilizando como monômero o modelo da proteína otimizado.

Figura 9 - Modelo da proteína F



(A) Representação do monômero da proteína colorido de acordo com suas estruturas secundárias. (B) Representação do trímero formado pelo modelo visto em (A), gerado pelo programa Wincoot. (C) Monômero da proteína F com destaque para a região N-terminal (rosa), C-terminal (preto) e região 130 a 137 (azul) Fonte: CRAVO, H. L. P., 2012.

O modelo também está de acordo com a previsão da estrutura secundária dos programas PREDICTPROTEIN e PSIPRED. Podemos observar uma predominância de alfa

hélices e *coils*, com pouca presença de folhas beta. O programa PDBSum quantificou as estruturas secundárias do modelo, como pode ser observado na tabela 3. Os resultados dos programas e do modelo concordam com os resultados pela técnica de Dicroísmo Circular da proteína F feito por Arcuri e colaboradores (2007), em que a ocorrência de alfa hélice e de estruturas randômicas (*coils*) é superior as de folha beta.

Tabela 3: Quadro de proporcionalidade das Estruturas Secundárias da Proteína F, de acordo com a análise do programa PDBSum.

Tipo de Estrutura	α -Hélice	Folha- β	Coil
Porcentagem na Proteína	39,9%	19,9%	40,2%

A otimização do modelo foi realizada inicialmente pelo programa WinCoot. O programa possui um pacote de validação em que os erros do modelo são indicados e podem ser corrigidos por novos cálculos feitos pelo programa ou mesmo pelo usuário. A partir da análise do gráfico de Ramachandran, o posicionamento dos átomos dos aminoácidos, bem como os valores dos ângulos *psi* e *phi* da cadeia polipeptídica principal, foram recalculados adequando o modelo aos valores permitidos, uma vez que devido aos impedimentos estéricos, a cadeia polipeptídica principal assume, preferencialmente, valores energeticamente favoráveis. Nenhuma alteração perceptível no esqueleto da proteína ou na estrutura modelada pelo programa I-TASSER foi verificada, mostrando que a estrutura apenas teve seus erros reparados.

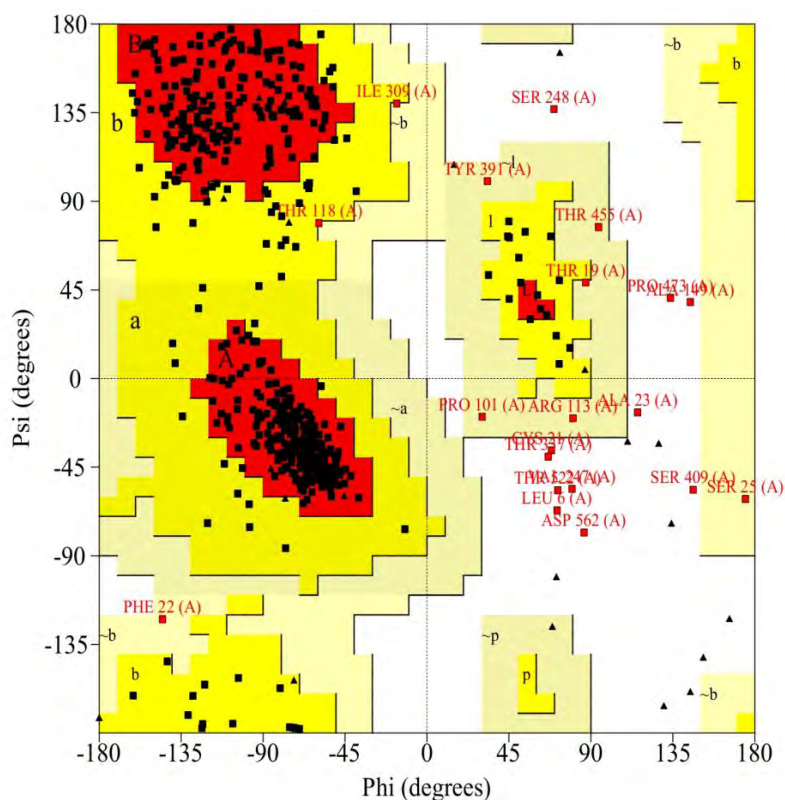
A segunda parte da otimização envolveu a minimização dos choques estéricos no modelo da proteína por meio do programa Chiron. A estrutura inicial possuía uma energia inicial de choques igual a 247,218 Kcal/mol e após a aplicação do protocolo de minimização a energia passou para 38,553 Kcal/mol. Para classificar a minimização como aceitável, os autores definiram um *score* como sendo a energia total de choques da proteína dividida pelo número de contatos. Assim, este *score* descreve os choques estéricos presentes na estrutura da proteína, mas é independente do seu tamanho. A energia total de choques da proteína é definida como a soma da energia de repulsão de Van der Waals de todos os choques estéricos na estrutura da proteína (RAMACHANDRAN et al., 2011).

A partir da análise da distribuição do *score* de choques de um conjunto de proteínas resolvidas experimentalmente com alta resolução, o *score* aceitável para uma proteína é de 0,02 Kcal/mol/contato. O modelo da proteína F em sua estrutura final teve um *score* de choque de 0,00439252 Kcal/mol/contato, indicando que a estrutura foi otimizada pelo

programa. O número de choques iniciais era de 269 choques e este número caiu para 77 choques não resolvidos pelo método.

A validação da estrutura feita pelo gráfico de Ramachandran pode ser vista na figura 10. De acordo com o gráfico, a estrutura possui 97,4% de aminoácidos em regiões permitidas, 1,5% de aminoácidos em regiões generosamente permitidas e 1,9% de aminoácidos na região não permitida. Estruturas válidas para o gráfico de Ramachandran possuem acima de 90% de aminoácidos na região permitida e a menor porcentagem possível nas regiões não permitidas. Uma vez que a proteína de fusão é uma proteína com muitos aminoácidos (574 resíduos), a porcentagem de 1,9% na região não permitida é aceitável, uma vez que representa uma pequena porção da proteína. Para modelos de baixa resolução, como os gerados por técnicas de modelagem, espera-se que a maioria dos resíduos estejam em regiões favoráveis e até 1% na região não favorável (EMSLEY et al., 2010).

Figura 10 - Gráfico de Ramachandran para o modelo final da proteína F.



A análise foi gerada pelo programa PROCHECK. A soma do número de aminoácidos em regiões permitidas e generosamente permitidas é de 98,9%. Apenas 1,9% dos resíduos estão em regiões desfavoráveis, o que caracteriza um bom número para uma proteína de muitos aminoácidos como a proteína F. Fonte: CRAVO, H. L. P., 2012.

O programa ProQ-server resultou em um LGScore de 1,629 e o *score* MaxSub de 0,226. Um valor LGScore acima de 1,5 classifica o modelo como bom e um *score* MaxSub acima de 0,1 classifica o modelo como bom. O programa fez a análise por uma combinação de características estruturais da proteína como: contatos átomo-átomo, resíduo-resíduo, área superficial exposta e concordância com a estrutura secundária predita, que foi usada como dado de entrada para a rede neural treinada, prever a qualidade do modelo (WALLNER & ELOFSSON, 2007).

4.3 – Análise das Estruturas Tridimensionais dos Flavonóides

Foram geradas 34 estruturas dos Flavonóides cedidos pelo Prof. Dr. Wagner Vilegas do Instituto de Química da UNESP de Araraquara. Todos os compostos foram desenhados em formato SMILE (1D) e convertidos para as estruturas tridimensionais (3D) pelo programa FROG2. Este é um método utilizado desde o início da década de 90 para auxílio na determinação de novos fármacos e na técnica de *Virtual Screening*. A modelagem inicial procura identificar certas características como anéis, ordens de ligação e centros estereoquímicos na molécula para gerar uma conformação de baixa energia (LEACH & GILLET, 2007). As agliconas, estruturas básicas sem glicosilação, são os únicos flavonóides com estruturas resolvidas experimentalmente. Os que possuem estrutura disponível são: Quercetina, Miricetina, Kaempferol e Apigenina. Estas estruturas foram usadas para validar as outras estruturas modeladas.

