



**UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”**

Campus de São José do Rio Preto

Dissertação de Mestrado

Mariana Pela Sabbag

**Caracterização estrutural e das interações entre a Proteína G do hRSV
e potenciais Inibidores**

São José do Rio Preto

2012

Mariana Pela Sabbag

**Caracterização estrutural e das interações entre a Proteína G do hRSV
e potenciais Inibidores**

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

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

São José do Rio Preto

2012

Mariana Pela Sabbag

**Caracterização estrutural e das interações entre a Proteína G do hRSV
e potenciais Inibidores**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Microbiologia 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

Mariana Pela Sabbag

**Caracterização estrutural e das interações entre a Proteína G do hRSV
e potenciais Inibidores**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Microbiologia, junto ao Programa de Pós-Graduação em Microbiologia, Área de Concentração - Virologia, 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

Profa. Dra. Fátima Pereira de Souza

Departamento de Física - UNESP- São José do Rio Preto

Profa. Dra. Karina Alves de Toledo

Departamento de Ciências Biológicas - UNESP - Assis

Profa. Dra. Teresa Cristina Cardoso

Depto. De Apoio, Produção e Saúde Animal UNESP- Araçatuba

São José do Rio Preto, 16 de janeiro de 2012.

Aos que torceram, acreditaram e estiveram presentes em mais essa etapa de minha vida.

AGRADECIMENTOS

A **Deus** por se fazer presente em todos os momentos de minha vida;

Aos meus pais, que me ensinaram os primeiros passos e me encorajaram nos primeiros vãos. Obrigada por tudo!

Às melhores irmãs do mundo, por sempre me apoiarem e torcerem por mim;

Ao meu namorado, que sempre me encorajou e me deu colo quando precisei. Com você minha caminhada sempre será segura, onde quer que eu esteja;

A Prof^a. Dr^a. Fátima P. Souza, pela amizade, confiança e comprometimento com este trabalho, me orientando sem poupar esforços, e aos participantes da banca de defesa, que contribuíram muito para a correção e aprimoramento deste trabalho;

Aos queridos amigos, Haroldo e Deriane, que me cederam amizade, tempo e conhecimento, e sem os quais esse trabalho jamais seria concluído;

Aos amigos de graduação, que nunca serão esquecidos, pelo tempo bom que passamos juntos, sentirei saudades eternas de vocês;

Aos amigos: Gabi, Luciana, Ícaro, Danúbia, Priscila e Gui, vocês ficarão para sempre em meus pensamentos;

A todos os amigos e funcionários do Departamento de Física, pelo apoio técnico-administrativo, amizade, descontração e incentivo;

À CAPES, FAPESP, CNPQ e Prope, pela bolsa de mestrado concedida e pelo auxílio financeiro, imprescindível para o desenvolvimento desta pesquisa. Este projeto recebeu auxílio financeiro das agências FAPESP processo:2010/50444-4/ CNPq processo: 306342/2010-3.

Ao Programa de Pós-graduação em Microbiologia, e à Pró Reitoria de Pesquisa e à todos seus funcionários, por toda a ajuda prestada ao longo deste trabalho.

Meus sinceros agradecimentos!

“A vida é um milhão de novos começos movidos pelo desafio sempre novo de viver e fazer todo sonho brilhar. Então, sonhe, imagine, viva os momentos de alegria! Porque deles vão sobrar lembranças, e elas tem que ser as melhores!”

RESUMO

As infecções respiratórias agudas (IRAs) constituem a principal causa de mortalidade infantil no mundo, e o Vírus Respiratório Sincicial Humano (hRSV – *Human Respiratory Syncytial Virus*) é um dos principais agentes etiológicos das IRAs. Este vírus pertencente à família *Paramyxoviridae*, é envelopado, de simetria helicoidal, cujo genoma é RNA de fita simples não segmentada. A infectividade do vírus está relacionada com suas proteínas de membrana e dentre elas a glicoproteína G, que é responsável pela ligação do vírus à célula hospedeira e conseqüente instalação da infecção. Esta glicoproteína exerce um importante papel como antígeno de reconhecimento, sendo alvo para identificação do RSV através de anticorpos. Existem evidências de que esta proteína se liga a receptores glicosilados na célula hospedeira, porém ainda não foi descrito um receptor para a proteína G na célula. Para elucidar estes mecanismos de interação, foram realizados estudos experimentais e teóricos desta proteína. Os domínios solúveis da região N-terminal (1 a 38 aa) e C-terminal (67 a 298 aa), com 231 aminoácidos da glicoproteína G do hRSV foram clonados e a região N-terminal foi expressa em bactéria BL21 pLysS. Em paralelo, foi realizada a caracterização teórica desta proteína, e foram avaliados os possíveis sítios de interação da mesma com glicosaminoglicanos (heparina). Foram obtidos dois modelos teóricos para a proteína G do hRSV, bem como dois modelos de interação com heparina, determinando portanto, um possível sítio de ocorrência de interação. O conhecimento da estrutura da proteína G é de grande importância para elucidar a composição da estrutura e os mecanismos de interação com potenciais ligantes e deste modo, em um passo posterior, propor mecanismos de reconhecimento celular pelo hRSV, através de glicosaminoglicanos.

Palavras chaves: Proteína G, hRSV, Modelagem, *Docking* Molecular.

ABSTRACT

Acute Respiratory Infections (ARI) are the leading cause of infant mortality in the world, and the Human Respiratory Syncytial Virus (hRSV) is one of the main agents of ARI. This virus belongs to Paramyxoviridae family, has a lipidic envelope, helical symmetry and its genome is a single-stranded RNA. The viral infectivity is related to its membrane proteins and among them the G glycoprotein, which is responsible for binding the virus to the host cell and consequent infection. This glycoprotein plays an important role as antigen recognition, being the target for hRSV identification through antibodies. There are evidences that this protein binds to host cell glycosylated receptors, but it has not been described a receptor for G protein in the cell yet. To elucidate these interaction mechanisms and understand the process of viral infectivity, we performed experimental and theoretical studies of this protein. The soluble domains of the N-terminal (1-38 aa) and C-terminal regions (67-298 aa), with 231 amino acids of the hRSV G glycoprotein have been cloned and the N-terminal region was expressed in BL21 pLysS bacteria. In a later trial these peptides will be purified and biophysical tests will be done. It was also performed a theoretical characterization of this protein, to assess the possible interaction sites with glycosaminoglycan (heparin). It were obtained two theoretical models for the hRSV G protein as well as two interaction models with heparin, in order to determine a possible site of occurrence of interaction. Knowledge of G protein structure is of great importance to elucidate the mechanism of viral infectivity and interaction mechanisms with potential ligands, and the results obtained in this work will allow us, in a later step, to propose mechanisms of cellular recognition by hRSV through glycosaminoglycans.

Key words: *G Protein, hRSV, Molecular Docking*

LISTA DE ILUSTRAÇÕES

Figura 1 – Classificação dos Membros da Família <i>Paramyxoviridae</i>	5
Figura 2 – Estrutura do Vírus Sincicial Respiratório	6
Figura 3 – Esquematização do genoma do hRSV, mostrando sua ordem gênica	6
Figura 4 – Representação esquemática da glicoproteína G do hRSV.....	9
Figura 5 – Mapa de restrição do plasmídeo pCR-XL-TOPO®.....	15
Figura 6 – Mapa de restrição do plasmídeo Pet28A (NOVAGEN)	17
Figura 7 – Esquemas das áreas geométricas do programa PatchDock.....	26
Figura 8 – Amplificação dos domínios N-terminal e C-terminal da proteína G	27
Figura 9 – Triagem dos transformantes pCR-XL-TOPO® - G1. e G2	28
Figura 10 – Resultado do seqüenciamento do domínio G1.....	28
Figura 11 – Perfil de restrição enzimática dos clones pCR-XL-TOPO® - G1	29
Figura 12 – Triagem dos transformantes contendo domínios do gene G1 do hRSV	29
Figura 13 – Miniprep do vetor pet28A contendo o inserto G1	30
Figura 14 – SDS Page 15%, mostrando a peptídeo G1 de ~ 3,0 KDa	31
Figura 15 – SDS Page 17% mostrando peptídeo G1 na forma solúvel.....	32
Figura 16 – SDS Page 17% mostrando purificação do peptídeo G1 com resina	33
Figura 17 – Modelos da proteína G do hRSV gerados pelo ROSETTA e I-TASSER ..	34
Figura 18 – Modelo do endodomínio (G1) do hRSV, gerado pelo ROSETTA.....	35
Figura 19 – Gráficos de Ramachandran para modelos da proteína G.....	36
Figura 20 – Localização dos resíduos de cisteínas, da proteína G do hRSV	37
Figura 21 – Esquema das interações da proteína G do hRSV	41
Figura 22 – Modelo da região de interação da proteína G do hRSV	42
Figura 23 – Resultados obtidos com o Q-Site Finder	42
Figura 24 – Esquema da interação da proteína G com heparina - PatchDock	43
Figura 25 - Esquema da interação da proteína G com heparina - AutoDock.....	44

LISTA DE TABELAS

Tabela 1 – Oligonucleotídeos iniciadores do gene para proteína G do hRSV.	14
Tabela 2 – Características avaliadas por programas de bioinformática.....	21
Tabela 3 – Estrutura secundária da proteína G obtida pelo programa Predictprotein....	33
Tabela 4 – Sítios de modificações pós traducionais da Proteína G.....	39
Tabela 5 – Ponto isoelétrico da Proteína G e seus domínios.....	40

SUMÁRIO

1. INTRODUÇÃO	1
1.1. Vírus Sincicial Respiratório Humano	1
1.1.1. Histórico	1
1.1.2. Características Estruturais e Epidemiologia	2
1.1.3. Organização Genômica e Proteínas Virais	3
1.1.4. Profilaxia e Tratamento	7
1.2. Glicosaminoglicanos Sulfatados - Heparina	8
1.3. Ferramentas de Bioinformática	9
2. OBJETIVOS	10
3. MATERIAL E MÉTODOS	11
3.1. Análise Experimental	11
3.1.1. Extração RNA e Transcrição Reversa	11
3.1.2. Desenho dos oligonucleotídeos iniciadores	12
3.1.3. Amplificação do Endodomínio, Ectodomínio da Proteína G	12
3.1.4. Clonagem - TOPO [®] XL PCR Cloning Kit	13
3.1.5. Sequenciamento dos produtos clonados	14
3.1.6. Clonagem em vetor pet28A	15
3.1.7. Indução da expressão da proteína G1	16
3.1.8. Purificação Protéica	17
3.2. Modelagem e Análise Teórica	18
3.2.1. Seleção das seqüências da Proteína G do Banco de Dados Genbank	18
3.2.2. Análise das seqüências pelos programas de Bioinformática	19
3.2.3. Estrutura da Heparina	19
3.2.4. Modelagem da Proteína G pelo programa ROSETTA	20
3.2.5. Modelagem da Proteína G pelo programa I-TASSER	22
3.2.6. Validação dos Modelos	22
3.2.7. <i>Docking</i> Molecular da Interação Proteína G com a Heparina	23
4. RESULTADOS E DISCUSSÃO	25
4.1. Amplificação do Endodomínio, Ectodomínio da Proteína G	25
4.2. Clonagem da Proteína G e seus domínios em pCR-XL-TOPO	25
4.3. Clonagem da Proteína G e seus domínios em vetor pet28A	26

4.4. Análise da expressão da proteína G1 em bactéria	28
4.5. Purificação do Endodomínio da Proteína G	30
4.6. Predição de Estrutura Secundária	31
4.7. Validação dos Modelos	33
4.8. Pontes Dissulfeto	34
4.9. Modificações Pós Traducionais.....	36
4.10. Ponto Isoelétrico.....	38
4.11. Verificação dos Possíveis Sítios de Interação	39
4.12. Verificação das possíveis interações G-heparina.	40
5. CONCLUSÕES.....	44
6. REFERÊNCIAS BIBLIOGRÁFICAS	45

1. Introdução

As infecções respiratórias agudas (IRAs) constituem a principal causa de mortalidade infantil no mundo (THOMAZELLI et al., 2007), estimando-se que ocorram anualmente 2 milhões de mortes entre crianças menores que 5 anos de idade (FAÇANHA & PINHEIRO, 2004).

No Brasil, as doenças respiratórias são responsáveis por mais de 50% das consultas médicas em centros de saúde (BOSSO et al., 2004). Na última década, a mortalidade por infecções respiratórias agudas vem diminuindo, mas mesmo assim, são responsáveis pela morte anual de 4 milhões de pessoas (WHO, 2008).

Os vírus são responsáveis por 50 a 90% das infecções do trato respiratório inferior em crianças pequenas (THOMAZELLI et al., 2007), sendo a maioria delas atribuída aos seguintes vírus: Vírus sincicial respiratório humano (VSRH), Vírus da Parainfluenza Humana (VPIH) I, II, e III, Influenza A e B, Adenovírus, e mais recentemente, o Metapneumovírus Humano (MPVH) (MANOHA et al., 2007; PITREZ et al., 2005).

1.1. Vírus Sincicial Respiratório Humano

1.1.1. Histórico

O Vírus Respiratório Sincicial Humano (hRSV – *Human Respiratory Syncytial Virus*) é reconhecido como o agente viral mais frequentemente relacionado a casos de bronquiolite e pneumonia durante a infância e idade pré-escolar. O pico de incidência da infecção por este vírus ocorre nos primeiros meses de vida (ANDERSON et al., 1990).

Em 1955, no *Walter Reed Army Institute of Research*, Washington DC, Estados Unidos, Morris e colaboradores isolaram, pela primeira vez, o hRSV a partir de *swab* de garganta de chimpanzés que apresentavam sinais clínicos comuns às doenças respiratórias: coriza, secreção nasal purulenta, tosse e espirro. O isolamento foi realizado em cultura de células hepáticas e inoculado em animais de laboratório: camundongos, ratos, cobaias, hamsters, coelhos e chimpanzés; porém somente os primatas desenvolveram os sinais da doença (CHANOCK ET al., 1957). O agente isolado passou a chamar-se *Chimpanzee Coryza Agent* (CCA) (MORRIS et al., 1956).

Em 1957, um vírus semelhante foi isolado em Baltimore em duas crianças, uma apresentando pneumonia e outra laringotraqueobronquite. O nome CCA foi, portanto, substituído por *Human Respiratory Syncytial Vírus* (hRSV), para designar sua origem

humana, sua afinidade pelo trato respiratório e sua capacidade de produção de sincícios em cultura celular (CHANOCK et al., 1957a; CHANOCK et al., 1957b).

No Brasil, o vírus foi isolado pela primeira vez por Candeias em 1964, em quatro crianças hospitalizadas com quadro respiratório agudo. A confirmação do isolamento em cultura celular foi feita através da prova de neutralização com soro padrão (CANDEIAS, 1967).

1.1.2. Características estruturais e epidemiologia

O RSV é o principal agente causador de infecções respiratórias agudas (IRA), em crianças no mundo todo (SWEDAN et al., 2009; NAKAMURA et al., 2009; OLSZEWSKA & OPENSHAW, 2009; MORRIS et al., 2009).

O vírus tem distribuição mundial (COLLINS et al., 2001), e vem sendo apontado também como uma das causas de doença respiratória na população adulta, principalmente em idosos, imunocomprometidos e pessoas com doenças cardiorrespiratórias (COLLINS et al., 2001; FALSEY & WASH, 2000; PAVLOVA et al., 2009). Reporta-se que 0,5% a 2% de todas as crianças infectadas por RSV são hospitalizadas. Destas, de 7% a 21% desenvolvem insuficiência respiratória, conduzindo à morte aproximadamente 0,5% a 1,5% dos casos (RICCETTO et al., 2006).

A transmissão do RSV ocorre através do contato direto com secreções de pessoas infectadas ou objetos contaminados. Em ambientes fechados, creches e escolas, a propagação do vírus é favorecida. Crianças lactentes infectadas com o hRSV excretam grandes quantidades de partículas virais, facilitando a contaminação de indivíduos próximos (VIEIRA et al., 2002).