Para a validação, os compostos que já possuem estruturas resolvidas experimentalmente também foram modelados e sobrepostos as estruturas experimentais para o cálculo do desvio médio quadrático (RMSD).

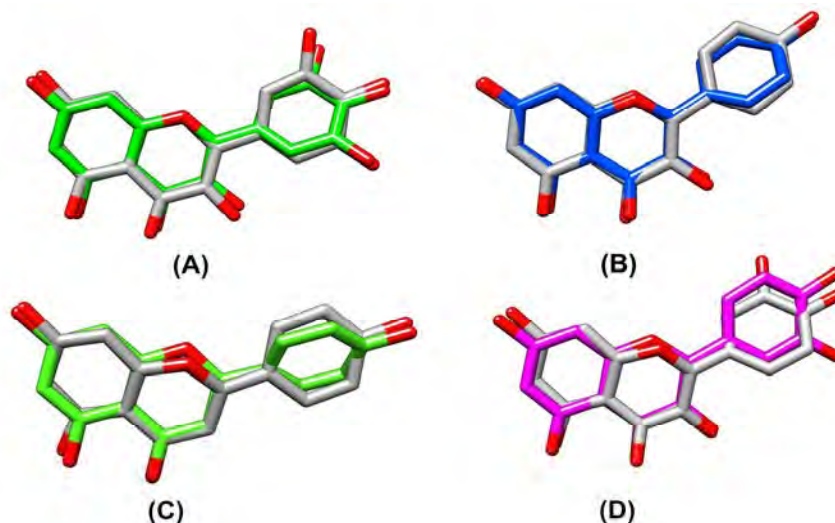
Tabela 4: Valores do RMSD da estrutura experimental de alguns flavonóides e a estrutura modelada pelo programa FROG2

Composto	RMSD (Å)
Quercetina	0,4276
Miricetina	0,6601
Kaempferol	0,4283
Apigenina	0,5458

Como podemos observar pela tabela 4, os valores do RMSD são pequenos e próximos de zero, o que garante diferenças pequenas entre as coordenadas das estruturas experimentais e dos modelos. Bons modelos apresentam valores de RMSD menores que 2 Å. Além disso, podemos observar na figura 11 as estruturas modeladas sobrepostas às estruturas resolvidas

experimentalmente. Podemos notar que as diferenças entre as estruturas são pequenas, garantindo que os modelos estejam de acordo com os valores esperados para estas estruturas.

Figura 11 - Comparação das estruturas modeladas dos flavonóides com as estruturas experimentais.



Em todas as imagens, branco representa as estruturas experimentais e as estruturas modeladas são representadas em verde, azul, verde e rosa, respectivamente. (A) Miricetina, (B) Kaempferol, (C) Apigenina, (D) Quercetina. Pequenas diferenças são observadas entre a estrutura experimental e estrutura modelada, o que garante a confiabilidade dos outros modelos. Fonte: CRAVO, H. L. P., 2012.

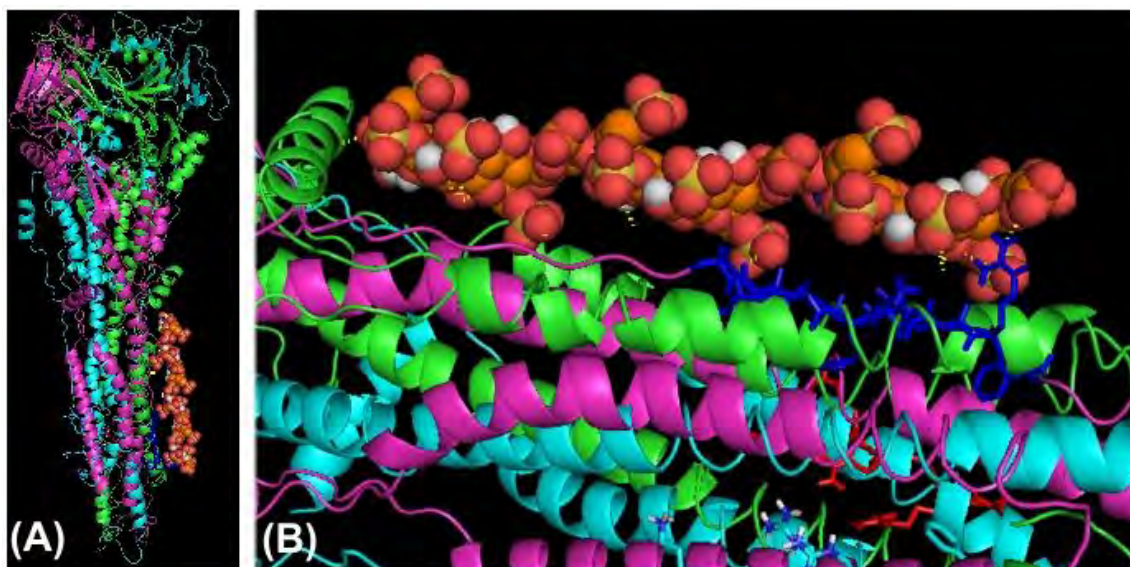
4.4 - Docking Molecular do Modelo da Proteína F com a Heparina

Dada as evidências experimentais da interação do hRSV com a Heparina, o *docking* molecular entre o modelo da proteína F e a estrutura resolvida da Heparina foi proposto para avaliar a interação. Uma vez que elementos estruturais podem influenciar na ligação entre dois compostos, apenas o estudo de possíveis seqüências de aminoácidos favoráveis a ligação com a Heparina não é suficiente para determinar se a ligação é possível.

O *docking* realizado pelo programa AUTODOCKVINA resultou em uma energia de -14,1 Kcal/mol, que representa um valor satisfatório para uma interação da Heparina com proteínas. Um estudo realizado por Gandhi e Mancera (2009) relatou que a média da energia de interação de proteínas com glicosaminoglicanos era de aproximadamente -13kcal/mol. A interação ocorreu no sítio 130 a 137, como indicado previamente por nosso estudo da busca de possíveis motivos de interação seqüencial entre a proteína e a Heparina (Figura 12). Este resultado vem reforçar a hipótese da explicação pela qual a infectividade do vírus diminui na presença da Heparina, como evidenciado por Rubio e colaboradores (2010). Uma vez que o sítio de clivagem II fica impedido para a ação das enzimas do tipo furina, a clivagem

necessária para obtenção da atividade de fusão não acontece, deixando a proteína inativa. Desta forma, o vírus não consegue ser adsorvido pela célula do hospedeiro.

Figura 12 - Modelo de interação entre a Heparina e a proteína de fusão.



(A) Visão geral do modelo de interação da proteína F com Heparina. Podemos observar o trímero (monômeros em roxo, verde e azul) interagindo com Heparina (esferas em vermelho, branco, laranja e amarelo) na região de 131-137 em azul. (B) Imagem ampliada da região de interação. Há ligações de hidrogênio entre o composto e proteína. Um deles é formado entre Heparina e o aminoácido Lys 131 e três ligações de hidrogênio foram estabelecidas com Arg 136. Esses dois resíduos pertencem à região do sítio de clivagem II da proteína, necessária para aquisição da atividade de fusão. Fonte: CRAVO, H. L. P., 2012.

As propriedades medicinais da Heparina são estudadas há muito tempo, com evidências de suas propriedades antivirais que datam desde a década de 60. Suas propriedades inibitórias já foram verificadas para o HIV, vírus da dengue, Hepatite C e a maioria dos vírus envelopados. Em geral, principalmente para o vírus Parainfluenza, vírus envelopados são susceptíveis a ação inibidora de glicosaminoglicanos sulfatados, como é o caso da Heparina, uma vez que ela impede a ligação destes vírus com os glicosaminoglicanos presentes na superfície celular, evitando assim o início da adsorção viral (BABA et al., 1988).

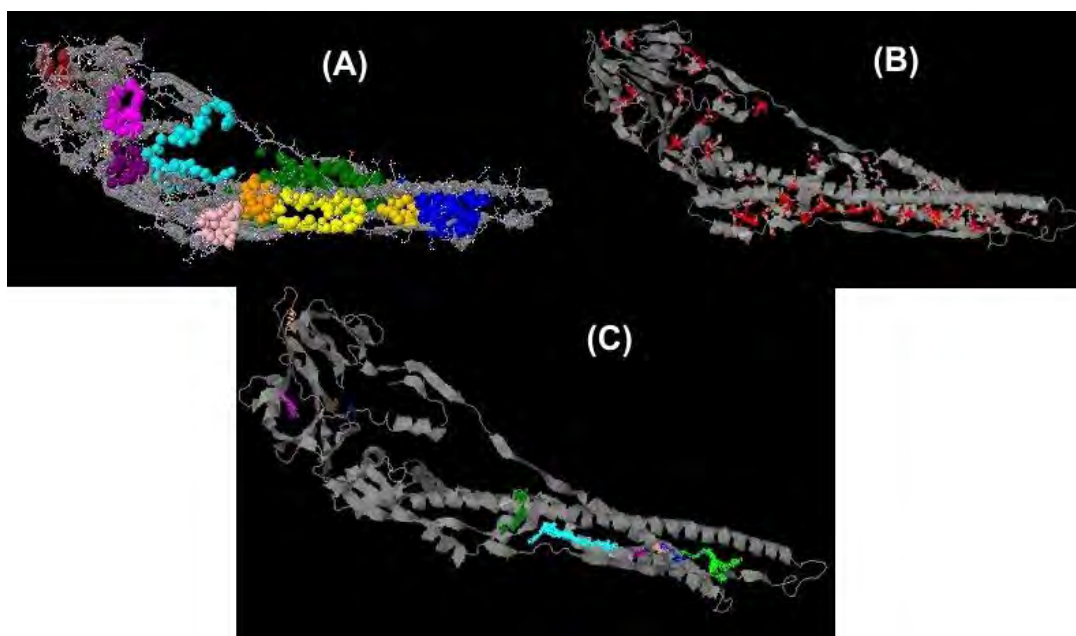
4.5 - Docking Molecular entre a Proteína F e Flavonóides

4.5.1 - Determinação de Sítios de Interação na Proteína de Fusão

O modelo da proteína de fusão foi submetido a três programas de predição de sítios de ligação: QSiteFinder, fpocket e CASTp. A predição foi feita para uma análise prévia de

possíveis regiões de interação da proteína que possam inibir a sua atividade. Os programas indicaram vários sítios de interação, sendo que os de maior importância são os primeiros da listagem dada nos resultados. A região indicada com a melhor probabilidade de interação é a dos resíduos 190 a 225. A segunda melhor região de interação é a dos resíduos 130 a 137, como já previamente identificado por nosso estudo pelo alinhamento das sequências da proteína F. Para os outros sítios indicados pelo programa, como em regiões de fosforilação e glicosilação, não foi evidenciada relevância na interrupção da atividade da proteína, como a o sítio de clivagem II ou mesmo da região da formação dos feixes de hélice da proteína. Na figura 13 podemos observar uma imagem das predições feitas pelos programas.

Figura 13 - Predição de sítios de interação da proteína F.



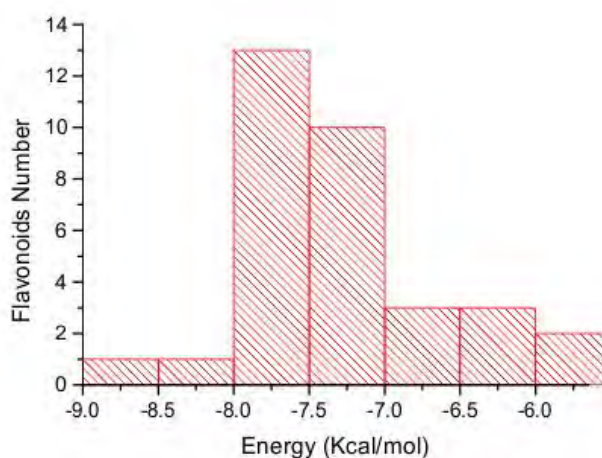
(A) Predição feita pelo programa CASTp. (B) Predição feita pelo programa fPocket. (C) Predição feita pelo programa QSite-Finder. Em todas as figuras notamos que sítios semelhantes foram encontrados como prováveis regiões de interação. Os três programas indicaram a região 130 a 137 como uma região favorável a interação com ligantes. Fonte: CRAVO, H. L. P., 2012.

4.5.2 - Interação entre o Modelo da Proteína de Fusão e Flavonóides: *Docking* Cego

No *docking* realizado entre o modelo inteiro da proteína de fusão e os flavonóides modelados, sem direcionar o composto para uma região específica, todas as estruturas interagiram com a proteína, sendo que a maioria se ligou no mesmo sítio de interação, que corresponde à região dos resíduos 190 a 225, como previamente indicado pelos programas de

predição de sítios de interação. Esta região não possui nenhuma importância significativa descrita na literatura relacionada à função da proteína. No entanto, de acordo com os resultados descritos acima, preditos pelo programa PREDICTPROTEIN, esta é uma região de fosforilação por enzimas do tipo Caseína Quinase II, o que sugere que as interações de compostos nessa região podem atrapalhar as mudanças conformacionais que a proteína sofre ao ser clivada para obtenção da sua atividade de fusão.

Gráfico 14 - Histograma da quantidade de flavonóides por energia de interação.



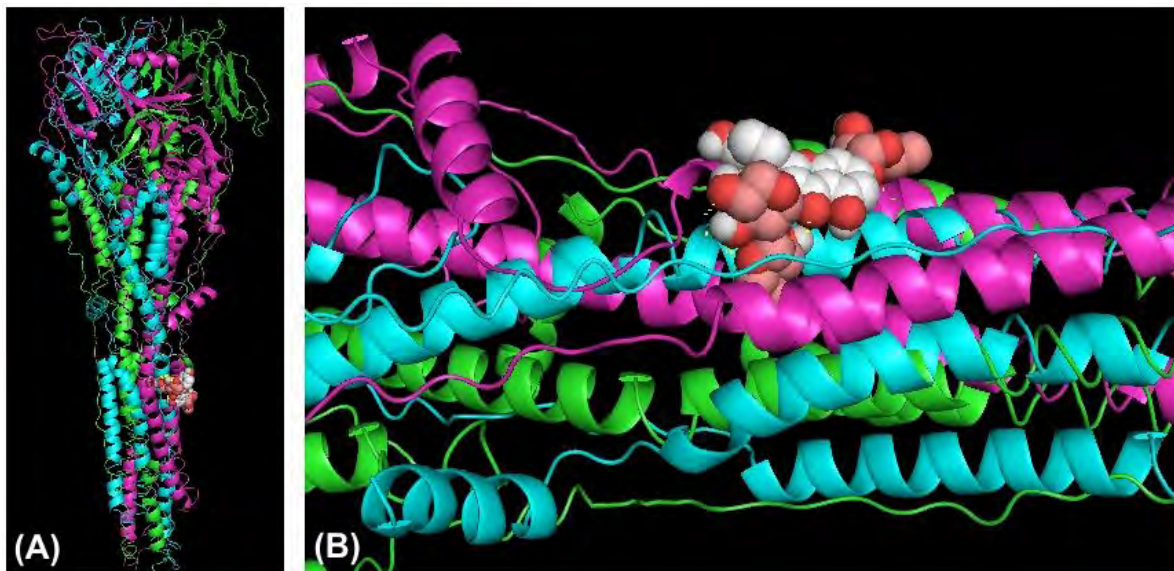
As barras representam a quantidade de flavonóides que interagiram com a proteína de fusão em uma determinada faixa de energia. Fonte: CRAVO, H. L. P., 2012.

A distribuição das faixas de energia de interação entre os flavonóides e a proteína de fusão pode ser observada no histograma do gráfico 14. Estudos experimentais, que já avaliaram a interação de flavonóides com outras proteínas, constatam que estes valores se enquadram na faixa de energia de uma interação favorável de flavonóides com proteínas. Um estudo da interação da Quercetina com a proteína Lisozima determinou uma energia de interação de -6,0 a -7,0 Kcal/mol (WANG et al., 2011). Em outro estudo, em que a interação da Quercetina foi avaliada com a proteína Lipoxigenase3, a energia de interação resultante foi de -6,9 Kcal/mol (FIORUCCI et al., 2008). Desta forma, as energias indicadas neste estudo, sugerem uma provável interação entre os flavonóides e a proteína F do hRSV, que pode ser melhor investigada a partir de ensaios experimentais.