O RSV pode ocasionar infecções tanto no trato respiratório superior (ITRS) quanto no inferior (ITRI) (WU et al., 2007). As ITRI geralmente são mais graves, sendo a bronquiolite e pneumonia as doenças mais comumente associadas, principalmente em crianças e pacientes imunocomprometidos (WU et al., 2007). As alterações clínicas das ITRI é caracterizada por tosse, dificuldade respiratória, coriza, dispnéia e febre (KIM et al., 2007). No início da infecção os vírus podem atingir os alvéolos e infectar as células epiteliais das vias aéreas, induzindo a secreção de fatores como, por exemplo, os IFN α e β e NF-kB. No período tardio da infecção, a infecção viral é acompanhada pela infiltração de células inflamatórias nos pulmões, induzidas por citocinas. Neste ponto, células mononucleadas de sangue periférico são observadas em torno dos alvéolos (BUENO et al., 2008).

Em São José do Rio Preto – SP, cerca de 31,8% das crianças que frequentam creche e tem sintomas de doença respiratória, apresentam infecções respiratórias virais, sendo o hRSV responsável por 20,7% das infecções (BONFIM et al., 2011).

1.1.3. Organização Genômica e Proteínas Virais

O vírus hRSV é membro da família *Paramyxoviridae*, subfamília *Pneumovirinae* e gênero *Pneumovirus* (MACLELLN et al., 2007). Dentro do gênero *Pneumovirus* existem, além do hRSV, o RSV bovino (bRSV) e o Vírus da Pneumonia do camundongo (*Murine Pneumonia Vírus*). Foram descritas também amostras de bRSV em caprinos e ovinos, não consideradas como espécies diferentes, mas, representantes de um subgrupo de amostras de ruminantes. (VAN REGENMORTEL et al., 2000). Os membros do gênero *Pneumovirus* podem ser visualizados na Figura 1.

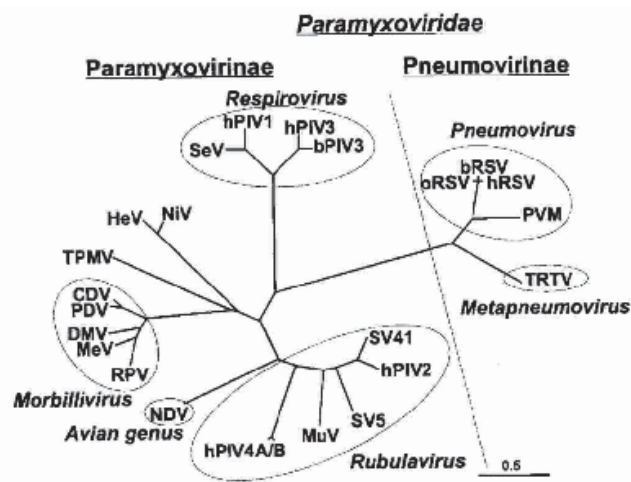


Figura 1. Classificação dos membros da família *Paramyxoviridae*. (COLLINS et al., 2001)

O RSV é um vírus envelopado, não-segmentado, com tamanho médio de 120 a 300nm e simetria helicoidal (LEMANSKE, 2004; BATONICK et al., 2008). Os membros da subfamília *Pneumovirinae* são vírus de tamanho variável (COLLINS et al., 2001). As partículas virais são compostas de um nucleocapsídeo e de um envelope lipoprotéico, derivado da membrana plasmática da célula hospedeira (COLLINS et al., 2001) (Figura 2). Sua superfície possui as seguintes glicoproteínas: Proteína de Ligação (G), Proteína de Fusão (F) e a Proteína Hidrofóbica Pequena (SH).

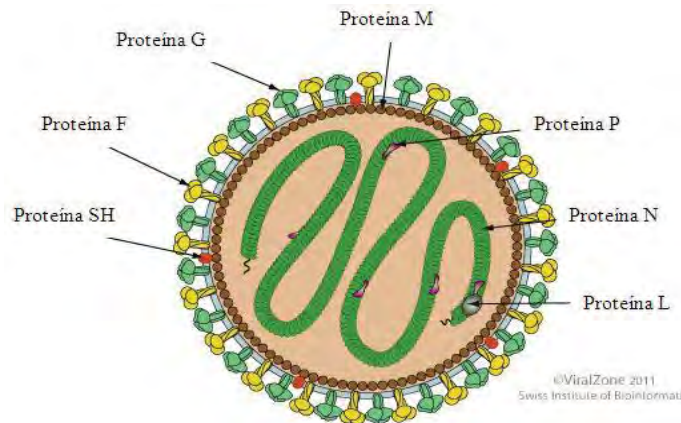


Figura 2. Representação esquemática do RSV. No esquema estão representadas as proteínas F, G, SH, P, L, N, M; e o envelope composto por bicamada lipídica. FONTE: Swiss Institute of Bioinformatics. ViralZone® [database]. Suíça: SIB; 2009. Disponível em: http://www.expasy.org/viralzone/all_by_species/90.html

O genoma dos *Pneumovirus* consiste em uma fita simples de RNA, com polaridade negativa, não segmentada, composta de aproximadamente 15.200 nucleotídeos e 10 genes, os quais codificam onze proteínas distintas (Figura 3). O peso molecular de seu genoma é de aproximadamente 5×10^3 kDa e contém cerca de 15.222 nucleotídeos (WELLIVER, 2003). Os nove primeiros genes são separados por regiões intergênicas que variam de 1 a 56 nucleotídeos. Os dois últimos genes fazem uma sobreposição (*overlap*) de 68 nucleotídeos, como também pode ser visto na figura 3 (COLLINS et al., 2001).

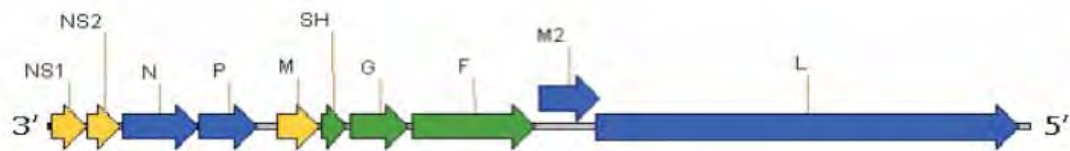


Figura 3. Esquema do genoma do hRSV, mostrando sua ordem gênica. Em azul estão os genes que codificam proteínas constituintes do nucleocapsídeo e em verde estão destacados os genes das glicoproteínas de superfície. O gene M2 possui duas ORFs: M2-1 e M2-2 (Adaptado de CARROMEU, 2007).

O RSV é dividido em dois grupos (A e B), baseado inicialmente em estudos com anticorpos monoclonais direcionados às proteínas F e G. Posteriormente estes estudos passaram a ser baseados no sequenciamento do gene G para classificação genotípica e, estudos epidemiológicos (BOTOSSO et al., 2009). Ensaio epidemiológicos demonstram que é possível a circulação de estirpes dos dois subgrupos dentro de uma mesma comunidade durante os surtos. Entretanto, suas proporções relativas podem diferir entre uma epidemia e outra (MORRIS et al., 2009; ZLATEVA et al., 2007, SATO et al., 2005).

O nucleocapsídeo presente nas partículas virais maduras do hRSV é formado pelo RNA genômico complexado com as proteínas N, P, L e M2-1 (COLLINS et al. 2001). A associação entre as proteínas N, P e L com o RNA viral já é suficiente para que o complexo apresente uma atividade de replicase. No entanto, a proteína M2-1 é necessária para uma eficiente transcrição do genoma pela polimerase viral (FEARNS E COLLINS, 1999).

A proteína N está associada ao RNA genômico e é a principal proteína estrutural do nucleocapsídeo (COLLINS et al., 2001). O RNA genômico do hRSV encontra-se sempre complexado com a proteína N e este complexo é usado como molde para a síntese de RNA. Este complexo proteína N-RNA previne a formação de estruturas secundárias no RNA, protegendo-o de degradação por nucleases e reduzindo a formação de RNAs dupla fita, podendo levar a resposta celular antiviral (COWTON et al., 2006).

As proteínas não estruturais NS1 e NS2 interferem na indução e sinalização do interferon (COLLINS & CROWE JR, 2007), o qual representa o principal mecanismo da imunidade inata no combate a infecções virais. O RSV, através das proteínas NS1 e NS2, pode agir inibindo ou neutralizando várias etapas da via dos interferons, o que favorece o maior sucesso da progênie viral (SWEDAN et al., 2009; LO et al., 2005).

A proteína P é altamente fosforilada, e desempenha função importante na transcrição e replicação do hRSV (COLLINS et al., 2001). A proteína P constitui uma subunidade da polimerase viral (formada pelas proteínas P e L) e parece mediar o contato entre a proteína L e o complexo N-RNA (TRAN et al., 2007).

A proteína L, a maior proteína do hRSV, contém seis segmentos altamente conservados e a proteína estrutural de menor abundância na partícula viral (COLLINS et al., 2001). Atribui-se a essa proteína as atividades de RNA-polimerase-RNA-dependente, poliadenilação e formação de estrutura de *cap* do RNAm viral (MASON et al., 2004).

A proteína M2-1 é interna e não glicosilada e interage com as proteínas N e P, atuando no processo de transcrição, evitando seu término prematuro (COLLINS et al., 2001). Essa proteína tem função na replicação do hRSV e alterações em sua sequência podem diminuir a infectividade do vírus (TANG et al., 2001).

A proteína M2-2 está relacionada à regulação da transcrição e replicação do RNA viral (ATREYA et al., 1998). A M2-2, quando expressa em altos níveis, inibe completamente a replicação do hRSV, sugerindo que essa proteína, quando produzida nas células infectadas em grande quantidade, pode ser um fator crítico na replicação do vírus (CHENG et al., 2005).

A proteína M está associada ao nucleocapsídeo do hRSV nas células infectadas e no vírion e, como as proteínas de matriz de outros vírus de RNA fita negativa, pode inibir a transcrição viral (GHILDYAL et al., 2002).

A glicoproteína de superfície SH é uma molécula pequena, transmembrana, cuja função ainda não está completamente elucidada, mas possivelmente sua atuação esteja relacionada ao Complexo de Golgi da célula hospedeira, no qual modificaria suas propriedades, para facilitar o transporte das proteínas virais pela via de secreção (RIXON et al., 2004).

A proteína F é uma glicoproteína responsável pela fusão do envelope viral com a membrana citoplasmática da célula hospedeira, fazendo com que o nucleocapsídeo seja liberado no interior do citoplasma (WALSH e HRUSKA, 1983). Essa proteína também proporciona a fusão de células infectadas formando sincícios (COLLINS et al., 2001).

A proteína G do hRSV é uma proteína de superfície, tem 289 a 319 aminoácidos, dependendo da amostra viral, (COLINS et al., 2002) e foi identificada como a responsável pela adesão do vírus à célula hospedeira, pela observação de que anticorpos específicos contra a mesma inibem esse processo (COLLINS et al., 2001, LEVINE et al., 1987; BOTOSSO et al., 2009). A proteína G tem sido alvo de estudos sobre a variabilidade genética do hRSV. Estudos indicam que ela é pouco conservada entre os grupos A e B, mostrando até 50% de divergência na seqüência de aminoácidos (ARBIZA et al., 2006). Alguns estudos mostraram que vários genótipos diferentes co-circulam, e alguns predominam em uma comunidade a cada ano (OSTLUND et al., 2008; PERET et al., 1998).

A glicoproteína G é uma proteína transmembrana integral do Tipo II (ou seja, está ancorada à membrana próximo à sua porção amino terminal por um domínio hidrofóbico que se estende do resíduo 38 ao 66), altamente glicosilada, com uma região localizada no citoplasma (N-terminal), uma região transmembrana e um grande ectodomínio (C-terminal) (GARCIA, O. et al., 1994). Esta proteína é sintetizada como uma precursor de peso molecular de 36kDa, sendo posteriormente glicosilada originando um complexo intermediário de 45kDa. Estes açúcares são convertidos para tipos complexos produzindo uma forma madura de 90 KDa.

Sua variabilidade e glicosilação estão localizadas principalmente no ectodomínio, que contém dois segmentos hipervariáveis com alto teor de serina e treonina, entre os aminoácidos 69-164 e 207-298, com aproximadamente 56% de divergência entre as estirpes A e B (JONHSON et al., 1987), separados por uma região altamente conservada entre os aminoácidos 164 e 199, e quatro resíduos de cisteínas (173, 176, 182, 186). Esta seqüência

motivo de 13 aminoácidos é considerada como provável sítio de ligação ao receptor da célula hospedeira (JOHNSON et al., 1987; CANE et al., 1991; SULLENDER et al., 2000). A proteína G também interage com o receptor celular CX3CR1, modulando assim a resposta imunológica do hospedeiro e facilitando a infecção (COLLINS et al., 2008). A representação esquemática da proteína G pode ser visualizada na figura 4.

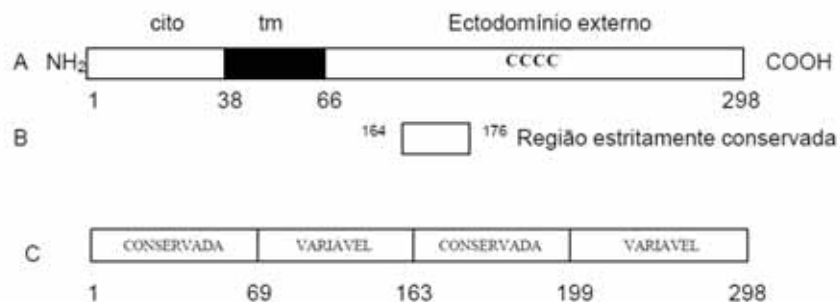


Figura 4. Representação esquemática da glicoproteína G DO hRSV. (A) Estrutura primária de aminoácidos sendo: Tm = domínio transmembrana; cito = domínio interno à membrana citoplasmática; CCCC = quatro resíduos de cisteínas (B) Região conservada nas amostras de hRSV; (C) Representação das duas regiões variáveis da proteína G, ricas em serina e treonina, compreendendo os aminoácidos 69-163 e 199-298 e região conservada compreendendo os aminoácidos 164-199. As regiões transmembrana e citoplasmática, compreendendo os aminoácidos 1-69, são conservadas. Extraído de SULLENDER, 2000.

1.1.4. Profilaxia e Tratamento

Muitos estudos têm sido realizados visando o desenvolvimento de um tratamento específico ou uma vacina eficiente contra o hRSV, porém a variabilidade antigênica viral e as características do sistema imune do hospedeiro dificultam a produção destas vacinas (PARVEEN et al., 2006).

Dentre as medidas de prevenção está a administração intravenosa de anticorpos monoclonais humanizados, como o Palivizumab, que age na atividade de fusão da proteína F. Entretanto, devido ao alto custo destas intervenções, as medidas têm sido aplicadas apenas em crianças com grande risco de desenvolver infecções graves (MEISSNER et al., 2003).

A única droga antiviral aprovada para tratamento do hRSV é a Ribavirina aerolizada, um análogo de nucleosídeo, responsável por inibir a síntese de ácido nucléico viral. A Ribavirina tem grande importância no tratamento de adultos imunocomprometidos, sobretudo transplantados de medula óssea, no entanto é responsável pela ocorrência de alguns efeitos colaterais indesejáveis. (NOKES e CANE, 2008).

1.2. Glicosaminoglicanos Sulfatados - Heparina

Os glicosaminoglicanos (GAGs) são polissacarídeos longos formados por unidades repetidas de dissacarídeos. Com exceção da hialurona, todos os GAGs são sulfatados e sintetizados como modificações pós-traducionais das proteínas chamadas proteoglicanas (POLLARD & EARNSHAW, 2007).

A principal função atribuída aos glicosaminoglicanos é relacionada à sua capacidade de interação com outras moléculas. Várias macromoléculas ligam-se aos GAGs, sendo a maioria proteínas ou proteínas conjugadas. A heterogeneidade dos GAGs em termos de padrão de sulfatação e estrutura química facilita a ligação de uma variedade de proteínas (RAMAN et al., 2005).

Estruturalmente, as heparinas são GAGs de peso molecular variável, compostos de unidades de Ácido Idurônico e Glucosamina, que se alternam entre ligações glicosídicas. 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 interação com ligantes carregados, como proteínas e íons bivalentes (VOET, 1995).

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). 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 apresentam maior afinidade de ligação.

Estudos identificaram GAGs na superfície celular, em particular a Heparina Sulfatada, como um dos fatores de maior importância para infecção do hRSV em células de cultura. Em particular, o nosso grupo experimental também verificou que a Heparina, em concentração acima de 3 μ M, inibe a infecção viral (RUBIO et al., 2010). 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 viral, porém, ainda não foi esclarecido qual seu papel em casos clínicos isolados.