O composto com melhor energia de interação foi Kaempferol-3-O- α -L-Arabinopiranosil-(2 \rightarrow 1)- α -L-Apiofuranosídeo-7-O- α -L-Ramnopiranosídeo, com uma energia de interação de -8,8 Kcal/mol. O Kaempferol sem suas glicosilações foi estudado como

inibidor da Tripsina e obteve uma energia de interação de -6,26 Kcal/mol (CHECA et al., 1997). Para a enzima Acetilcolinesterase, enzima envolvida na doença de Alzheimer, a interação com o Kaempferol foi de -8,82 Kcal/mol (PARAMESWARI et al., 2011). A interação do Kaempferol com a proteína F é mostrada na figura 15.

Figura 15 - Modelo da interação do flavonóide com menor energia de interação e a proteína de fusão.



(A) Visão geral da interação e da posição do flavonóide interagindo com o modelo. Pode ser visto o trîmero com seus monômeros em roxo, verde e azul. O Kaempferol-3-O- α -L-Arabinopiranosil-(2 \rightarrow 1)- α -L-Apiofuranosideo-7-O- α -L-Rhamnopiranosideo é representado por esferas em amarelo, vermelho e branco. (B) Uma visão ampliada da região de interação, onde observa-se que as linhas amarelas pontilhadas indicam a formação de ligações de hidrogênio entre a proteína e o flavonóide. Fonte: CRAVO, H. L. P., 2012.

4.5.3 - Interação entre o Modelo da Proteína de Fusão e Flavonóides: *Docking* Dirigido

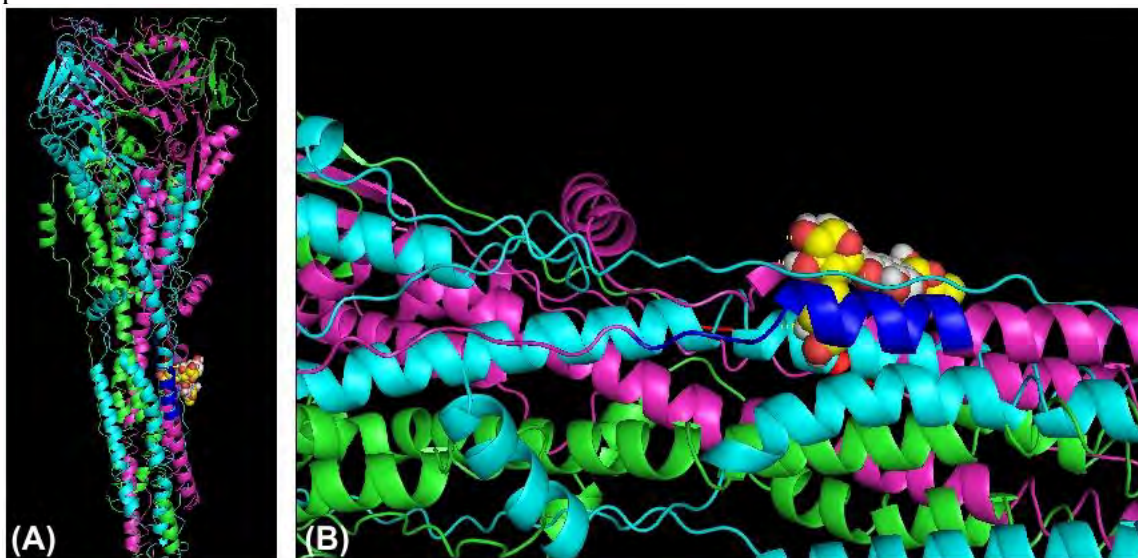
Recentemente, pequenas moléculas tornaram-se o principal objeto de estudo para inibidores da proteína F. Uma série de derivados de benzimidazol (composto orgânico aromático heterocíclico que consiste da fusão do benzeno e imidazol. Estes compostos são representados comumente pelo albendazol e mebendazol, que atuam como antihelmintícos) possuem relatos de interação com a proteína F e, embora possuam estruturas diferentes, todos eles parecem se ligar aos sítios hidrofóbicos das regiões HRA (*heptad repeat* A, resíduos 153-209) e HRB (*heptad repeat* B, resíduos 482-520). A ligação nestes sítios interferiria na interação da HRA com a HRB e na formação do feixe de hélices (6HB) necessário para a finalização da fusão das membranas (COLLINS & MELERO, 2011).

Uma vez que os flavonóides também são pequenas moléculas aromáticas, e a análise da interação em toda a proteína não chegou a uma região que influencia na atividade da proteína, realizou-se um *docking* dirigido às regiões HRA e HRB do modelo da proteína de fusão, uma vez que esta região mostrou-se um sítio importante para a ação de inibidores contra o hRSV.

Assim, como no *docking* cego, todas as moléculas interagiram com a proteína, no entanto desta vez, interagiram especificamente no sítio hidrofóbico, com uma faixa de energia que variou de -7,9 Kcal/mol a -10,4 Kcal/mol. Como já discutido anteriormente, estes valores de energia caracterizam uma interação favorável dos compostos com a proteína, em que análises experimentais poderiam esclarecer melhor estes resultados.

O composto que teve melhor interação nessa região foi Kaempferol-3-O- α -L-Arabinopiranosil-(2 \rightarrow 1)- α -L-Apiofuranosideo-7-O- α -L-Rhamnopiranosideo com uma energia livre de interação de -10,4 Kcal/mol, que como previamente discutido, representa um bom valor de interação entre um composto Kaempferol e uma proteína. A imagem do modelo pode ser vista na figura 16. Foram formados duas ligações de hidrogênio com a Leu 489, resíduo que faz parte da região do feixe de hélices. Este resultado sugere que os flavonóides podem ser utilizados como prováveis inibidores da formação dessa região, influenciando no mecanismo de fusão.

Figura 16 - Modelo a interação dos flavonóides no local da interrupção da atividade da proteína de fusão F.



(A) Uma visão geral da posição do ligante no sítio de ligação. (B) uma imagem ampliada da interação mostrando as ligações de hidrogênio entre o ligante e sítio de ligação. Fonte: CRAVO, H. L. P., 2012.

Acredita-se que a estrutura do feixe de hélices é mantida por uma interação proteína-proteína. Evidências experimentais para outros vírus, que também possuem proteína de fusão similares a proteína F do hRSV, como HIV, Simian Vírus 5 e Visna Vírus, tem demonstrado que pequenas moléculas nessa região interagem com resíduos importantes na formação do feixe, simulando assim a interação que ocorre entre as duas regiões HRA e HRB. Esta interferência das pequenas moléculas desestabiliza a formação do feixe, impedindo a formação da estrutura, o que resulta na perda da atividade de fusão, justificando assim o uso da ação de pequenas moléculas nessa região (CIANCI et al., 2004).

Um trabalho demonstra que a molécula VP-14637 é um inibidor da proteína F que interage neste sítio hidrofóbico e que atua desestabilizando a formação do feixe de hélices (DOUGLAS et al., 2003). Para compararmos a ação de inibição dos flavonóides com a da molécula VP-14637, realizamos o *docking* dirigido da molécula com o modelo da proteína F, sob as mesmas condições realizadas para os flavonóides. A melhor energia de interação da molécula VP-14637 com o modelo da proteína F foi de -7,2 Kcal/mol, valor próximo ao encontrado para a melhor interação do flavonóide com a proteína F, mas que ainda representa uma energia de interação inferior ao encontrado neste trabalho. Este indício sugere que a interação dos flavonóides com a proteína F pode ser mais favorável e representar uma opção melhor para inibição do vírus. Evidentemente, estes resultados necessitam ser mais explorados, através de ensaios experimentais que confirmem os modelos encontrados pelas simulações realizadas neste trabalho.