1.3. Ferramentas de Bioinformática

O termo bioinformática refere-se, inicialmente, a análise de seqüências de proteínas e DNA (TRIFONOV, 2000). 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.

Inicialmente, a bioinformática foi definida como um campo interdisciplinar que envolvia a biologia, a ciência da computação, a matemática e a estatística para analisar dados de seqüências biológicas, conteúdo de genomas e para a predição de função e estruturas de macromoléculas (MOUNT, 2004). A bioinformática tem crescido exponencialmente devido, principalmente, a enorme quantidade de dados gerados pela Biologia Molecular (BALDI & BRUNNAK, 2001), e atualmente é imprescindível para a manipulação dos dados biológicos (SANTOS & ORTEGA, 2003).

Uma das áreas de atuação da bioinformática é no estudo e predição das estruturas protéicas. Métodos computacionais para predição de estruturas tridimensionais de proteínas são alvos de grande interesse e focos de muitas pesquisas nos dias atuais. A predição de estruturas 3D de uma proteína a partir de sua seqüência de aminoácidos remete a um problema científico fundamental e é frequentemente considerado um dos grandes desafios da química e da bioinformática (SCHWEDE & PEITSCH, 2008).

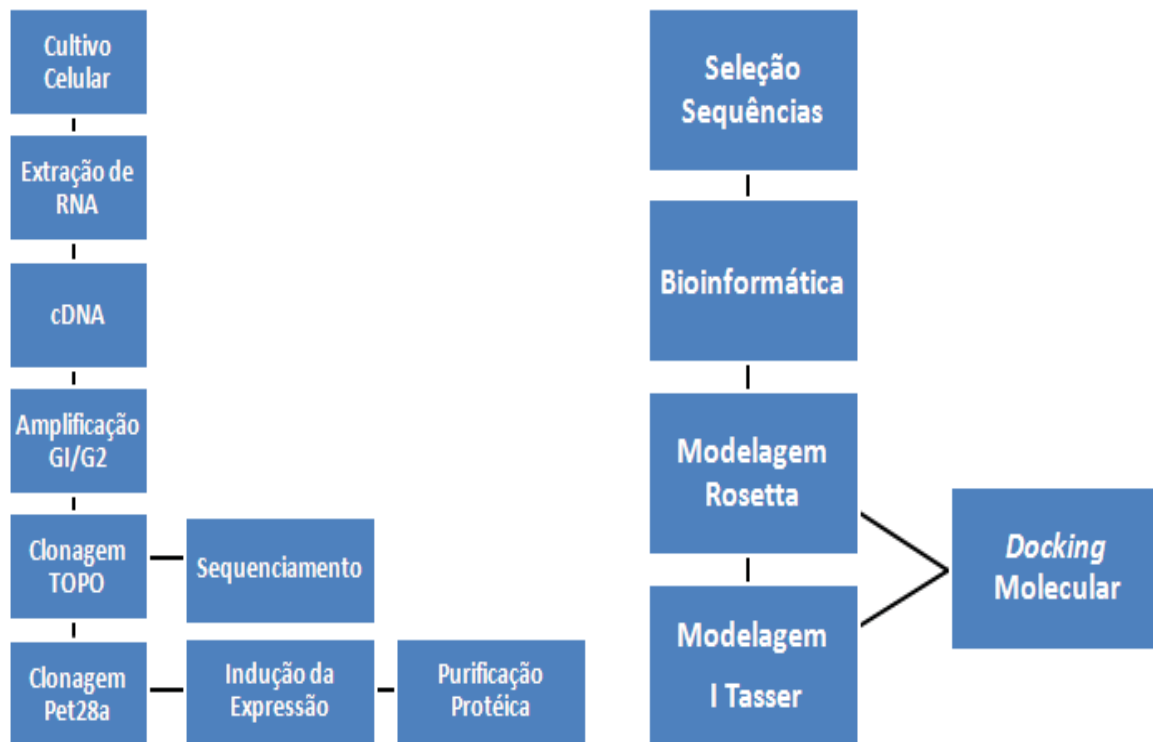
2. OBJETIVOS

O processo de ligação do vírus à célula hospedeira é realizado por intermédio da proteína G, sendo este mecanismo essencial para o vírus iniciar a entrada na célula. Conhecer sua estrutura nos permitirá caracterizar funcionalmente esta proteína.

- 1) Clonar, expressar e purificar os domínios solúveis da região N-terminal (1 a 38 aa) e C-terminal (67 a 298) com 231 aminoácidos da proteína G do hRSV tipo A;
- 2) Realizar a modelagem da proteína e analisar as características estruturais e funcionais da proteína G;
- 3) Buscar as interações da heparina com a Proteína G por *Docking Molecular*.

3. MATERIAL E MÉTODOS

Toda a metodologia utilizada ao longo do projeto foi resumida e esquematizada no fluxograma abaixo.



Fluxograma: O fluxograma ilustra as etapas a serem desenvolvidas para realização do presente projeto.

3.1. Análise experimental

3.1.1. Extração do RNA e transcrição reversa

O RNA foi extraído de amostras da cultura celular Hep2 inoculadas com hRSV. O cDNA foi transcrito a partir de 4µl de RNA extraído, diluído em tampão contendo 500mM de Tris-HCl pH 8.3 a 25°C; 75mM de KCL; 3mM de MgCl₂ (*Hight capacity* - Applied Biosystems), 10 mM de Diltreitol (Applied biosystems), 200 U de Transcriptase Reversa *Hight capacity* (Applied biosystems), 40U de Inibidor de Ribonuclease (RNAout Invitrogen), 50pmol de *random primer*, 1,5 mM de cada dNTP (Applied biosystems) e água Milli Q tratada com DEPC q.s.p. 40µl. A mistura foi incubada em um termociclador a 25°C por 10 minutos e a 37°C por 120 minutos.

3.1.2. Desenho dos oligonucleotídeos iniciadores

Para que as regiões correspondentes aos domínios solúveis da região N-terminal (1 a 38 aa) e C-terminal (67 a 298) com 231 aminoácidos da proteína G fossem amplificadas, foi necessário o desenho de pares de oligonucleotídeos iniciadores específicos para cada domínio.

Foram selecionadas sequências da proteína G do hRSV disponíveis no banco de dados *GeneBank*, as quais foram alinhadas utilizando-se o programa BioEdit (HALL et al., 1999). De posse do alinhamento, foram selecionadas as regiões amino e carboxi terminal da proteína G, que por sua vez, serviram de molde para o desenho dos oligonucleotídeos senso e anti-senso de cada domínio, totalizando 2 pares (G1+ e G1-, G2+ e G2-). As seqüências de cada par de *primers* podem ser visualizadas na Tabela 1. Na seqüência dos *primers* pode-se visualizar em vermelho a região de ancoragem da enzima de restrição a ser utilizada (Xho I – Fermentas), e em verde o sítio de restrição que será reconhecido por essa enzima.

Tabela 1. Características principais dos oligonucleotídeos iniciadores do gene da proteína G do hRSV.

<i>Primer</i>	Seqüência	Tamanho	Temp melting	% CG	Enzima Restrição
rRSVA G1+	CCGCTCGAGATGTCCAAAAACAAGGACCAA CGCA	34	73,5	59,2	Xho I
RSVA G1-	CCGCTCGAGTCATATAGATTTAAGATTTAAC	31	63,0	35,5	Xho I
RSVA G2+	CCGCTCGAGATGCACAAAGTCACACTAACA ACT	33	70,09	48,5	Xho I
RSVA G2-	CCGCTCGAGTTATTTGCCCCAGAGTTGATTT TG	33	70,03	48,5	Xho I

3.1.3. Amplificação do endo e ectodomínio da Proteína G.

A reação de PCR para ambos os domínios da proteína G foi feita utilizando-se tampão 1x concentrado, 1,5mM de cloreto de magnésio, 1µl de DNA (100ng), 0,5µM de cada *primer*, 0,2mM de dNTPs, 1,25U de *Taq DNA polymerase* e água para um volume final de 25µl. A reação para o endodomínio foi submetida à uma desnaturação inicial a 94°C por 5 minutos, seguida de 35 ciclos de 94°C por 1 minuto, 65°C por 1 minuto e 72°C por 1 minutos, finalizando com um ciclo de 72°C por 10 minutos para extensão final. Já a reação para o

ectodomínio foi submetida à desnaturação inicial a 94°C por 5 minutos, seguida de 35 ciclos de 94°C por 1 minuto, 68°C por 1 minuto e 72°C por 1,5 minutos, finalizando com um ciclo de 72°C por 10 minutos para extensão final.

Os produtos amplificados foram analisados por eletroforese em gel de agarose 0,8%. Uma mistura de 5µl de produto amplificado e 1µl de *Loading Buffer* 6x foi aplicada no gel de agarose 0,8% em tampão Tris-Borato-EDTA 1x (90 mM de Tris-Borato e 2mM de EDTA pH 8.0) com adição de Brometo de etídeo em concentração final de 0,5 µg/mL.

3.1.4. Clonagem dos domínios da Proteína G em TOPO®XL PCR Cloning Kit

Após a amplificação dos domínios solúveis da região N-terminal (1 a 38 aa) e C-terminal (67 a 298) da proteína G, estes foram clonados utilizando-se o kit de clonagem TOPO® XL PCR Cloning Kit (INVITROGEN), segundo recomendações do fabricante.

Nessa primeira clonagem as regiões gênicas de interesse foram inseridas no vetor pCR®-XL-TOPO® (Figura 5) e transformadas em bactéria quimicamente competente TOP10. As bactérias que receberam os plasmídios recombinantes foram semeadas em placas contendo ágar e antibiótico kanamicina (50µg/mL), e submetidas à temperatura de 37°C em estufa para crescimento *overnight* e posterior verificação de clones positivos.

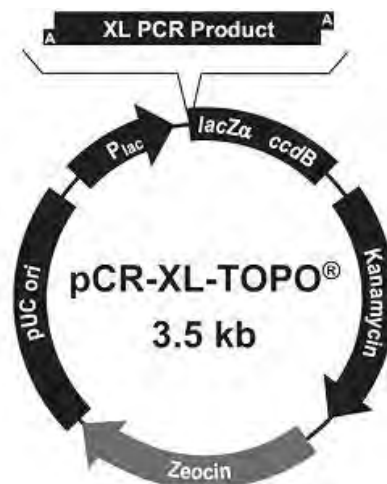


Figura 5. Mapa de restrição do plasmídeo pCR-XL-TOPO® (INVITROGEN), destinado à formação do vetor de clonagem de proteínas recombinantes. Adaptada do protocolo TOPO® XL PCR Cloning Kit.

Após a incubação, 96 colônias de cada grupo de transformantes pCR®-XL-TOPO–G1, pCR®-XL-TOPO–G2 foram selecionadas para triagem inicial quanto à presença de clones positivos. Para tanto, cada um dos 96 transformantes foi inoculado em 180µL de

meio LB contendo antibiótico kanamicina (50,0 µg/mL), em placas com fundo em U, seguido por incubação a 37 °C, sob agitação de 120 rpm por 16 horas. Em seguida foi analisada a positividade dos clones, utilizado-se 2,0 µL de cada cultura para reação de PCR, usando os *primers* de amplificação de cada domínio, como descrito no item 3.1.3. Para verificar a amplificação dos produtos de PCR foi feita eletroforese em gel de agarose 0,8%.

Após a verificação da positividade dos clones pCR[®]-XL-TOPO-G1 e pCR[®]-XL-TOPO-G2, estes foram cultivados em 4 ml de meio LB contendo 50 µg/mL de kanamicina, crescidos a 37°C por 14 horas. Os plasmídios recombinantes foram purificados utilizando o QIAprep[®] Spin Miniprep Kit (QUIAGEN), segundo recomendações do fabricante. Após a extração, os DNAs plasmidiais foram quantificados em espectrofotômetro.

3.1.5. Sequenciamento dos produtos clonados

Para certificar que os domínios da proteína G do hRSV foram clonados no vetor pCR[®]-XL-TOPO de forma correta, foram padronizadas as reações de seqüenciamento dos produtos de mini-prep. Para a reação foram utilizados 100ng de plasmídio contendo o gene de interesse clonado e purificado (mini-prep), 10pmol de cada *primer* (reações separadas para *primers* senso e anti-senso), *Big Dye Buffer* 1X, 2µL de *Big Dye reaction* e água para um volume final de 10µL. A ciclagem utilizada para os clones pCR[®]-XL-TOPO-G1 foi de 96°C por 1 minuto, seguida de 25 ciclos de desnaturação de 96°C por 15 segundos, 55°C por 15 segundos (temperatura de anelamento) e 4 minutos a 60°C (temperatura de extensão).

Aos 10µL da reação de Big Dye foram acrescentados 2,6µL de EDTA 0,1M (concentração final de 21mM). Adicionou-se 9µL de Acetato de Sódio 1M (concentração final de 0,42mM) e em seguida adicionou-se 50µL de etanol 100%. A placa foi brevemente homogeneizada em agitador *vórtex* e incubada em temperatura ambiente por 15 minutos. A placa foi centrifugada a 4.000rpm por 30 minutos à 4°C e o sobrenadante descartado. A placa foi centrifugada invertida por 1 minuto a 1.000rpm à 4°C. Foram acrescentados 70µL de etanol 70% e a placa foi novamente centrifugada a 4.000rpm por 15 minutos à 4°C e o sobrenadante foi descartado. A placa foi centrifugada invertida por 1 minuto a 1.000rpm à 4°C. Em seguida a placa foi seca à temperatura ambiente por cerca de 30 minutos ao abrigo da luz.

Em cada poço da placa foram acrescentados 10µL de formamida HiDi (Applied Biosystems), inclusive os poços sem reação, pois o seqüenciador reconhece apenas múltiplos de 16. A placa foi centrifugada invertida por 1 minuto a 1.000rpm à 4°C e em seguida, desnaturada por 5 minutos a 95°C em termociclador. Foi então incubada em gelo por 2 minutos e inserida no seqüenciador Sequencer Analyser 3130 (Applied Biosystems). As

5mM, DTT 5mM, PEG-8000 25%), 2 µL de enzima e água Milli-Q para volume final de 20,0 µL. A razão molar inserto/vetor usada foi 3:1, de acordo com a seguinte equação:

$$\frac{50 \text{ ng do vetor} \times \text{tamanho do inserto (Kb)}}{\text{Tamanho do vetor (Kb)}} \times \text{RAZÃO MOLAR inserto} = \text{ng do inserto vetor}$$

As reações de ligação foram incubadas a 16°C por 16h. Finalizado este período, seguiu-se à transformação bacteriana. Para tanto, 5,0µL dos produtos de (construções Pet28a-G1) foram gentilmente adicionados a 50,0µL de *E. coli* DH5α quimicamente competente. Esta mistura foi incubada no gelo por 30 minutos, com subsequente choque térmico a 42°C por dois minutos, seguido de incubação no gelo por dois minutos.

Foram adicionados 250,0µL de meio SOC (meio de cultura enriquecido para crescimento bacteriano) à temperatura ambiente seguido de incubação a 37°C, com agitação de 100rpm, por uma 1h30min. Ao final da incubação, a mistura foi centrifugada por um minuto a 8.000 rpm e o precipitado de células foi gentilmente ressuspendido e plaqueado em meio seletivo, contendo o antibiótico kanamicina. Após secagem, as placas foram incubadas a 37°C, durante 16 horas, para crescimento dos transformantes.

Após a incubação, 48 colônias de cada grupo de transformantes pet28A-G1 foram selecionadas para triagem inicial quanto à presença de clones positivos. Para tanto, cada um dos 48 transformantes foi inoculado em 4,0 mL de meio LB contendo antibiótico kanamicina (50,0 µg/mL), seguido por incubação a 37 °C, sob agitação de 100 rpm. Decorrido o período de incubação (14-16h), 2,0 µL de cada cultura foram analisados por PCR usando os *primers* de amplificação do domínio G1, conforme apresentado no item 3.1.3. Para verificar a amplificação dos produtos de PCR foi feita eletroforese em gel de agarose 0,8%.

Os plasmídeos recombinantes das culturas positivas na PCR foram isolados utilizando o QIAprep® Spin Miniprep Kit (QIAGEN), segundo recomendações do fabricante. Após a extração, os DNAs plasmidiais foram quantificados em espectrofotômetro. As construções Pet28a-G1 foram clonadas em bactérias quimicamente competentes da linhagem BL21-pLysS, para posterior ensaios de indução da expressão protéica.

3.1.7. Indução da expressão do peptídeo G1

Foram selecionados 3 tubos de pré-inóculos da colônia selecionada contendo 5mL de meio LB líquido juntamente com 50µg/mL de kanamicina e 34µg/mL de cloranfenicol.

Após incubação a 37°C sob agitação de 100 rpm por aproximadamente 16h, o pré-inóculo foi adicionado a um novo meio de cultura (20 mL de meio LB, 50µg/mL de kanamicina e 34 µg/mL de cloranfenicol) que foi incubado sob as mesmas condições.

Quando a cultura de células atingiu a densidade óptica de 0,6 em 600nm, foi coletada uma alíquota de 1ml de cada tubo, a qual foi centrifugada a 13.000xg durante um minuto e posteriormente o *pellet* resultante foi diluído em 10,0µL de tampão de amostra (Tris 60 mM, pH 6,8; Glicerol 25%; SDS 2%; 2-Mercaptoetanol 14,4 mM; Azul de Bromofenol 0,1%). Para a desnaturação das proteínas, estas foram incubadas em termociclador a 96°C, por dez minutos.

Com o restante da cultura foi feita a indução de expressão, com três concentrações diferentes de isopropil-β-D-tiogalactopiranosida (IPTG): 0,1mM, 0,2mM e 0,4mM, a 37°C sob agitação de 100rpm. Neste período, em intervalos de uma hora, foram feitas coletas de 1,0 mL de cada cultura para posterior verificação da expressão protéica em cada tempo, e em cada concentração de IPTG. Em cada coleta, as células foram separadas do meio por centrifugação, o *pellet* resultante foi diluído em tampão de amostra (20, 30, 40 e 50µl, em 1h de crescimento, 2ho, 3h e 4h respectivamente), e posteriormente lisadas a 96°C por 10 minutos.

Eletroforese em gel de poliacrilamida (SDS-PAGE) 15% foi realizada para verificação da expressão da proteína G1, utilizando-se 10,0µL de amostra previamente desnaturada e 5,0µL de marcador de peso molecular PageRuler™ Prestained Protein Lader (10 - 170 KDa) (Fermentas). O procedimento de eletroforese teve duração de 3 horas a uma voltagem de 120V. Após esta etapa o gel foi corado com solução corante (1g de *comassie blue R-25*, 450ml metanol 100ml ácido acético glacial e 450ml de água) por um período de 5 horas e em seguida o gel foi descorado em solução descorante (100ml metanol, 100ml ácido acético glacial e 800ml de água)por um período de 5 horas.

3.1.8. Purificação protéica com resina de afinidade

Como o vetor Pet28a permite a expressão protéica com uma cauda de 6 histidinas em sua porção C-terminal, optou-se por purificação com resina de níquel, a qual apresenta afinidade com a histidina. A proteína teve o mesmo tratamento descrito no item 3.1.7., após a sonicação, o lisado foi centrifugado por 20 minutos a 4.000 x g a 4°C. Os passos da purificação protéica foram realizados em condições nativas.

A matriz utilizada foi His60 Ni Superflow Resin (CLONTECH). Esta foi previamente ressuspensa para evitar perda de resina, e posteriormente foram adicionados 2ml na coluna, por aproximadamente 30 minutos, até que ocorresse seu empacotamento. A

seguir foram feitos tratamentos com diferentes tampões, e em cada aplicação foram recolhidas alíquotas de 1 ml para posterior verificação do conteúdo das mesmas por SDS Page.

A resina foi primeiramente equilibrada com 10ml de tampão de equilíbrio (50 mM Fosfato de Sódio, 300mM de cloreto de Sódio e 20mM de imidazol; pH 7.4). O sobrenadante do lisado foi aplicado na coluna previamente equilibrada, recolhendo-se 3 alíquotas de 1ml cada contendo o conteúdo que atravessou a coluna (F1, F2, F3). Posteriormente a coluna foi lavada com 10ml de tampão de lavagem (50 mM Fosfato de Sódio, 300 mM de cloreto de Sódio e 40 mM de imidazol; pH 7.4), recolhendo-se 10 alíquotas de 1ml cada contendo o conteúdo que atravessou a coluna (L1 - L10).

A amostra foi eluída com 10ml de tampão de eluição (50mM Fosfato de Sódio, 300mM de cloreto de Sódio e 300mM de imidazol; pH 7.4), coletadas em 10 alíquotas de 1ml cada (E1–E10). A resina foi regenerada para posterior utilização com tampão de regeneração (20mM MES, 0.3M NaCl; pH 5.0) e armazenada a 4°C.

As amostras resultantes de todas as alíquotas foram desnaturadas com tampão de amostra, assim como o pellet do inóculo, e posteriormente foram verificadas em SDS-PAGE 17%, coradas com nitrato de prata. As etapas de coloração seguiram os seguintes passos: primeiramente o gel foi incubado por 30 minutos em solução fixadora (30ml etanol, 10ml ácido acético, 60ml de água), posteriormente incubado por 30 minutos em solução redutora (0,1g de tiosulfato de sódio, 250µl de acetato de sódio 4M, 35ml de etanol, 500µl de glutaraldeído e 64ml de água). Foram feitas 3 lavagens de 30 minutos cada com água destilada para retirar os resíduos das soluções anteriores, e o gel foi incubado por 20 minutos em solução de prata (0,1g de nitrato de prata, 100µl de formaldeído e água q.s.p. 100ml). A seguir o gel foi rapidamente lavado com água e revelado com solução reveladora (2,5g de carbonato de sódio, 50µl de formaldeído e água q.s.p. 100ml). Assim que as bandas começaram a aparecer a reação foi interrompida adicionando-se ácido acético glacial.

3.2. Modelagem e Análise Teórica

3.2.1 Seleção das seqüências da Proteína G do banco de dados Genbank

Foi realizada uma busca de seqüências da proteína G do hRSV no Banco de Dados Internacional GenBank, mantida pelo National Center for Biotechnology Information (NCBI). (WHEELER et al., 2007). Foram encontradas 54 seqüências, algumas contendo os 298 aa da proteína G, algumas contendo o genoma completo do hRSV, nas quais foi

selecionada a região correspondente à proteína G, e algumas conteúdo número menor de aminoácidos, que no alinhamento mostraram compreender uma região específica da proteína.

3.2.2. Análise das seqüências pelos programas de bioinformática

Todas as seqüências foram salvas em formato FASTA em arquivos do bloco de notas para serem importadas pelo programa BIOEDIT (HALL et al., 1999). Após a importação das seqüências escolhidas para o BIOEDIT, aquelas que eram seqüências de nucleotídeos do genoma completo do hRSV foram selecionadas as regiões correspondentes à proteína G. Em seguida, as seqüências nucleotídicas foram traduzidas em aminoácidos, e a ferramenta de alinhamento CLUSTALW foi aplicada, sem alterar nenhuma configuração padrão. Foram utilizados três programas de predição teórica de algumas características protéicas, como segue na Tabela 2:

Tabela 2: Características avaliadas por programas de bioinformática

Características Estruturais e Funcionais				
Estruturas 2 ^{as}	Pontes Dissulfeto	Modificações pós Traducionais	P. isoelétrico	Tipos de Interações
PredictProtein	PredictProtein	PredictProtein	Compute PI ProtParam	PredictPotein

O aplicativo PREDICTPROTEIN (ROST, B. et al., 2004) é um conjunto de vários programas que prediz a estrutura secundária da proteína e do motivo de interesse, bem como determina possíveis sítios de glicosilação, sulfatação e fosforilação, possíveis seqüências de localização nuclear, possíveis regiões de formações de pontes dissulfeto, além de verificar sítios de interação com outras proteínas ou até mesmo DNA. Todos esses resultados fornecem informações para uma interpretação sobre a função e o tipo de interação que a proteína pode realizar. Dois outros programas utilizados foram o COMPUTE PI (BJELLQVIST et al., 1993) e PROTPARAM (GASTEIGER et al., 2005) que possibilitam calcular o valor do ponto isoelétrico da proteína estudada, bem como seu peso molecular.

3.2.3. Estrutura da heparina

Como existem estudos experimentais que sugerem uma interação entre a proteína G do hRSV e a heparina (FELDMAN et al., 1999, BOURGEOIS et al., 1998) também foram realizados estudos teóricos para verificar essa interação, e para tal, utilizou-se a estrutura

depositada no bando de dados de proteínas (PDB- *Protein Data Bank*) da estrutura de heparina resolvida por ressonância magnética nuclear, de código 1HPN.

O PDB, administrado pelo RCSB (Bioinformatics) fornece informações sobre estruturas tridimensionais que compreendem, principalmente, proteínas e ácidos nucleicos. Cientistas de todo o mundo depositam estruturas e informações de suas proteínas, e pode ser acessado (<http://www.rcsb.org>) (PDB ANNUAL REPORT, 2009).

3.2.4. Modelagem da Proteína G pelo programa ROSETTA

O melhor método para a modelagem da estrutura de uma proteína depende se ela possui sequência homóloga ao de uma proteína de estrutura conhecida. Se existe essa homologia, modelos relativamente precisos podem ser construídos usando a estrutura conhecida como molde. Na ausência de tal característica, os modelos podem ser construídos utilizando métodos de previsão *ab initio*, que não dependem de uma estrutura modelo. O método *ab initio* foi escolhido uma vez que a modelagem comparativa não foi possível por causa da baixa homologia e identidade da sequência da proteína G com outras sequências de proteínas de estruturas já conhecidas.

Para a modelagem da Proteína G do hRSV foi utilizado um programa automático de predição de estruturas, o ROSETTA (CHIVIAN, 2003). O Rosetta é um programa de predição de estruturas de proteínas totalmente automatizado que usa o método de inserção de fragmentos para a modelagem protéica.

A biblioteca de fragmentos é composta por estruturas determinadas experimentalmente, que o programa utiliza para orientar a busca do espaço conformacional e as previsões das estruturas, bem como, as conformações de *loop* em modelos de homologia. Essa biblioteca foi obtida para a proteína G tipo A pelo seguinte processo:

- **1º passo:** A sequência protéica é dividida em fragmentos (3 e 9 aminoácidos);
- **2º passo:** Busca de estruturas determinadas experimentalmente similares aos fragmentos da proteína a ser modelada;
- **3º passo:** Filtragem de todos os fragmentos de estruturas determinadas experimentalmente. São mantidos aqueles com similaridade igual ou superior a 20%.

Tendo a biblioteca de fragmentos, os modelos são gerados utilizando o algoritmo de Monte Carlo seguindo o seguinte protocolo:

- 1 - Uma posição na sequência da proteína G é escolhida;
- 2 - Ocorre a inserção de todos os fragmentos candidatos a esta posição;
- 3- O melhor fragmento (que apresentou maior semelhança com aquela região,é mantido;

4 - Uma nova posição na sequência é escolhida e os passos anteriores são repetidos até a montagem completa da proteína.

Para a construção dos modelos, o Rosetta faz uso do método de Monte Carlo. Este método utiliza a seguinte estratégia: uma conformação inicial é formada, com base na biblioteca de fragmentos, com uma determinada energia. Um segundo modelo é predito, se este possui uma conformação de menor energia que o anterior ele é aceito e o anterior é descartado. Este processo é realizado até que os mil modelos tenham sido comparados uns com os outros e a conformação de menor energia tenha sido obtida. No término do processo, o Rosetta fornece os possíveis modelos da proteína. A quantidade de modelos pode ser escolhida e quanto maior a mesma, melhor a qualidade do modelo a ser selecionado. Para a modelagem da proteína G foram gerados 1000 modelos.

O programa MaxCluster foi utilizado para avaliar os modelos gerados, pois fornece vários métodos diferentes para a seleção das estruturas representativas de um conjunto de proteínas: o agrupamento do vizinho mais próximo (clustering Restricted Nearest Neighbour - RNN) e 3D-Jury.

O 3D-Jury utiliza o algoritmo MaxSub (SIEW et al., 2000) para a seleção do melhor modelo. O algoritmo MaxSub visa identificar o maior subconjunto de átomos de C (alfa) de um modelo que se sobrepõem "bem" sobre as outras estruturas, e produz um resultado único normalizado que representa o modelo de melhor qualidade. Desta forma, a proteína que compartilha a maioria de sua estrutura com as outras estruturas do conjunto de dados ficará no topo do ranking dos modelos.

O método de Vizinhos mais Próximos (RNN) é baseado no método de SHORTLE et al. (1998), e consiste em uma estratégia para prever a estrutura da proteína baseando-se na hipótese de que há um maior número de conformações de baixa energia em torno da conformação correta. Desta forma, analisam-se os cálculos da raiz do desvio quadrático médio (rms) das coordenadas dos átomos de C (alfa) entre cada conformação com todos os outros das demais conformações do conjunto.

Duas proteínas são consideradas parte do mesmo grupo (*cluster*), se elas estão mais próximas do que um limiar de corte, ou seja, elas são quase vizinhas. Para cada proteína no conjunto de dados, o número de vizinhos mais próximos do que o ponto de corte são incluídos no *cluster* da proteína. Assim, as proteínas podem ser atribuídas a vários *clusters*. O item com a maioria dos vizinhos é escolhido como o centro do *cluster* e este é retirado do conjunto. Isso envolve a remoção do centróide e seus n vizinhos mais próximos. O processo é

repetido até que os *clusters* contenham menos do que um número mínimo de membros. Por fim, os representantes são selecionados em cada *cluster*.

3.2.5. Modelagem da Proteína G pelo programa I-TASSER

Os métodos de *threading* compartilham algumas características com a modelagem comparativa: tais como a realização de alinhamentos e a utilização de informações sobre estruturas protéicas conhecidas. Porém, o método de *threading* utiliza similaridades menores que 30% entre as seqüências.

A idéia básica dos métodos de *threading* é ajustar a conformação do esqueleto peptídico da seqüência alvo a uma coleção de proteínas de referência (biblioteca de enovelamentos). Um valor de “qualidade de ajuste” é calculado para cada estrutura gerada.

Atualmente esses métodos são largamente utilizados principalmente porque se acredita que exista um número limitado de enovelamentos na natureza, seja por resultado da evolução ou por restrições de natureza física da cadeia peptídica. Existe então uma boa chance (atualmente entre 70 e 80%) de que uma proteína com um enovelamento similar ao da seqüência alvo já tenha sido estudada por cristalografia de raios-X ou RMN. Atualmente existem aproximadamente 1100 enovelamentos diferentes conhecidos (ORENGO et al., 1997; REEVES et al., 2006).

O método I-TASSER (ZHANG et al., 2007) utiliza as informações de estruturas de referência, de forma idêntica ao *threading*, em um algoritmo especial para realizar múltiplos alinhamentos da proteína alvo a perfis tridimensionais de estruturas de referência, que podem apresentar baixa similaridade entre si e com a seqüência alvo.

Em seguida as estruturas resultantes da etapa de *threading* sofrem um processo de agrupamento e os maiores grupos (motivos estruturais) passam por uma etapa de refinamento por meio de simulações de Monte Carlo (SWENDSEN e WANG, 1986).

3.2.6. Validação dos modelos

Após a obtenção dos modelos, é necessário identificar os ajustes necessários relacionados à escolha das estruturas de referência, e ao alinhamento entre seqüência-alvo e molde. Caso o modelo seja caracterizado de baixa qualidade, todo o processo de modelagem deve ser revisto, no intuito de se melhorar o alinhamento inicial ou utilizar outros métodos.

O software utilizado para a validação dos modelos gerados foi o PROCHEK (LASKOWSKI et al., 1993), o qual avalia diversos parâmetros estereoquímicos, gerando o gráfico de Ramachandran (RAMACHANDRAN & SASISEKHARAN, 1968), que é muito

útil, pois define os resíduos de aminoácidos que se encontram nas regiões energeticamente mais favoráveis e desfavoráveis, além de orientar a avaliação da qualidade para modelos teóricos e experimentais de proteínas.