5 – Conclusões

Por meio dos resultados deste trabalho, podemos concluir que:

1 – As diversas sequências da proteína de Fusão possuem uma alta identidade entre si, o que implica na conservação de regiões importantes para o funcionamento da proteína, como sítios de glicosilação, fosforilação e interação com enzimas. Estas características fazem da proteína F um bom alvo para inibidores que atuem nestes sítios. Em especial, a região 130 a 137, formada pelos resíduos [S-K-K-R-K-R-R-F], se mostrou uma região importante na interação com compostos de carga negativa;

2 – O modelo da proteína F obtido foi válido e satisfaz as características da proteína já relatadas na literatura. A estrutura da proteína possui regiões hidrofóbicas que possibilitam a interação com pequenas moléculas, sendo a região 130 a 137 (indicada previamente pelo estudo sequencial) uma das principais delas;

3 – O *docking* molecular do modelo da proteína de Fusão com a Heparina Sulfatada mostrou que esta interação é possível e favorável, uma vez que a energia de interação obtida foi melhor que a de valores já relatados na literatura. A interação ocorreu no sítio de clivagem II, região 130 a 137 da proteína. Pela técnica do *Virtual Screening* também constatamos que os flavonóides interagem favoravelmente com a proteína de Fusão, representando, assim, moléculas candidatas à inibição de um sítio ativo da proteína. O composto que melhor interagiu foi o Kaempferol-3-O- α -L-Arabinopiranosil-(2 \rightarrow 1)- α -L-Apiofuranosideo-7-O- α -L-Rhamnospiranosídeo. Este flavonóide apresentou alta probabilidade de interação no sítio hidrofóbico do feixe de hélices da proteína, região importante no mecanismo de fusão da proteína. A energia de interação nesta região foi favorável e dentro dos valores esperados para uma interação entre uma proteína e flavonóides.

4 – Este estudo sugeriu que a Heparina Sulfatada e os Flavonóides são bons candidatos a inibidores da proteína de Fusão. A utilização destes compostos em futuros ensaios experimentais esclarecerá melhor estes resultados, demonstrando qual o papel e a eficácia destes compostos sobre a infectividade e ação do vírus.

6 – Referências bibliográficas

- ALTSCHUL, S. F. et al. **Gapped BLAST and PSI-BLAST**: a new generation of protein database search programs. **Nucl. Acids res.**, [S.l.], v. 25, n. 17, p. 3389-3402, 1997.
- ANDRES, A.; DONOVAN, S. M.; KUHLENSCHMIDT, M. S. Soy isoflavones and virus infections. **Journal of Nutritional Biochemistry**, v. 20, p. 563-569, 2009.
- ANDRONICO, A. et al. **Data-Driven High-Throughput Prediction of the 3-D Structure of Small Molecules**: Review and Progress. **Journal of Chemical Information and Modeling**, [S.l.], v. 51, p. 760-776, 2011.
- ARCURI, H. A. et al. Expression and purification of human respiratory syncytial virus recombinant fusion protein. **Protein Expression and Purification**. [S.l.], v. 68, p. 146-152, 2008.
- AVEDAÑO, L. F.; PARRA, C.; PALOMINO, M. A. **Impacto en salud infantil del invierno 2002**: disociación entre factores ambientales y virus respiratorio sincicial, en Santiago. **Rev Méd**, v. 8, 2003.
- BABA, M. et al. Sulfated polysaccharides are potent and selective inhibitors of various enveloped viruses, including herpes simplex virus, cytomegalovirus, vesicular stomatitis virus, and human immunodeficiency virus. **Antimicro. Agents Chemother.**, [S.l.], v. 32, n. 11, p. 1742-1745, 1988.
- BARROSO, R. C. et al. Evaluation of the protamine in the neutralization of heparin after cardiopulmonary bypass. **Rev Bras Circ Cardiovasc**, v. 17, n.1, 2002.
- BARROWCLIFFE, T. W. Heparin assay and standardization. In: Lane, D. A.; Linsdahl, V. Heparin. **Chemical and Biological properties, Clinical Applications**. C.R.C. Press, Boca Raton, 1989.
- BENSON, D. A. et al. Genbank. **Nucl. Acids. Res.**, [S.l.], v. 25, n. 1, 1997.
- BINKOWSKI, T. A.; NAGHIBZADEH, S.; LIANG, J. **CASTp**: Computed Atlas of Surface Topography of proteins. **Nucl. Acids. Res.**, [S.l.], v. 31, n. 13, p. 3352-3355, 2003.
- BLACK, C. P. Systematic Review of the Biology and Medical Management of Respiratory Syncytial Virus Infection. **Respiratory Care**, [S.l.], v. 48, n. 3, p. 209-233, 2003.
- BROOKS, G. F.; BUTEL, J. S.; MORSE, S. A. **Microbiologia Médica**. [S.l.]: Guanabara Koogan, 2000.
- BUCHAN, D. W. et al. Protein annotation and modelling servers at University College London. **Nucl. Acids. Res.**, [S.l.], v. 38, p. 563-568, 2010.
- BYLKA, W.; MATLAWASKA, I. Natural Flavonoids as Antimicrobial Agents. **JANA**, v. 7, p. 24-31, 2004.
- CAPILA, I; LINHARDT, R. J. Heparin-protein interactions. **Angewandte Chemie**. [S.l.], v. 41, p. 391-412, 2002.

- CARDIN, A. D.; WETRAUB, H. J. R. Molecular modeling of protein-glycosaminoglycan interactions. **Arteriosclerosis**, [S.l.], v. 9, p. 21-32, 1989.
- CASU, B. Structure and biological activity of mammalian glycosaminoglycans. **Mod. Probl. Pharmacopsychiatry**, [S.l.], v. 23, p. 56-67, 1989.
- CHECA, A. et al. **Assessment of Solvation Effects on Calculated Binding Affinity Differences: Trypsin Inhibition by Flavonoids as a Model System for Congeneric Series.** **J. Med. Chem.**, [S.l.], v. 40, n. 25, p. 4136-4145, 1997.
- CIANCI, C. et al. **Targeting a binding pocket within the trimer-of-hairpins: Small-molecule inhibition of viral fusion.** **PNAS**, [S.l.], v. 101, n. 42, p. 15046-15051, 2004.
- CINTRA, O. A. L. et al. Occurrence and severity of infections caused by subgroup A and B respiratory syncytial virus in children in Southeast Brazil. **Journal of Medical Virology**, [S.l.], v. 65, p. 408-412, 2001.
- COLLINS, P. L.; GRAHAM, B. S. Viral and Host Factors in Human Respiratory Syncytial Virus Pathogenesis. **J. Virol.**, [S.l.], v. 82, n. 5, p. 2040-2055.
- COLLINS, P. L.; MELERO, J. A. **Progress in understanding and controlling respiratory syncytial virus: Still crazy after all these years.** **Virus Research**, [S.l.], v. 162, p. 80-99, 2011.
- CORNE, D. W.; MARTIN, A. C. Artificial Intelligence in Bioinformatics. **Computers & Chemistry**, [S.l.], v. 26, p. 1-3, 2001.
- COWAN, M. M. Plant Products as Antimicrobial Agents. **Clinical Microbiology Reviews**, v. 12, p. 564-582, 1999.
- CRISTOBAL, S. et al. A study of quality measures for protein threading models. **BMC Bioinformatics**, [S.l.], v. 2, 2001.
- DIMITROV, D.S. **Virus Entry: Molecular Mechanisms And Biomedical Applications.** In: **Nature Reviews**, v. 2, p. 109-122, 2004.
- DOUGLAS, J. L. et al. Inhibition of Respiratory Syncytial Virus Fusion by the Small Molecule VP-14637 via Specific Interactions with F Protein. **J. Virol**, [S.l.], v. 77, n. 9, 5054-5064, 2003.
- DOWELL, S. F. et. al. Respiratory syncytial virus in an important cause of community-acquired lower respiratory infection among hospitalized adults. **J Infect Dis**, v. 62, p. 174-456, 1996.
- DUDAS, A. R.; KARRON, A. R. Respiratory Syncytial Virus Vaccines. **Clinical Microbiology Reviews**, v. 11, p. 430-439, 1998.
- DUTCH, R. E. Entry and Fusion of Emerging Paramyxoviruses. **PLoS Pathog**, [S.l.], v. 6, n. 6, 2010.