A estrutura protéica possui três ângulos principais, ω , ϕ e ψ . Como o ângulo ω é fixo os ângulos ϕ e ψ são os responsáveis por toda a variação conformacional da cadeia principal, pois tornam a cadeia polipeptídica flexível. Devem-se desconsiderar os resíduos de glicina, já que eles não possuem cadeia lateral, os de prolina, nos quais o Carbono alpha está ligado à cadeia lateral, e os resíduos das extremidades (C-terminal e N-terminal), que apresentam padrões estereoquímicos diferentes dos outros resíduos (SILVA, 2008).

3.2.7. Docking Molecular da interação proteína G com a heparina

Com o objetivo de identificar os sítios ligantes na proteína, esta foi submetida ao programa Q-SiteFinder (LAURIE et al., 2005). Estes sítios seriam possíveis locais onde os ligantes se posicionariam para interagir com maior afinidade, segundo interação de Van der Waals, com o sítio receptor. O programa detecta cavidades na superfície de proteína, particularmente na região do sítio receptor, tipicamente onde os compostos se ligariam.

Os *dockings* entre a proteína G e heparina foram realizados através dos programas PATCHDOCK (DUHOVNY et al., 2002) e AUTODOCK (MORRIS et al., 1998). O algoritmo PatchDock é inspirado no reconhecimento de objetos e técnicas de segmentação de imagens. O *docking* pode ser comparado à montagem de um quebra-cabeça. Para resolver o quebra-cabeça ele tenta combinar duas peças, escolhendo uma e procurando sua complementar. O PatchDock emprega uma técnica similar. Dadas duas moléculas, suas superfícies são divididas em áreas de acordo com a forma superficial. Essas áreas são identificadas e sobrepostas. O algoritmo tem três fases principais:

Representação da Forma Molecular - Calcula-se a superfície da molécula. Em seguida, aplica-se um algoritmo de segmentação para detecção de áreas geométricas (côncavo, convexo e plana – Figura 7). As áreas são filtradas, e somente as áreas da superfície que têm elevada probabilidade de pertencer a um sítio de ligação são mantidas.

Correspondência das Áreas da Superfície – Aplicam-se técnicas de correspondência para combinar as áreas detectadas na etapa anterior. Áreas côncavas são combinadas com convexas e vice-versa e áreas planas com qualquer tipo de área.

Filtragem e Pontuação - Complexos candidatos são analisados. Descartam-se os complexos de penetrações inaceitáveis dos átomos do receptor para os átomos do ligante. Os restantes são classificados de acordo com uma pontuação (*score*) da forma de complementaridade.

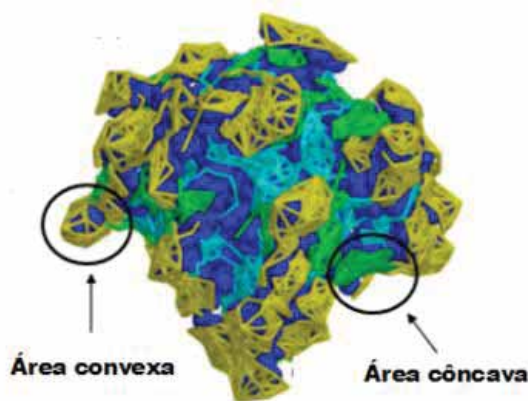


Figura 7. Esquemas ilustrativo das áreas geométricas de uma proteína, determinadas pelo algoritmo de segmentação do programa PachtDock. Em amarelo estão representadas as áreas convexas e em verde as áreas côncavas.

Outro programa usado para ensaios de interação foi o Autodock, o qual utiliza um método inspirado no conceito de evolução. Inicialmente, em uma população aleatória, cada membro corresponde a uma solução potencial do problema. Um membro da população é representado por um cromossomo, e cada cromossomo contém um número de genes, os quais correspondem aos valores de certas variáveis. No caso do *docking*, as variáveis são os ângulos de translação, rotação e de torção do ligante. Operadores genéticos (cruzamento e a mutação) são aplicados à população inicial para gerar uma nova população. No cruzamento, os genes de dois cromossomos distintos são trocados para gerar dois novos indivíduos e na mutação o gene é modificado aleatoriamente.

Indivíduos com melhores *scores* recebem uma melhor chance de serem selecionados como membros de uma nova população, e assim uma maior chance de sobreviver e reproduzir. Portanto, o ajuste médio aumenta de geração para geração até que em certo ponto o processo termina. Quanto menor o valor da energia livre de ligação, maior a afinidade ligante-proteína.

4. RESULTADOS E DISCUSSÃO

4.1. Amplificação do endodomínio, ectodomínio da Proteína G

Os genes de interesse foram amplificados resultando em bandas de tamanhos esperados, referentes a cada domínio. O resultado pode ser observado por eletroforese, que mostra a amplificação destes domínios, e podem ser visualizados na figura 8.

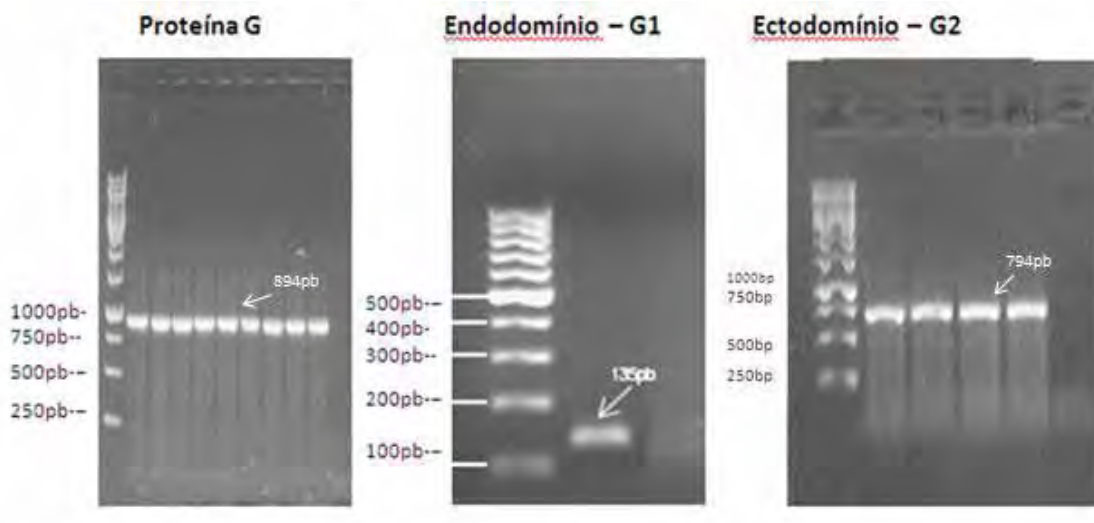


Figura 8. Amplificação da proteína G do hRSV e dos seus domínios solúveis da região N-terminal (1 a 38 aa - Endodomínio) e C-terminal (67 a 298 - Ectodomínio). M (Marcador Molecular de 100bp e 1000bp) G (Proteína G completa com aproximadamente 894pb) G1 (Endodomínio da glicoproteína G com aproximadamente 135 pares de base) G2 (Ectodomínio da glicoproteína G com aproximadamente 794 pares de base)

4.2. Clonagem dos domínios da proteína G em pCR®-XL-TOPO

Após a realização da transformação, o resultado mostrado na figura 9 demonstra a presença de colônias nas respectivas placas contendo meio de cultura LB e antibiótico kanamicina (50µg/ml). Os resultados de verificação da positividade dos clones são mostrados na figura 9, onde podemos observar a presença de bandas de 135pb e 794pb referentes aos transformantes contendo o domínio G1 e G2 respectivamente.

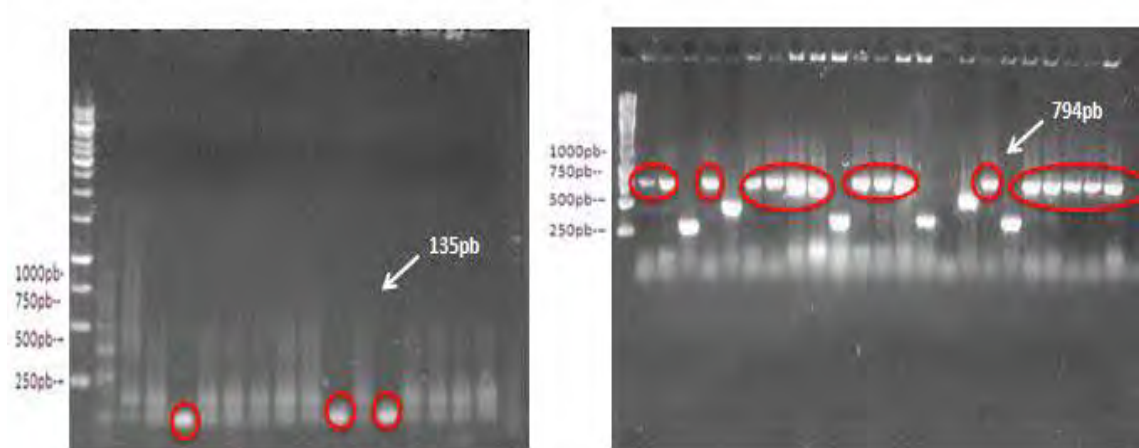


Figura 9. Eletroforese em gel de agarose 1% para triagem dos transformantes positivos contendo domínios do gene G do hRSV. A) Amplificação dos fragmentos de 135pb referentes aos transformantes pCR-XL-TOPO®-G1. PM de 1kb; B) Bandas de 794pb contendo os transformantes pCR-XL-TOPO®-G2. PM de 1Kb.

Os clones positivos foram utilizados para extração de plasmídeos, os quais foram seqüenciados, e o resultado pode ser visualizado na figura 10, mostrando o fragmento clonado. Desta forma os resultados obtidos permitem o prosseguimento dos experimentos de clonagem.

GTAAGAAATTAGGAAGGACCTGGGACACTCTCAATCATTATTATTTCATATCATCGTGCTTATATAAGTTAAATC
TTAAATCTATA

Figura 10. Sequência resultante do seqüenciamento do domínio G1 da proteína G do RSV utilizada nas clonagens.

4.3. Clonagem dos domínios da Proteína G em vetor pet28A

Após a certificação da clonagem por seqüenciamento os clones foram utilizados para subsequente reação de digestão os resultados mostram que a banda correspondente a 135pb foi liberada e que o plasmídeo foi linearizado (Figura 11).

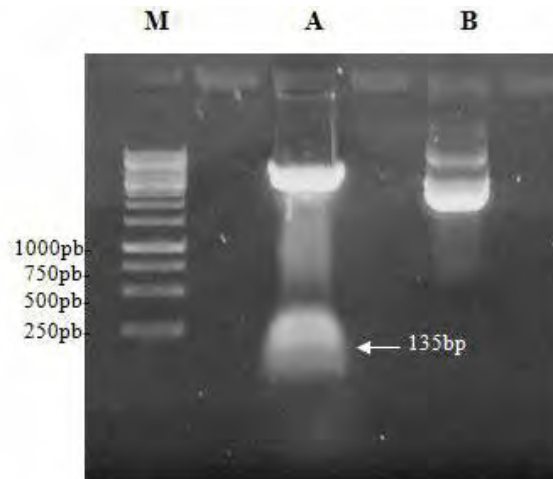


Figura 11. Perfil de restrição enzimática *Xho I*, A: dos clones positivos *pCR-XL-TOPO@-G1* da biblioteca selecionados da triagem. B: do vetor *pet28A (NOVAGEN)*. M: marcador de pares de base (1 Kb).

Após a transformação em bactéria *DH5 α* , foi verificada a positividade dos clones por PCR. Como pode ser mostrado na figura 12, o produto amplificado apresenta 135pb, correspondente ao fragmento G1. Em seguida estes plasmídios recombinantes foram extraídos por miniprep (Figura 13) e posteriormente transformados em bactérias *BL21* para expressão.

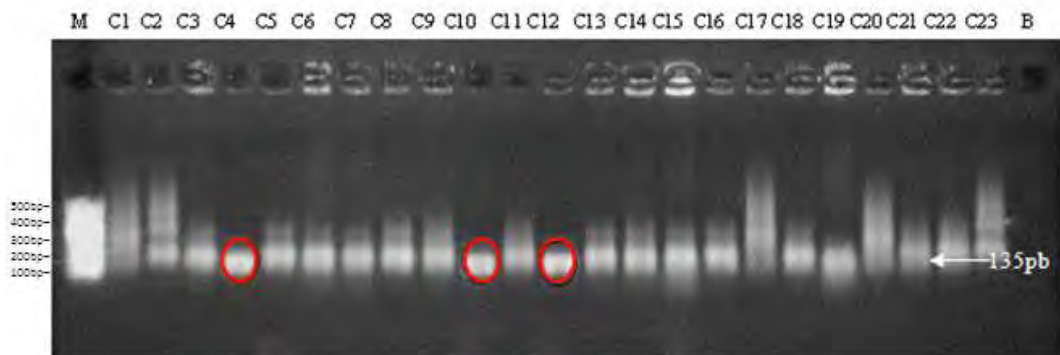


Figura 12. Eletroforese em gel de agarose 0,8% para triagem dos transformantes positivos contendo domínios do gene *G1* do *hRSV*. Amplificação dos fragmentos de 135pb referentes aos transformantes *pet28A- G1*. M: marcador de pares de bases (100 bp). B: controle negativo, sem amostra.

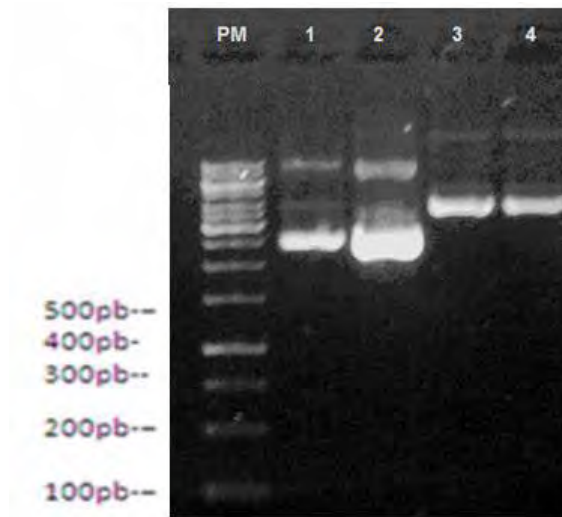


Figura 13. Eletroforese em gel de agarose 0,8% para verificação do produto obtido após a purificação dos plasmídeos contendo o peptídeo G1 clonado anteriormente, de 1- 4 clones purificados de plasmídeos contendo G1. (PM) marcador de pares de bases (100pb).

4.4. Análise da expressão da proteína G1 em bactéria

O domínio G1 foi expresso em células de *E. coli* linhagem BL21 pLysS induzidas nas seguintes condições:

- Concentrações de IPTG: 0,1mM; 0,2mM e 0,4mM.
- Coletas dos inóculos nos tempos: T0 (Antes de colocar IPTG), T1 (1h após adição IPTG) T2 (2h após adição de IPTG), T3 (3h após adição de IPTG) e T4 (4h após adição de IPTG).

A análise da expressão protéica foi realizada através de SDS-PAGE 15%. Após a indução da expressão foi verificado a expressão da proteína com peso molecular de aproximadamente 4,5 kDa, correspondente ao peso molecular da G1, para todas as concentrações de IPTG utilizadas conforme observado na figura 14.

Na figura A (clone 4) observa-se que a proteína foi produzida em baixa concentração, correspondente a 0,1 e 0,2mM de IPTG, sendo que esta sofre um leve aumento no clone 12, na concentração de 0,1mM de IPTG (figura B).

Na figura C, referente ao clone 12 podemos observar que a expressão protéica não pode ser visualizada muito bem. Na figura D a expressão protéica foi mais acentuada, e é referente aos clones 12 na concentração de 0,4mM de IPTG, e clone 24 na concentração 0,1mM de IPTG. A figura E demonstra expressão protéica menos acentuada, e é referente ao clone 24 nas concentrações 0,2 e 0,4mM de IPTG. A condição escolhida para expressão protéica em maior escala foi o clone 24, na concentração de IPTG de 0,1mM com crescimento a 37°C por 2 horas.