- EMSLEY, P. et al. Features and development of Coot. **Acta Cryst.**, [S.l.], v. 66, p. 486-501, 2010.
- EMSLEY, P.; COWTAN, K. **Coot**: model-building tools for molecular graphics. **Acta Cryst.**, [S.l.], v. 60, p. 2126-2132, 2004.
- EVANS, J. E.; CANE, P. A.; PRINGLE, C. R. Expression and characterisation of the NS1 and NS2 proteins of respiratory syncytial virus. **Virus Reseachr**, v. 43, p. 155-161, 1996.
- FAÇANHA, M. C., PINHEIRO A. C. Doenças respiratórias agudas em serviços de saúde entre 1996 e 2001. **Rev. Saúde Pública.**, Fortaleza, v. 38, n.3, p. 346-350, 2004.
- FALSEY, A. R. et. al. Respiratory syncytial virus infection in elderly and high-risk adults. **N. Engl. J. Med.**, v. 352, p. 1749-1759, 2005.
- FALSEY, A. R.; WALSH, E. E. Respiratory Syncytial Virus Infection in Adults. **Clin. Microbiol. Rev.**, [S.l.], v. 13, n. 3, p. 371-384, 2000.
- FELDMAN, S. A.; AUDET, S.; BEELER, J. A. The Fusion Glycoprotein of Human Respiratory Syncytial Virus Facilitates Virus Attachment and Infectivity via an Interaction with Cellular Heparan Sulfate. **J. Virol**, [S.l.], v. 74, n. 14, p. 6442-6447, 2000.
- FIORUCCI, S. et al. Molecular simulations enlighten the binding mode of quercetin to lipoxygenase-3. **Proteins: Structure, Function and Bioinformatics**, [S.l.], v. 73, n. 2, p. 290-298, 2008.
- FOSSEY, M. A. et al. Study Of The Sulfated Polysaccharides And Polypeptides Interaction. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE BIOQUÍMICA E BIOLOGIA MOLECULAR, 23., 2003, Caxambu. Caxambu: [s.n.], 2003.
- GADJA, M. J.; PAWLOWSKI, M.; BUJNICKI, J. M. **Protein Structure Prediction**: From Recognition of Matches with Known Structures to Recombination of Fragments. **Multiscale Approaches to Protein Modeling**, [S.l.], p. 231-254, 2011.
- GONZALEZ-REYES, L. et al. Cleavage of the human respiratory syncytial virus fusion protein at two distinct sites is required for activation of membrane fusion. **PNAS**, [S.l.], v. 98, n. 17, p. 9859-9864, 2001.
- GREEN, M.; BRAYER, A. F.; SCHENKMAN, K. A.; WALD, E. R. Duration of hospitalization in previously well infants with respiratory syncytial virus infection. **The Pediatric Infectious Disease Journal**, [S.l.], v. 8, p. 601-605, 1989.
- GROOTHUIS, J. R.; HOOPES, J. M.; HEMMING, V. G. **Prevention of Serious Respiratory Syncytial Virus-Related Illness**: Immunoprophylaxis. **Advances in Therapy**, [S.l.], v. 28, p. 110-125, 2011.
- HACKING, D.; HULL, J. Respiratory Syncytial Virus – Viral Biology and the Host Response. **Journal of infection**, v. 45, p. 18-24, 2002.
- HALL, T. **BioEdit**: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium**, [S.l.], v. 41, n. 41, p. 95-98, 1999.

HENRICH, S. et al. Computational approaches to identifying and characterizing protein binding sites for ligand design. **Journal of Molecular Recognition**, [S.l.], V. 23, n. 2, p. 209-219, 2010.

HU, J.; ROBINSON, J. L. **Treatment of respiratory syncytial virus with palivizumab: a systematic review.** **World Journal of Pediatrics**, [S.l.], v. 6, p. 296-300, 2010.

JOHNSON, P. R.; COLLINS, P. L. **The fusion glycoproteins of human respiratory syncytial virus of subgroups A and B: sequence conservation provides a structural basis for antigenic relatedness.** **The Journal of General Virology**, [S.l.], v. 69, p. 2623-2628, 1988.

KARRON, R. A. et al. **Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication in vitro: Clinical evaluation and molecular characterization of a cold-passaged, attenuated RSV subgroup B mutant.** **PNAS**, [S.l.], v. 94, n. 25, p. 13961-13966, 1997.

KAUFMANN, K. W. et al. **Practically Useful: What the Rosetta Protein Modeling Suite Can Do for You.** **Biochemistry**, [S.l.], v. 49, n. 10, p. 2987-2998, 2010.

KLENK, H. D.; GARTEN, W. Host cell proteases controlling virus pathogenicity. **Trends in Microbiology**, [S.l.], v. 2, p. 39-43, 1994.

KOES, R. E.; QUARTTROCCHIO, F. **The flavonoid biosynthetic pathway in plants: function and evolution.** **Bioessays**, v.16, p.123-132, 1994.

KOSAKA, I. M.; CARROMEU, C.; DURIGON, E. L. et al. Construction of adenoviral vectors expressing F and G glycoproteins of human respiratory syncytial virus (HRSV). **Brazilian Journal of Microbiology**, [S.l.], v. 35, n. 1-2, p. 167-172, 2004.

LARKIN, M. A. et al. Clustal W and Clustal X version 2.0. **Bioinformatics**, [S.l.], v. 23, n. 21, 2007.

LASKOWSKI, R. A. et al. **PDBsum: A Web-based database of summaries and analyses of all PDB structures.** **Biochemical Sciences**, v. 22, n. 12, p. 488-490, 1997.

LASKOWSKI, R. A. PDBsum new things. **Nucl. Acids. Res.**, [S.l.], v. 37, p. 355-359, 2009.

LASKOWSKI, R. A. et al. **PROCHECK: a program to check the stereochemical quality of protein structures.** **J. Appl. Cryst.**, [S.l.], v. 26, p. 283-291, 1993.

LAURIE, A. T. R.; JACKSON, R. M. **Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites.** **Bioinformatics**, [S.l.], v. 21, n. 9, p. 1908-1916, 2005.

LEACH, A. R.; GILLET, V. J. **An Introduction to Chemoinformatics.** [S.l.]:Springer, 2007, 255 p.

LE GUILOX, G.; SCHMIDTKE, P.; TUFFERY, P. **Fpocket: An open source platform for ligand pocket detection.** **BMC Bioinformatics**, [S.l.], v. 10, 2009.

LOPES, R. M. et al. Flavonóides. **Biotechnologia: Ciência & Desenvolvimento**, [S.l.], p. 18-22, 2003.

LOURENÇÃO, L. G. et al. Infecções pelo Vírus Sincicial Respiratório em crianças. **Pulmão**, Rio de Janeiro, v. 12, p. 59-68, 2005.

LUSCOMBE, N. M.; GREENBAUM, D.; GERSTEIN, M. What is Bioinformatics? A Proposed Definition and Overview of the Field. **Method Inform Med**, [S.l.], v. 4, p. 346-357, 2001.

MANOHA, C. et al. Rapid and sensitive detection of metapneumovirus in clinical specimens by indirect fluorescence assay using a monoclonal antibody. **Journal of Medical Virology**, [S.l.], v. 80, p. 154-158, 2008.

MATHEWS, J.M. *et al.* The Core of the Respiratory Syncytial Virus Fusion Protein Is a Trimeric Coiled Coil. **Journal of Virology**, [S.l.], v.74, n. 13, p. 5911-5920, 2000.

MCDONALD, T. P. et al. **Evidence that maturation of the N-linked glycans of the respiratory syncytial virus (RSV) glycoproteins is required for virus-mediated cell fusion:** The effect of α -mannosidase inhibitors on RSV infectivity. **Virology**, [S.l.], v. 350, n. 2, p. 289-301, 2006.

MCGUFFIN, L. J.; BRYSON, K.; JONES, D. T. The PSIPRED protein structure prediction server. **Bioinformatics**, [S.l.], v. 16, n. 4, 2000.

MELERO, J. A. et al. The Soluble Form of Human Respiratory Syncytial Virus Attachment Protein Differs from the Membrane-Bound Form in Its Oligomeric State but Is Still Capable of Binding to Cell Surface Proteoglycans. **Journal of Virology**, [S.l.], v. 78, p. 3524-3532, 2004.

MELLO, W.; SILVA, C. Epidemiological aspects of RSV subgroups in Belém, Brazil. *Boletim Latino-americano* 1. 1992.

MITEVA, M. A.; GUYON, F.; TUFFERY, P. **Frog2:** Efficient 3D conformation ensemble generator for small compounds. **Nucl. Acids Res.**, [S.l.], v. 38, p. 622-627, 2010.

MORRIS, G. M. et al. **AutoDock4 and AutoDockTools4:** Automated docking with selective receptor flexibility. **Journal of Computational Chemistry**, [S.l.], v. 30, n. 16, p. 2785-2791, 2009.

MORTON, C. J. et al. **Structural characterization of respiratory syncytial virus fusion inhibitor escape mutants:** homology model of the F protein and a syncytium formation assay. **Virology**, [S.l.], v. 311, p. 275-288, 2003.

MOURA, F. E. A. et al. **Short Report: Respiratory Syncytial Virus Infections in Northeastern Brazil:** Seasonal Trends and General Aspects. **Am. J. Tro. Med. Hyg.**, [S.l.], v. 74, p. 165-167, 2006.

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. **The Lancet**, [S.l.], v. 375, p. 1545-1555, 2010.

NI, L. et al. Design and characterization of human respiratory syncytial virus entry inhibitors. **International Medical Press**, [S.l.], v. 10, p. 833-840, 2005.

NIJVELDT, R. J. et al. **Flavonoids**: a review of probable mechanisms of action and potential applications. **Am J Clin Nutr**, v. 74, p. 418-425, 2001.

OHUCHI, R. et al. Oligosaccharides in the stem region maintain the influenza virus hemagglutinin in the metastable form required for fusion activity. **J. Virol.**, [S.l.], v. 71, n. 5, p. 3719-3725, 1997.

OGRA, P. L. **Respiratory syncytial virus**: the virus, the disease and the immune response. **Paediatr Respir Rev**, Supplement A, p. S119-S126, 2004.

OLSZEWSKA, W.; OPENSHAW, P. Emerging drugs for respiratory syncytial virus infection. **Expert Opin. Emerg. Drugs**, [S.l.], v. 14, p. 207-217, 2009.

PAES, B. A.; BANERJI, A.; LANCTOT, K. L.; LANGLEY, J. M. **A decade of respiratory syncytial virus epidemiology and prophylaxis**: Translating evidence into everyday clinical practice. **Canadian Respiratory Journal**, [S.l.], v. 18, p. 10-19, 2011.

PARAMESWARI, A. R. et al. **Understanding the Binding Modes and Affinities of different Drug Molecules of Alzheimer Disease in Acetylcholinesterase Active Site Gorge**: A Comparative Molecular Docking Study. **International Journal of Pharmaceutical Research**, [S.l.], v. 3, n. 3, p. 52-58, 2011.

PARVEEN, S. et al. Genetic diversity among respiratory syncytial viruses that have caused repeated infections in children from rural India. **Journal of Medical Virology**, [S.l.], v. 78, p. 659-665, 2006.

POLLARD, T. D.; EARNSHAW, W. C. **Cell Biology**. Nova Iorque: W.B. Saunders, 2002.

RAMACHANDRAN, S. et al. Automated minimization of steric clashes in protein structures. **Proteins: Structure, Function and Bioinformatics**, [S.l.], v. 79, n. 1, p. 261-270, 2011.

RAMAN, R. et al. **Glycomics**: an integrated systems approach to structure-function relationships of glycans. **Nature Methods**, v. 2, p. 817-824, 2005.

RHOL, C. A. et al. Protein Structure Prediction Using Rosetta. **Numerical Computer Methods**, [S.l.], v. 383, p. 66-93, 2004.

ROST, B.; YACHDAV, G.; LIU, J. The PredictProtein server. **Nucl. Acids. Res.**, [S.l.], v. 32, p. 321-326, 2004.

RUBIO, M.; CALÇA, J.; BONFIM, C. M.; Gomes, E. D.; [DURIGON, E. L.](#) ; [FOSSEY, M. A.](#) ; RAHAL, P.; S. F. P. **HEPARIN AND DEXTRAN SULFATE: THEIR ROLE ON RSV INFECTIVITY**. **Virus Reviews and Research**, v. 15, p. X-z, 2010.

RUIZ-ARGUELLO, M. B. et al. Effect of Proteolytic Processing at Two Distinct Sites on Shape and Aggregation of an Anchorless Fusion Protein of Human Respiratory Syncytial Virus and Fate of the Intervening Segment. **Virology**, v. 298, p. 317-326, 2002.

SANDHAR, H. K. et al. A Review of Phytochemistry and Pharmacology of Flavonoids. **International Pharmaceutica Scientia**, [S.l.], v. 1, p. 25-41, 2011.

SANTOS, F. R.; ORTEGA, J. M. Bioinformática aplicada à Genômica. In: Biowork IV, 2003, Viçosa. Viçosa: [s.n.], 2003.

SANTOS FILHO, O. A.; ALENCASTRO, R. B. Modelagem de proteínas por homologia. **Quim. Nova**, [S.l.], v. 26, n.2, 2003.

SCHIMITT, A.P.; LAMB, R.A. **Escaping from the cell**: assembly and budding of negative-strand RNA viruses. **Current Topics Microbiology Immunology**, [S.l.], 283: 145-196, 2004.

SCHLENDER, J. et al. Inhibition of Toll-Like Receptor 7- and 9-Mediated Alpha/Beta Interferon Production in Human Plasmacytoid Dendritic Cells by Respiratory Syncytial Virus and Measles Virus. **Journal of Virology**, [S.l.], v. 79, n. 9, p. 5507-5515, 2005.

SCHNEIDER, G.; BOHM, H. J. Virtual screening and fast automated docking methods. **Drug Discovery Today**, [s.l.], v. 7, n. 1, p. 64-70, 2002.

SCHOIET, B. K. Virtual screening of chemical libraries. **Nature**, [S.l.], v. 432, p. 862-865, 2004.

SIEW, N. et al. **MaxSub**: an automated measure for the assessment of protein structure prediction quality. **Bioinformatics**, [S.l.], v. 16, n. 9, p. 776-785, 2000.

SILVA, L. H. A., et al. Vírus Sincicial Humano e Metapneumovírus Humano. **Rev HCPA**, v. 29, p. 139-146, 2009.

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. **J Clin Microbiol**, v. 29, p. 557-559, 1991.

SKROPETA, D. The effect of individual N-glycans on enzyme activity. **Bioorganic & Medicinal Chemistry**, [S.l.], v. 17, n. 7, p. 2645-2653, 2009.

STENSBALLE, L. G. et al. **Respiratory syncytial virus epidemics**: the ups and downs of a seasonal virus. **Pediatric Infectious Disease Journal**, [S.l.], v. 22, p. 21-32, 2003.

STRALIOTTO, et. al. **Respiratory syncytial virus (RSV) bronchiolitis**: comparative study of RSV groups A and B infected children. **Rev Soc Bras Med Trop**, v. 27, p. 1-4, 1994.

SULLENDER, W. M. Respiratory syncytial virus genetic and antigenic diversity. **Clin Microbiol Rev**, v. 13, p. 1-15, 2000.

TAPAS, A. R.; SAKARKAR, D. M.; KAKDE, R. B. **Flavonoids as Nutraceuticals**: A Review. **Tropical Journal of Pharmaceutical Research**, v. 7, n. 3, 2008.

TENG, M. N.; COLLINS, P. L. Altered Growth Characteristics of Recombinant Respiratory Syncytial Viruses Which Do Not Produce NS2 Protein. **Journal of Virology**, v. 73, p. 466-473, 1999.