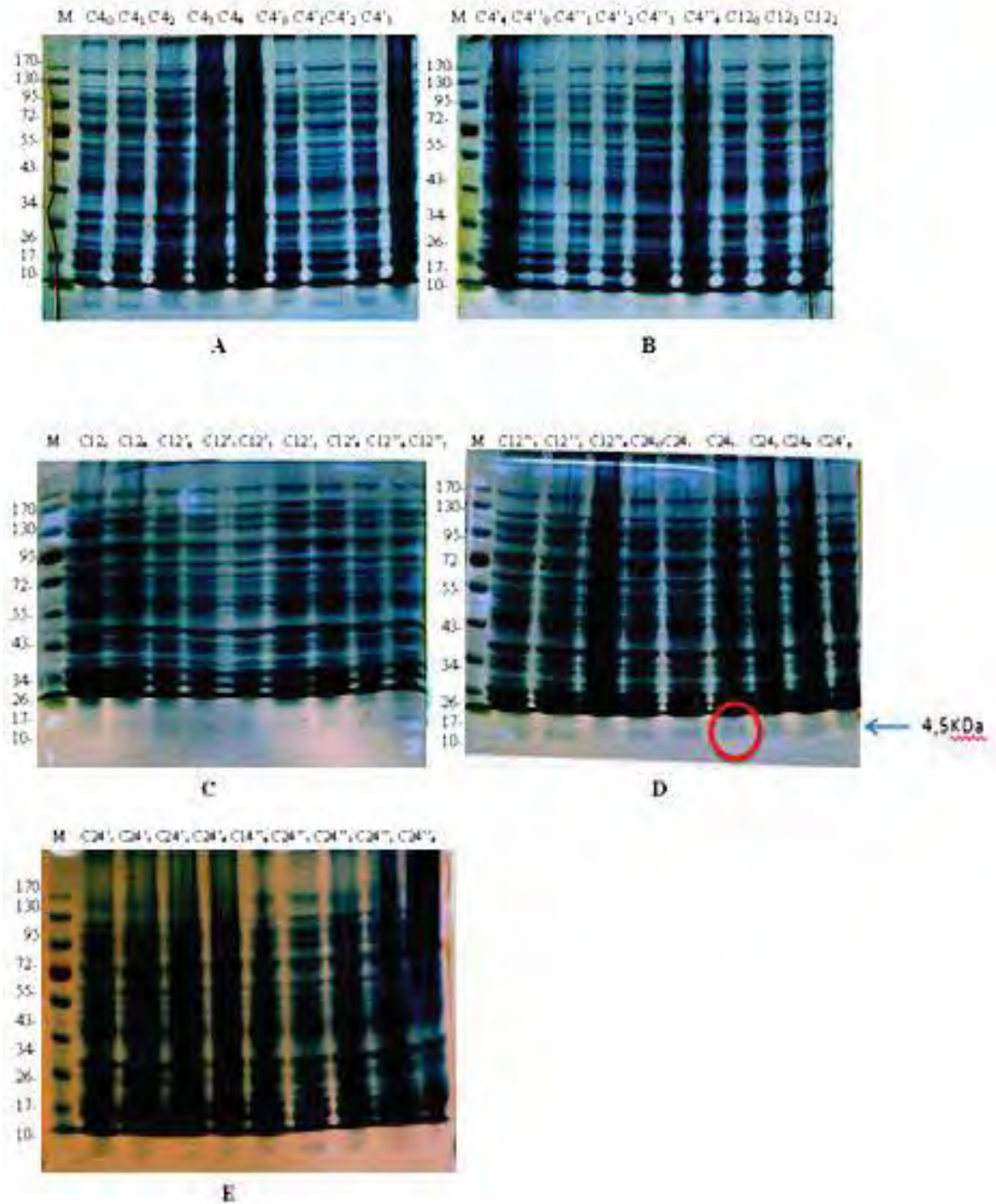


Figura 14. SDS Page 15%, mostrando a proteína G1 de ~ 4,5 KDa (Seta). M: Marcador de massa molecular (kDa). C4: Clone 4 com concentração de IPTG 0,1mM; C4': Clone 4 com concentração de IPTG 0,2mM; C4'': Clone 4 com concentração de IPTG 0,4mM; C12: Clone 12 com concentração de IPTG 0,1mM; C12': Clone 12 com concentração de IPTG 0,2mM; C12'': Clone 12 com concentração de IPTG 0,4mM; C24: Clone 24 com concentração de IPTG 0,1mM; C24': Clone 24 com concentração de IPTG 0,2mM; C24'': Clone 24 com concentração de IPTG 0,4mM.. Todos os clones estão sendo demonstrados nas coletas do tempo zero (antes da indução com IPTG), ao tempo quatro (4 horas após indução com IPTG).

As condições anteriores foram repetidas para o clone 24 para verificar se a proteína estava na forma solúvel ou insolúvel. Na figura 15 observa-se que na condição de expressão em 0,1 mM de IPTG a 37 °C por 2 horas a proteína estava na forma solúvel, como pode ser visualizada no gel quando se comparar a canaleta 1 que corresponde ao pellet, com as canaletas 2 a 7 que correspondem ao sobrenadante, em diferentes diluições.

Pode-se observar que a banda esta ausente no pellet e que proteína esta no sobrenadante em baixa quantidade na forma solúvel, sendo esta condição importante para a etapa de purificação.

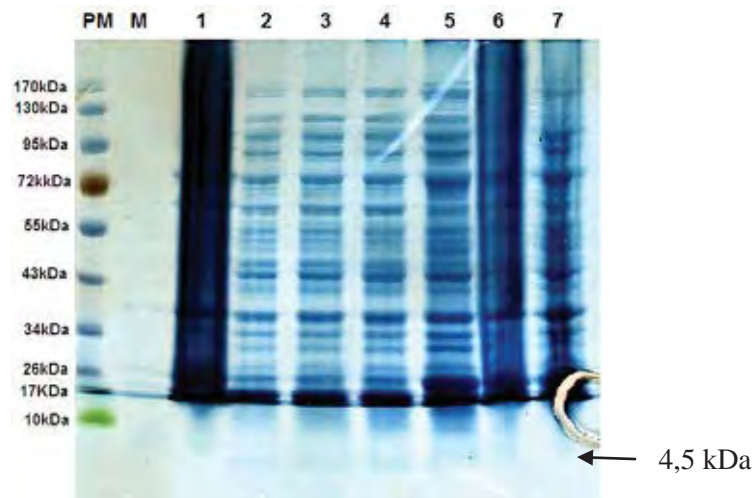


Figura 15. Gel de poliácridamida 17%, PM: Marcador peso molecular, M: meio de cultura, 1: pellet após indução por duas horas com 0,1mM de IPTG, 2 a 7: Sobrenadante após indução de 2 horas com 0,1mM de IPTG. Corado com comassí, seta apontando banda de 4,5 kDa.

4.5. Purificação do endodomínio da Proteína G

Mesmo em baixa quantidade tentamos purificar esta proteína como mostra a figura 24. Podemos observar no gel corado com prata que a proteína está no sobrenadante (s1). O protocolo de purificação precisa ser otimizado, as canaletas p1 p2 e p3 mostra ausência da proteína nas etapas de purificação, o que pode ser devido a baixa quantidade de proteína expressa.

Na figura 16 podem-se visualizar o gel de SDS-Page realizado para confirmar se a proteína havia sido purificada por resina de afinidade

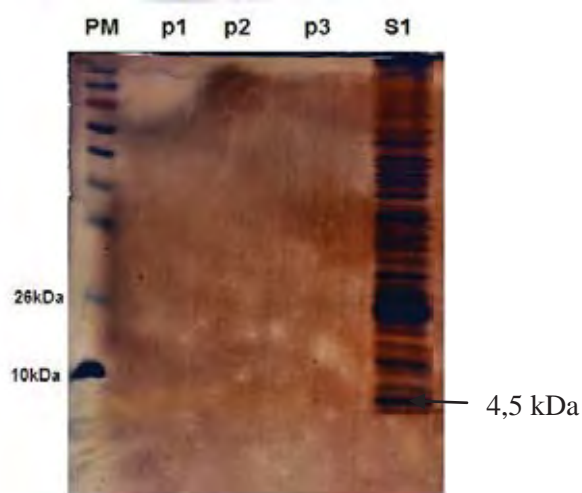


Figura 16. Gel de Poliacrilamida 17%, PM: Marcador peso molecular, p1, p2, p3: frações de purificação após indução por duas horas com 0,1mM de IPTG, S1: Sobrenadante após indução de 2 horas com 0,1mM de IPTG. Corado com nitrato de prata. Seta indicando a proteína de 4,5kDa.

4.6. Predição da estrutura secundária

A partir dos sítios selecionados, realizou-se uma predição das características particulares da proteína por meio do programa PREDICTPROTEIN. Informações sobre a composição de estruturas secundárias podem ajudar na interpretação sobre a função e o tipo de interações que a proteína pode realizar. Os resultados mostram a porcentagem teórica de estruturas secundárias (Tabela 3), sendo que a proteína inteira apresenta 85,5% de *coil*, 10% de folha beta e 4% de α -hélice. O ectodomínio apresentou 100% de *coil*, e o endodomínio 24% *coil* e 76% de α -hélice.

Tabela 3. Proporcionalidade da estrutura secundária da proteína G obtida pelo programa PredictProtein.

Proteína G	Hélice (%)	Folha (%)	Random coils (%)
Proteína Inteira	4	10	85
Ectodomínio	0	0	100
Endodomínio	76	0	24

As proteínas integrais de membrana são anfífilas: os segmentos imersos no interior apolar da membrana possuem resíduos predominantemente hidrofóbicos, enquanto as porções que se estendem para o meio aquoso estão em geral cobertas por resíduos polares (VOET et al., 2000).

Geralmente regiões hidrofóbicas que atravessam a membrana possuem uma conformação α -hélice, pois proporcionam que as cadeias laterais dos resíduos de aminoácidos projetem-se para a face externa da hélice e interajam com os lipídeos de membrana, formando uma estrutura muito estável (ALBERTS et al., 1994; LEHNINGER et al., 1999).

Como a proteína G é integral de membrana, ela possui uma região transmembrana, (do aminoácido 39 até 67, totalizando 29 aminoácidos, aproximadamente 9,7 % da proteína total). (GORMAM et al., 2001). Os resultados encontrados pelos programas, portanto, são coerentes com as características da proteína quanto à porcentagem teórica de alfa-hélices. A conformação em folhas beta é mais comum em proteínas globulares, e consiste em cadeias polipeptídicas vizinhas ligadas através de ligações de hidrogênio (diferentemente da ligação de hidrogênio dentro da própria cadeia, como é o caso das alfa-hélices), apresentando uma aparência plissada (VOET et al., 2000).

Regiões de *coils* são mais irregulares, sem estrutura secundária definida, e, portanto, mais difíceis de serem descritas. Essas regiões podem significar sequências protéicas que sofrem modificações pós traducionais (como glicosilações, por exemplo). Essa informação pode ser reforçada pela predição teórica de vários sítios de modificações sofridas pela proteína no ectodomínio (região predominantemente formada por *coils*), item a ser discutido na seção 5.7. A conformação da estrutura secundária obtidas pelos programas Rosetta e I-Tasser, do modelo representante da proteína G pode ser observada na figura 17.

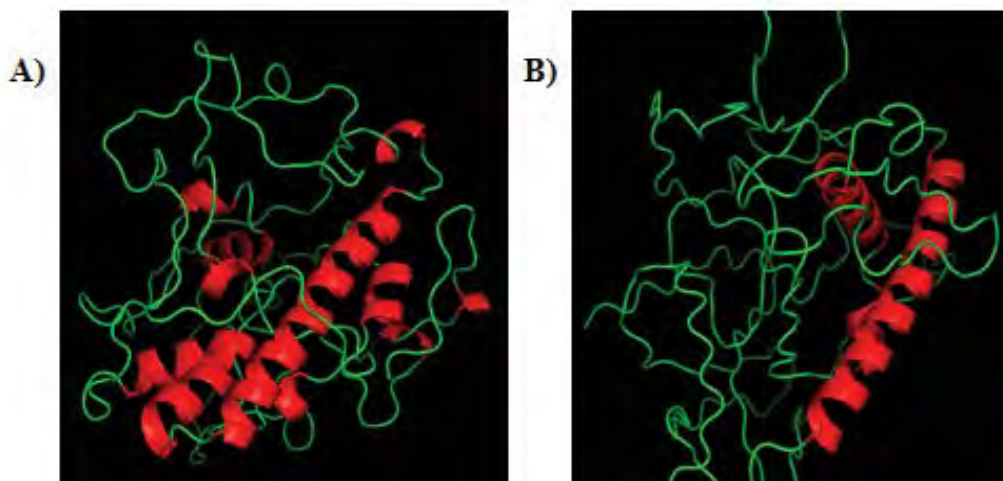


Figura 17. A) Modelo da proteína G do hRSV do tipo A, gerado pelo programa I-Tasser, B) Modelo da proteína G do hRSV do tipo A, gerado pelo programa ROSETTA. A proteína é composta predominantemente por *coils* (regiões em verde), e também por alfa-hélices (regiões em vermelho).

Na figura 17 observa-se que o modelo da proteína G apresentou algumas α -hélices e é composto principalmente de *coils*, o que confirma os resultados da predição teórica

realizada pelo PREDICTPROTEIN. A grande quantidade de *coils* é característica de proteínas que apresentam múltiplas funções celulares, como é o caso da proteína G do hRSV. Esta é uma proteína de membrana responsável para ligação do vírus a célula que será infectada e pode ter importante função de passar informações para o meio intracelular.

Com a sequência do endodomínio da proteína G (G1) obtida nos resultados experimentais através de seqüenciamento nucleotídico, foram preditos dois modelos de estrutura secundária, para verificar a semelhança com os modelos obtidos para a proteína G completa. Os resultados obtidos podem ser visualizados na figura 18.

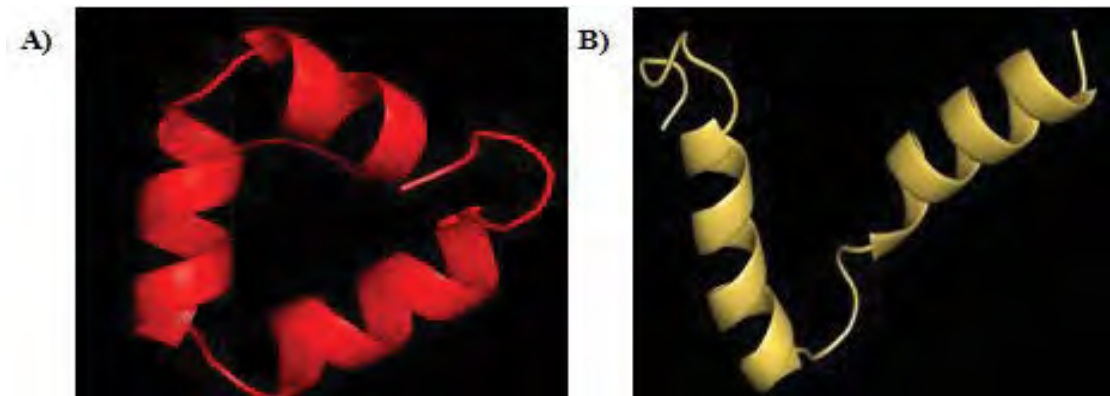


Figura 18. Modelo do endodomínio da proteína G do hRSV do tipo A, A) Modelo gerado pelo programa ROSETTA. B) Modelo gerado pelo programa I-Tasser, mostrando estruturas em *coil* e alfa-hélice.

Pode-se verificar que ambos os modelos gerados isoladamente para o endodomínio da proteína G do hRSV ficaram muito parecidos em composição secundária, e com os modelos gerados para a proteína G completa, o que reforça que o provável enovelamento dessa sequência seja esse. Os resultados teóricos obtidos pelos programas PREDICTPROTEIN, ROSETTA e I-TASSER mostram que o endodomínio da proteína G apresenta estrutura em *coil* e alfa-hélice. Provavelmente essa região da proteína G que se projeta para o interior celular apresenta estruturas em alfa-hélices devido a sua capacidade de se ancorar à membrana hidrofóbica e, portanto interagir com esses domínios, gerando uma estrutura espiralada, mais estável.

4.7. Validação dos modelos

O principal gráfico gerado pelo PROCHEK é o gráfico de Ramachandran que apresenta uma correlação entre os ângulos torcionais da cadeia principal (ϕ e ψ) para cada resíduo, e pode ser observado na Figura 19.