TOOMULA, N.; KUMAR, D. S.; KUMAR A, V. V. L. Computational Methods for Protein Structure Prediction and Its Application in Drug Design. **Proteomics & Bioinformatics**, [S.l.], 2011.

- TRIFONOV, E. N. Earliest pages of bioinformatics. **Bioinformatics**, [S.l.], v. 16, n. 1, p. 5-9, 2000.
- 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**, [S.l.], v. 31, n. 2, p. 455-461, 2010.
- VIEIRA, S. A. et al. Clinical Patterns and Seasonal Trends in respiratory Syncytial Virus Hospitalizations in São Paulo, Brazil. **Rev. Inst. Med. Trop.**, São Paulo, v. 43, n. 3, p. 125-131, 2001.
- VOET, D.; VOET, J. G. **Biochemistry**. 2 ed. [S.l.]: John Wiley Sons, INC, 1995.
- WALSH, E. E. et al. Severity of respiratory syncytial virus infection is related to virus strain. **J Infect Dis**, v. 143, p. S112-117, 1997.
- WALLNER, B.; ELOFSSON, A. Can correct protein models be identified. **Protein Sci.**, [S.l.], v. 12, n. 5, p. 1073-1086, 2003.
- WALLNER, B.; ELOFSSON, A. Prediction of global and local model quality in CASP7 using Pcons and ProQ. **Proteins: Structure, Function, and Bioinformatics**, [S.l.], v. 69, p. 184-193, 2007.
- WANG, D. et al. **Immunoprophylaxis against respiratory syncytial virus (RSV) with palivizumab in children**: a systematic review and economic evaluation. [S.l.]:[s.n.], 2008. 107 p.
- WANG, G. et al. Binding of Quercetin to Lysozyme as Probed by Spectroscopic Analysis and Molecular Simulation. **J. Fluoresc.**, [S.l.], v. 21, p. 1879-1886, 2011.
- WEISMAN, L. E.; GROOTHIUS, J. R. **Contemporary diagnosis and management of respiratory syncytial virus**. [S.l.]: Handbooks in Health Care, 2000. 253 p.
- WELLIVER, R. C. Pharmacotherapy of respiratory syncytial virus infection. **Current Opinion in Pharmacology**, [S.l.], v. 10, p. 289-293, 2010.
- WHITE, J. M. Viral and Cellular Membrane Fusion Proteins. **Annual Review of Physiology**, [S.l.], v. 52, p. 675-697, 1990.
- WRIGHT, M.; PIEDIMONTE, G. **Respiratory Syncytial Virus Prevention and Therapy**: Past, Present, and Future. **Pediatric Pulmonology**, [S.l.], v. 46, p. 324-347, 2011.
- WU, S.; SKOLNICK, J.; ZHANG, Y. Ab initio modeling of small proteins by iterative TASSER simulations. **BMC Biology**, v. 5, n. 17, p. 1-10, 2007.
- WU, Q. et al. Oral Respiratory Syncytial Virus (RSV) DNA Vaccine Expressing RSV F Protein Delivered by Attenuated Salmonella typhimurium. **Human Gene Therapy**, [S.l.], v. 18, p. 746-752, 2007.
- ZHANG, Y. I-TASSER server for protein 3D structure prediction. **BMC Bioinformatics**, [S.l.], v. 9, n. 40, 2008.

ZHANG, Y. **Protein structure prediction**: when is it useful. **Current Opinion in Structural Biology**, [S.l.], v. 19, n. 2, p. 145-155, 2009.

ZHAO, X. et al. Structural characterization of the human respiratory syncytial virus fusion protein core. **PNAS**, [S.l.], v. 97, p. 14172-14177, 2000.

ZHU, J. et al. Design and analysis of post-fusion 6-helix bundle of heptad repeat regions from Newcastle disease virus F protein. **Protein Engineering**, [S.l.], v. 16, p. 373-379, 2003.

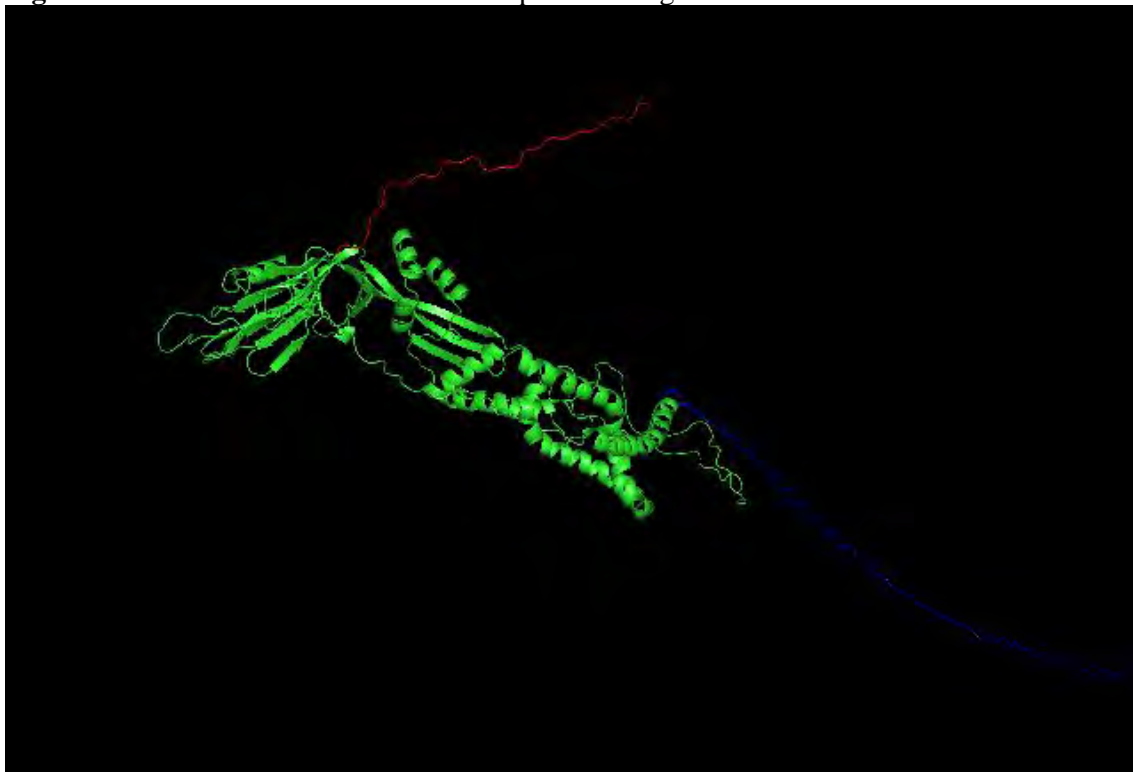
ZIMMER, G.; CONZELMANN, K.; HERRLER, G. Cleavage at the Furin Consensus Sequence RAR/KR and Presence of the Intervening Peptide of the Respiratory Syncytial Virus Fusion Protein Are Dispensable for Virus Replication in Cell Culture. **PNAS**, [S.l.], v. 98, n. 17, p. 9859-9864, 2001.

ZIMMER, G.; TROTZ, I.; HERRLER, G. N-Glycans of F Protein Differentially Affect Fusion Activity of Human Respiratory Syncytial Virus. **J. Virol.**, [S.l.], v. 75, n. 10, p. 4744-4751, 2001.

7. ANEXOS

ANEXO A – Modelo da Proteína de Fusão gerado por homologia pelo programa MODELLER.

Figura 17 – Modelo da Proteína de Fusão por homologia



Em vermelho e azul estão destacadas as regiões N-terminal e C-terminal da proteína, respectivamente. Como pode se observar, estas regiões não foram modeladas, o que invalidou o modelo. Fonte: CRAVO, H. L. P., 2012.

ANEXO B – Modelo da Proteína de Fusão gerado por *Ab Initio* pelo programa ROSETTA.

Figura 18 – Modelo da Proteína de Fusão por *Ab Initio*



O modelo gerado não corresponde com as características da proteína relatadas na literatura, sendo assim descartado. Fonte: CRAVO, H. L. P., 2012.