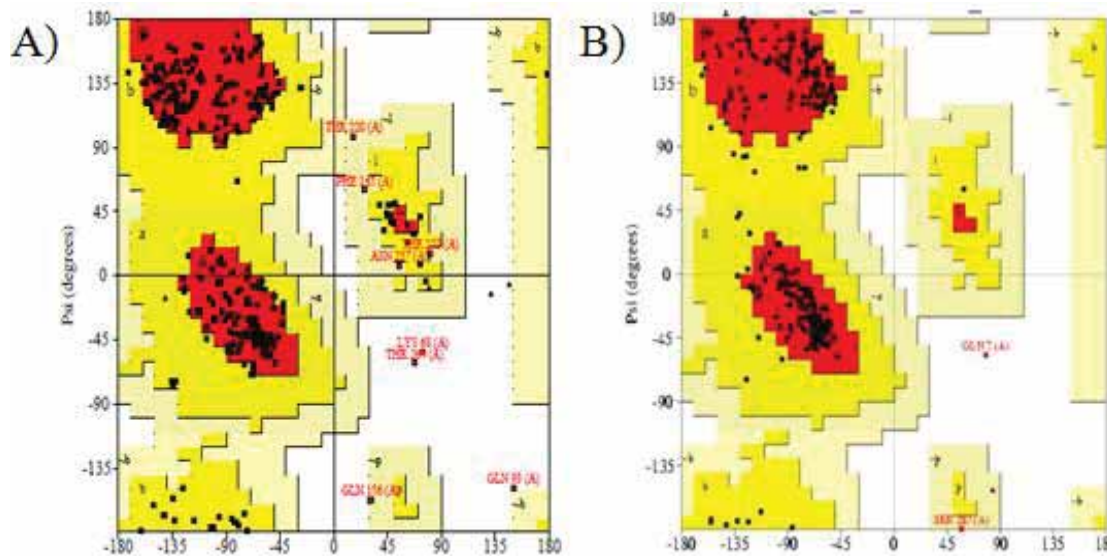


Figura 19. A) Gráfico de Ramachandran do modelo da proteína G do RSV gerado pelo Rosetta. B) Gráfico de Ramachandran do modelo da proteína G do RSV gerado pelo I-Tasser.

Em relação ao modelo gerado pelo Rosetta, a porcentagem de resíduos encontrados nas regiões favoráveis (vermelho) foi de 77,2%, nas regiões adicionalmente permitidas (amarelo escura) foi de 19,8%, nas regiões generosamente permitidas (amarelo claro) foi de 1,9% e nas regiões desfavoráveis (região em branco) foi de 1,1%.

Em relação ao modelo gerado pelo I Tasser, a porcentagem de resíduos encontrados nas regiões favoráveis (vermelho) foi de 86,6%, nas regiões adicionalmente permitidas (amarelo escura) foi de 12,7%, nas regiões generosamente permitidas (amarelo claro) foi de 0,4% e nas regiões desfavoráveis (região em branco) foi de 0,4%.

Os resíduos que se encontram em regiões desfavoráveis são um reflexo das estruturas usadas como molde, algumas são glicinas e alguns se encontram nas extremidades da proteína. Sendo assim, os gráficos garantem a qualidade dos ângulos torcionais da cadeia principal e revelam uma boa validação dos modelos.

4.8. Pontes dissulfeto

As pontes dissulfeto são formadas pela oxidação de dois grupos – SH, cada um dos quais estão presentes na cadeia lateral de um resíduo de cisteína. Para que ocorra a formação de uma ponte, a distância entre dois átomos de enxofre deve ser de 2,07 Å (STRYER, 1988).

Os resultados de análise das ligações dissulfeto podem ser visualizados abaixo, mostra que a sequência da proteína G possui 4 resíduos de cisteínas, porém não há ocorrência de pontes dissulfeto. A confiabilidade dessa análise é alta (8), o que permite dizer que a

probabilidade da ocorrência dessas pontes dissulfeto é praticamente nula. Esta afirmação pode ser feita por levar em consideração também a disposição dos resíduos de cisteína na seqüência primária da proteína, já que podemos observar que estas estão bem próximas umas das outras.

.....160.....170.....180.....190.....200.....210.....				
AA	P	N	N	D
	F	H	F	E
	V	F	N	F
	V	P	C	S
	I	C	S	N
	N	P	T	C
	W	A	I	C
	K	R	I	P
	N	K	K	P
	G	K	K	T
	T	T	T	K
	P	T	K	K
	P	T	F	K
	T	F	K	T
DB state	0	0	0	0
DB conf	8	8	8	8

Baseado no modelo construído para a proteína G, as distâncias entre resíduos de cisteínas foram medidas (Figura 20). Os modelos indicaram que as distâncias foram muito maiores que a distância entre dois átomos de enxofre, confirmando a não existência destas pontes dissulfeto na proteína G. Os resíduos C em rosa são as cisteínas.

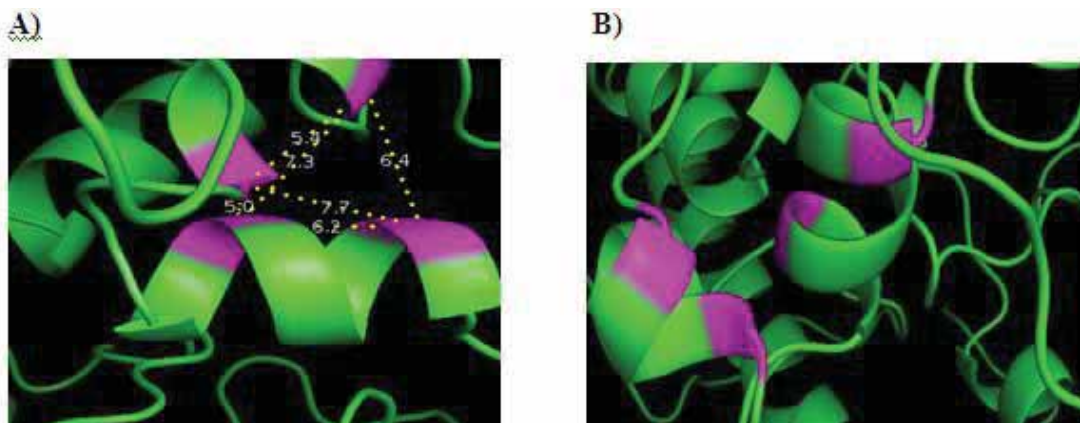


Figura 20. A) Localização dos resíduos de cisteínas, mostrados em cor rosa nas posições 173, 176, 182 e 186 do modelo da proteína G do hRSV. B) Distâncias entre os resíduos de cisteínas do modelo da proteína G, mostrados em cor verde (todas superiores a 2,07 Å).

Segundo alguns estudos, os resíduos conservados de cisteínas presentes no ectodomínio da proteína G do RSV formam pontes dissulfetos em uma estrutura em *loop*. (GORMAN et al., 1997; JONHSON et al., 1987; LANGEDIJK et al., 1998). Os resíduos que se ligariam formando as pontes são: 173-186 e 176-182, e que esta região, juntamente com mais 9 aminoácidos conservados, seria interpretada como possível sitio de reconhecimento para ligação celular. No entanto, estudos recentes tem mostrado que esta região não é necessária para infecção *in vitro*, o que leva a acreditar que o papel desta região de 13 aminoácidos conservados ainda não está esclarecido (TENG et al, 2001; TENG et al, 2002).

Os estudos teóricos realizados no âmbito de verificar se houve formação de pontes dissulfeto na proteína G do RSV mostraram a ausência desta, fato que pode ser explicado por

considerar apenas a sequência protéica, sem a mesma estar inserida em um ambiente *in vivo*, e, portanto não considerar possíveis conformações protéicas que poderiam ser formadas em um ambiente real.

4.9. Modificações pós-traducionais

Com o objetivo de entender melhor quais os tipos de modificações que a proteína de adesão pode ter, foram analisados pelo programa PREDICTPROTEIN possíveis sítios de modificações pós traducionais.

Todos os tipos de modificações pós traducionais podem alterar diversos fatores da proteína: sua localização, estabilidade e conformação, sua capacidade de interação com membranas ou domínios hidrofóbicos de outras proteínas e até mesmo sua função (LEHNINGER et al., 1999).

O ectodomínio carboxila-terminal da glicoproteína G possui alto teor de serina e treonina, sendo possíveis sítios de ligação de glicídios por ligações glicosiladas do tipo “O” (o glicídio liga-se a átomos de oxigênio das cadeias desses aminoácidos) ou do tipo “N” (ligação entre glicídio e átomo de nitrogênio da cadeia lateral da asparagina) (WERTZ et al., 1995).

A infectividade do hRSV é sensível à remoção das ligações do tipo N e do tipo O dos oligossacarídeos por endoglicosidases, indicando a necessidade desses carboidratos para que a glicoproteína G desempenhe sua função. Alterações na glicosilação do domínio C-terminal da proteína diminui a reatividade com soros humanos, indicando que esta variação pode contribuir para a evasão do sistema imunológico humoral pelo hRSV (RAWLING e MELERO, 2007).

Para muitas outras glicoproteínas virais, já foi demonstrado que os N-glicanos são importantes componentes estruturais que influenciam os processos de enovelamento e transporte, assim como a atividade, estabilidade e propriedades imunológicas destas moléculas (ZIMMER et al., 2001). Acredita-se que pressões seletivas atuam no sentido de manter um número mínimo de locais de N-glicosilação (ESTEVES et al., 2002), e uma vez que o padrão de glicosilação pode ser importante na replicação e transmissibilidade do vírus, as regiões altamente glicosiladas constituem deste modo uma estratégia importante para escape do vírus ao sistema imunitário do hospedeiro (LOLE et al, 1999).

Os resultados obtidos pelo programa mostraram sítios de glicosilação nas posições: 103, 135, 237, 242, 250. O processo é denominado N-glicosilação, e as posições encontradas estão no ectodomínio e são conservadas nas seqüências selecionadas, o que sugere uma significativa importância na estrutura e função da proteína G. Entretanto, poucos

estudos já relataram qual o verdadeiro papel da N-glicosilação para a proteína G do hRSV. Acredita-se que são necessários para proteção da proteína contra o reconhecimento de anticorpos, manutenção de 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 prováveis sítios de fosforilação por Quinase dependente de AMPc (posições 196 e 204), Fosforilação por Quinase C (posições 9,130, 147, 199, 203, 207, 210 219, 227, 231, 238 e 260), Fosforilação por Caseína Quinase II (nas posições 211, 260, 268 e 281), N-Miristilação (na posição 98) e Amidação (na posição 194), que estão dispostos na Tabela 4.

Tabela 4. Sítios de modificações pós traducionais da Proteína G separados pelo tipo de enzima que executa o processo, segundo a análise do programa PREDICTPROTEIN.

Modificação	Tipo de Modificação	Posições
Glicosilação	N-Glicosilação	103, 135, 237, 251, 294
Miristilação	N-miristilação	98
Amidação	Sítio de Amidação	194
Fosforilação	Quinase dependente cAMP	196, 204
	Quinase C	9,130, 147, 199, 203, 207, 210 219, 227, 231, 238, 260
	Caseína Quinase II	211, 260, 268, 281

Os sítios fosforilados estão presentes majoritariamente no ectodomínio protéico. 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. A fosforilação resulta na alteração de conformação espacial protéica (estrutura terciária), e é realizado por proteínas quinases, que são importantes para transdução do sinal. A grande quantidade destes sítios indicada pelo programa pode direcionar um estudo específico que vise à compreensão do real papel das modificações conformacionais da proteína G no processo de infecção viral.

Em relação ao processo de miristilação, um grupo N-miristoil se liga geralmente à uma glicina (G) N-terminal. As proteínas integrais de membrana são mantidas na mesma por fortes interações com os domínios hidrofóbicos (aminoácidos apolares) e as cadeias laterais dos lipídios. Assim, algumas proteínas contêm um ou mais lipídios de vários tipos covalentemente ligados, fornecendo uma âncora hidrofóbica na membrana plasmática e direcionando a proteína para sua localização correta (LEHNINGER et al., 1999). A miristilação é essencial para a associação da proteína com a membrana celular, bem como para a realização de quase todas as suas funções biológicas (GEYER et al., 2001).

O grau de amidação indica a porcentagem de grupos carboxilas na forma amida, e influencia fortemente algumas propriedades funcionais, tais como: solubilidade, capacidade de gelificação, entre outras. Este tipo de modificação confere maior quantidade de cargas positivas à molécula, e pode estabilizar regiões helicoidais (PUKALA et al., 2006).

4.10. Ponto isoelétrico

O Ponto isoelétrico teórico da proteína G foi obtido utilizando os programas Compute PI (BJELLQVIST et al., 1993) e ProtParam (GASTEIGER et al., 2005). Os resultados obtidos podem ser visualizados na tabela abaixo.

Tabela 5. Pontos isoelétricos da Proteína G e seus domínios.

Programa usado	ProtParam	Compute PI
Proteína G inteira	9.90	9.90
Proteína G Endodomínio	9.88	9.88
Proteína G Ectodomínio	9.74	9.74

O resultado foi exatamente o mesmo em ambos os programas, o que gerou uma maior confiança em relação ao possível ponto isoelétrico real da proteína. Segundo VOJDANI et al., (1996); quando uma solução protéica está no seu ponto isoelétrico, ou seja, quando a proteína apresenta carga líquida nula num sistema aquoso, as interações proteína-proteína aumentam, pois as forças eletrostáticas moleculares são mínimas, conseqüentemente, menos água interage com as moléculas de proteína, condição favorável para que as moléculas de proteína se aproximem, agreguem e precipitem. Ou seja, quanto mais próximo o pH de uma solução protéica do seu ponto isoelétrico, mais baixa será sua solubilidade. Em pHs diferentes do ponto isoelétrico a solubilidade protéica aumenta, devido ao aparecimento de cargas positivas ou negativas em excesso sobre as cadeias de proteínas, que favorecem a interação carga-momento do dipolo da água (CARNEIRO, 1997).

Estes dados são importantes, uma vez que, com base neles, pode-se estimar a carga líquida da proteína em um determinado pH, e será muito útil para realização de estudos experimentais como no processo de purificação protéica e manutenção da mesma em uma solução que favoreça sua manutenção na forma solúvel. Nos próximos ensaios de purificação protéica serão realizados utilizando-se tampões de pH 7.0 até pH 8.0, nos quais provavelmente a proteína não irá precipitar.

4.11. Verificação dos possíveis sítios de interações

A busca de sítios de interação na proteína G foi feita utilizando-se o programa PREDICTPROTEIN. Os resultados (Figura 21) mostraram apenas 1 sítio de interação da proteína G com outra proteína, e nenhum sítio de interação com DNA foi encontrado

```

MSKTKDQRTAKTLERTWDTLNHLLFFISSCLYKLNLSIAQITLSILAMIISTSLIIAAIIFIASANHKVTL
-----
TTAIMQDAT NQIKNTTPTYLTQNPQLGISFSNLSETTSQPTTILASTTPSAESTPQSTTVKTKNTTTTQIQPSK
-----
TTKQRQNKPNKPNNDHFHFEVFNFPVPCSI CSNNPTCWAICKRIPNKKPGKTTTKPTKKPTIKTTKKDLKPQTTK
-----P-----
PKEVLTTKPTKPTINTTKTNIRTTLLTSNTTGNPEHTSQKETLHSTTSEGNPSPSQVYTTSEYLSWSLSPSNTT
-----

```

Figura 21. Esquema das interações da proteína G do hRSV tipo A. A letra P indica as regiões de interação proteína-proteína obtidas pelo programa PredictProtein.

Como o hRSV é um vírus que não tem acesso ao núcleo da célula hospedeira tal interação realmente não foi demonstrada até o momento. Todas as etapas do ciclo replicativo do vírus sincicial respiratório ocorrem no citoplasma da célula, e a síntese de seu envelope ocorre na membrana plasmática (MCINTOSH & CHANOCK, 1990). Com base nessas informações acredita-se que a proteína G do vírus provavelmente não interage com moléculas de DNA, já que durante todas as etapas da replicação viral (que ocorrem no citoplasma), não há contato com o material genético celular.

Observando a figura 32 verificou-se que a proteína G provavelmente não realiza interações com o DNA, mas talvez realize com outras proteínas. Somente um sítio foi indicado como provável de interação com outras proteínas, localizado na posição 157, correspondente ao aminoácido asparagina. A asparagina é um aminoácido não essencial que possui três potenciais sítios de coordenação podendo atuar como ligante mono, bi ou tridentado.

Este potencial sítio de interação proteína-proteína pode indicar interação com as demais proteínas do RSV e também podem estar evidenciando importantes regiões de interações com as proteínas celulares. Esta informação é de grande importância, pois pode fornecer o possível sítios de interação das proteínas não só com outras proteínas, mas também com moléculas de características semelhantes, que podem atuar nestes sítios. Após a obtenção do modelo da proteína G do hRSV, observou-se que esta região de interação está localizada externamente na superfície da proteína, no ectodomínio, o que pode reforçar ainda mais essa

região como possível sítio de interação. A figura 22 mostra claramente esta região de interação da proteína G, na região em azul.

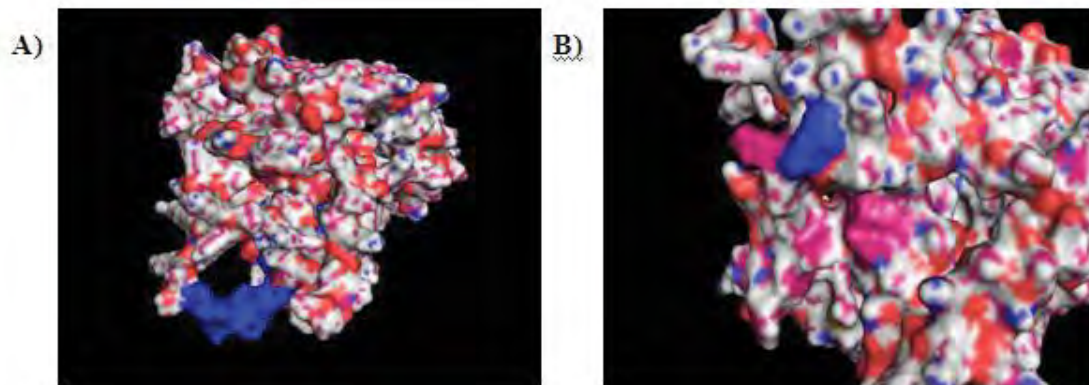


Figura 22. A) Modelo da região de interação da proteína G do hRSV. Região em azul representa a região de interação proteína-proteína da proteína G do tipo A, fornecida pelo programa PREDICTPROTEIN e localizada externamente no modelo. B) Aumento da região de interação da proteína G com outras proteínas.

A posição encontrada pela predição teórica (157), não está inserida no sítio conservado da proteína G do RSV (164-176), que se acredita ser o possível sítio de ligação com a célula hospedeira. No entanto, TENG e COLLINS, 2002 mostraram em um estudo com camundongos e cultura de células Vero, que a remoção do motif da proteína G tem pouca ou nenhuma interferência na infectividade ou replicação viral *in vitro* ou *in vivo*.

4.12. Verificação das interações G-heparina

Os modelos gerados para proteína G do hRSV foram submetidos ao Q-Site Finder, onde se buscou identificar sítios de interação. Na Figura 23, podem ser visualizados os modelos da proteína, sendo possível identificar os principais sítios protéicos.

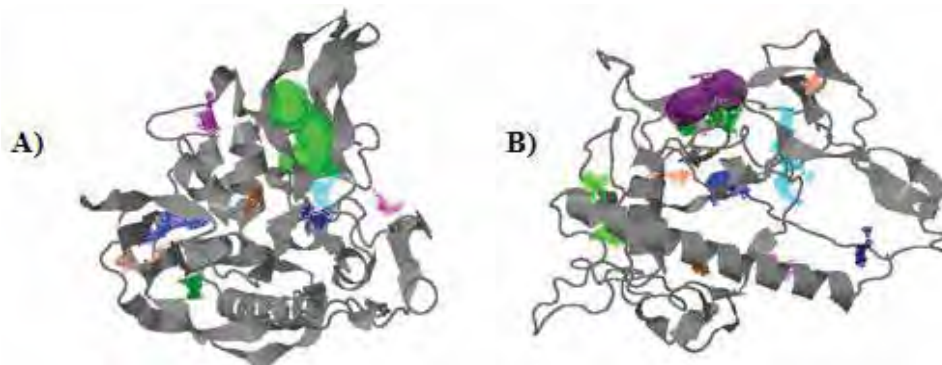


Figura 23. A) Resultado obtido com Q-SiteFinder para o modelo da proteína G gerado pelo Rosetta. B) Resultado obtido com Q-SiteFinder para o modelo da proteína G gerado pelo i-Tasser. Em colorido estão os sítios de interação obtidos pelo programa e em destaque (roxo) está o sítio descrito na literatura como ligante de heparina.

No total foram encontrados 10 prováveis sítios de interação da proteína G do RSV. Essa primeira análise foi preliminar, para, posteriormente, ser mais bem avaliada com estudos de Docking Molecular. Entretanto, os resultados encontrados foram satisfatórios, sendo que há evidências na literatura de alguns sítios indicados pelo programa são sítios de ligação para heparina, um possível inibidor da proteína G. Em ambos os modelos os sítios foram predominantemente indicados no ectodomínio, o que corrobora com as informações anteriores de que a região externa da proteína G tem um grande número de sítios de possível interação com diversos compostos.

O programa *PatchDock* forneceu vários modelos de interação da proteína G com a heparina (Figura 24). O modelo de interação entre proteína G (gerado pelo Rosetta) e heparina apresentou *score* 9972 enquanto que o modelo de interação da proteína G (gerado pelo I-Tasser) e heparina apresentou *score* 10720. Os valores de *score* obtidos foram considerados bons, uma vez que estudos de *docking* utilizando o mesmo programa, no intuito de verificar interação de uma proteína transmembrana (NA) do vírus *Influenza* com um inibidor utilizado experimentalmente resultou em *score* 4240 (SHARMA, 2010).

A) MSKTKDQRTAKTLERTWDTLNLHLLFISSCLYKLNLSIAQITLSILAMIISTSLIIAAIIFIASANH
 KVTLLTAIMQDATNQIKNTTPTYLQNPQLGISFSNLSETTSQPTTILASTTPSAESTPQSTTVKT
 KNTTTTQIQPSKPTTKQRQNKPNNDHFHFEVFNFPVPCISCSNNPTCWAIICKRIPNKKPGK
 KTTTKPTKKPTIKTTKKDLKPQTTKPKVLTTPKTEKPTINTTKTNIRTTLLTSNTTGNPEHTS
 QKETLHSTTSEGNPSPSQVYTTSEYLSQSLSPSNTTN

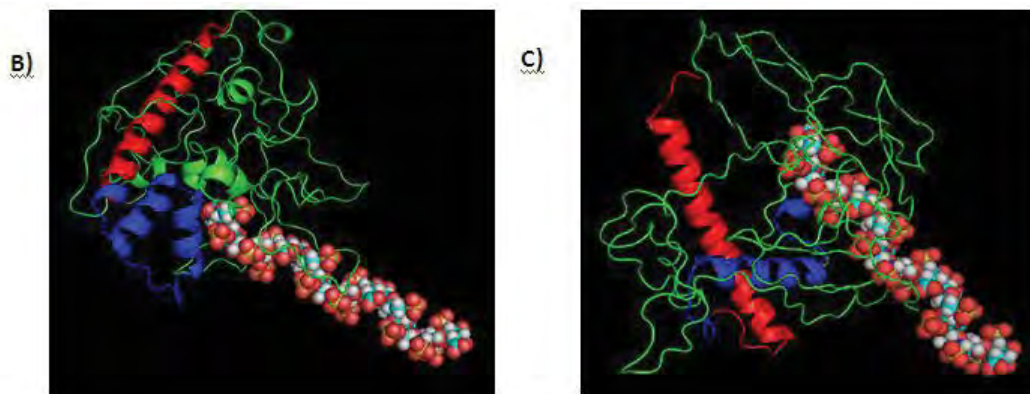


Figura 24. Esquema da interação da proteína G com heparina. A) Sequência de aminoácidos da proteína G do tipo viral A, utilizada para realização da modelagem. Em destaque estão as regiões de interação proteína-proteína da proteína G. B) Docking entre o modelo da proteína G (programa Rosetta) e a heparina, mostrando que a interação ocorre no mesmo sítio mostrado em A. C) Docking entre o modelo da proteína G (programa I-Tasser) e a heparina, mostrando que a interação também ocorre no mesmo sítio mostrado em A. O endodomínio está mostrado em vermelho, a região transmembrana em azul, o ectodomínio em verde e a estrutura da heparina está em forma de esferas.

O programa *AutoDock* forneceu 20 modelos de interação da proteína G com a heparina (Figura 25). O modelo de interação entre proteína G (gerada pelo Rosetta) e heparina apresentou energia livre de ligação de $-7,5\text{kcal/mol}$ enquanto que o modelo de interação entre proteína G (gerada pelo I-Tasser) e heparina apresentou energia livre de ligação de $-6,9\text{kcal/mol}$. Este valor de energia foi considerado bom, uma vez que quanto mais negativo é o valor obtido maior a estabilidade do complexo proteína-ligante.

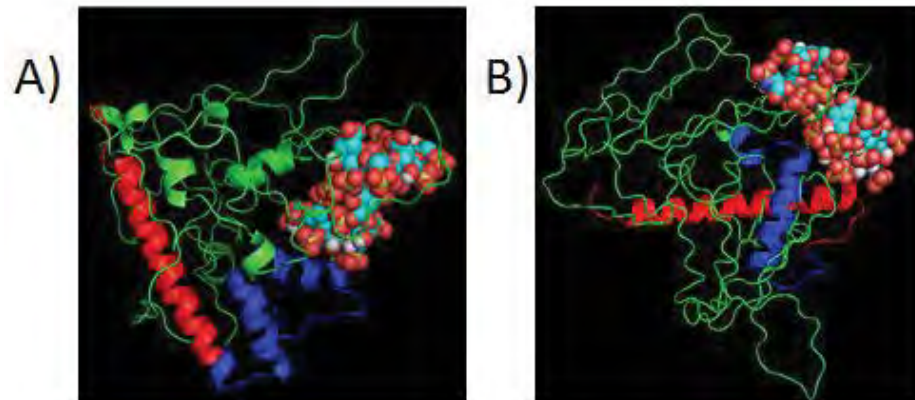


Figura 25. Esquema da interação da proteína G com heparina. A) Docking entre o modelo da proteína G gerado pelo Rosetta e a heparina B) Docking entre o modelo da proteína G gerado pelo I-Tasser e a heparina. Em vermelho – endodomínio, em azul - região transmembrana, em verde – ectodomínio, Estrutura da heparina - em forma de esferas.

Estes resultados indicam que parte da interação com a heparina ocorreu em uma região de importância da proteína G, independentemente do programa utilizado para gerar o modelo da proteína G. Desta forma, sugere-se que a heparina, quando ligado à proteína G é um provável inibidor das interações entre esta proteína com outras proteínas, pois esta se acomoda conforme mostrado no modelo no mesmo sítio de interação protéica.

Segundo evidências experimentais (FELDMAN et al.1998), uma seqüência de aminoácidos localizada no ectodomínio (AICKRIPNKKPGKKT (184 – 198) da proteína G do hRSV do tipo A seria responsável por ligar-se à heparina. Essa seqüência está localizada no sítio de interação obtido pelo *docking* molecular, o que indica este sitio com potencial alvo para investigação experimental. Essa região está localizada no ectodomínio, em uma porção que é localizada externamente à proteína, e que tem estrutura em *coil*, o que favorece a interação protéica.

Devido ao importante papel desempenhado pelas interações eletrostáticas, procurou-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. Procurou-se também a ausência de aminoácidos como histidina,

carregada negativamente e que provocaria uma repulsão com moléculas ácidas como Heparina. O principal aminoácido encontrado nesta região foi a Lisina (K – em verde), e apenas 1 Arginina (A – em vermelho) foi encontrada, enquanto nenhuma histidina esteve presente (H – em azul). Este resultado é esperado para motivos consenso deste tipo, já que a concentração destes resíduos (A e K), forma centros de alta densidade de carga positiva, que interage eletrostaticamente com os grupos ácidos da Heparina (CARDIN & WEINTRAUB, 1989).

5. CONCLUSÕES

Os resultados obtidos ao longo deste trabalho, com uso de ferramentas de bioinformática aliadas a estudos experimentais, contribuem para melhor entendimento da estrutura da proteína G do RSV, e como esta proteína pode ser alvo de compostos inibidores, podendo então ser utilizada em ensaios posteriores para avaliar a infecção viral baseando-se no processo de interação protéica.

Os resultados permitidos permitiram concluir que:

- A região correspondente aos segmentos G1 e G2 da proteína G foram clonadas, o peptídeo G1 foi expresso na concentração de 0,1mM de IPTG a 37° C por 2 horas, em bactéria BL21- pLysS; em sua forma solúvel;
- Existem na proteína G sítios de modificações pós traducionais, presentes principalmente no ectodomínio protéico, os quatro resíduos de cisteínas localizados no ectodomínio não formam pontes dissulfeto; os modelos gerados por predição *ab initio* tiveram estrutura secundária semelhantes, mostrando uma proteína composta predominantemente por estruturas indefinidas (*random coils*), seguidas de alfa-hélices, foi detectado um sítio de interação para a heparina na proteína G, localizado no ectodomínio protéico.
- Existe um sítio de interação para a heparina na proteína G, que esta localizado no ectodomínio.

6. REFERÊNCIAS

- ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WATSON, J. D. **Molecular biology of the cell**. 3rd ed. New York: Garland Publishing. 1994.
- ANDERSON, L. J.; PARKER, R. A.; STRIKAS, R. L. Association between respiratory syncytial virus outbreaks and lower respiratory tract deaths of infants and young children. **Journal of Infectious Diseases**, Chicago, v. 161, n. 4, p. 640-646, Apr. 1990.
- ARBIZA, J. et al. Genetic characterization of respiratory syncytial viruses isolated from consecutive acute respiratory infections in a HIV infected child. **Journal of Clinical Virology**, Amsterdam, v. 35, p. 41-45, 2006.
- BALDI, P.; BRUNNAK, S. **Bioinformatics: the machine learning approach**. 2nd ed. Cambridge: MIT Press, 2001.
- BARROSO, R. C. et al. Evaluation of the protamine in the neutralization of heparin after cardiopulmonary bypass. **Revista Brasileira de Cirurgia Cardiovascular**, São Paulo. v. 17, n. 1, p. 54-60, jan./mar. 2002.
- BATONICK, M.; OOMENS, A. G. P.; WERTZ, G. W. Human respiratory syncytial virus glycoproteins are not required for apical targeting and release from polarized epithelial cells. **Journal of Virology**, Washington, v. 82, n. 17, p. 8664-8672, Sept. 2008.
- BJELLQVIST, B.; HUGHES, G. J.; PASQUALI, C.; PAQUET, N.; RAVIER, F. et al. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. **Electrophoresis**, Deerfield Beach, v. 14, n. 1, p. 1023-1031, 1993.
- BONFIM, C. M.; NOGUEIRA, M. L.; SIMAS, P. V. M.; GARDINASSI, L. G. A.; DURIGON, E. L.; RAHAL, P.; SOUZA, F. P. Frequent respiratory pathogens of respiratory tract infection in children attending daycare centers in Brazil. **Jornal de Pediatria**, Rio de Janeiro, v. 87, n. 5, p. 439-444, 2011.
- BOSSO, P. A. R. et al. Human respiratory syncytial virus detection in children admitted at a community hospital in Botucatu, SP. **Brazilian Journal of Microbiology**, São Paulo, v. 35, n. 4, p. 348-351, Oct./Dec. 2004.
- BOTOSO, V. F.; ZANOTTO, P. M. A.; UEDA, M. et al. Positive selection results in frequent reversible amino acid replacements in the G protein gene on human respiratory syncytial virus. **PLoS Pathogens**, San Francisco, v. 5, n. 1, p. faltam as páginas, 2009.
- BOURGEOIS, C.; BOUR, J. B.; LIDHOLT, K.; GAUTHRAY, C.; POTHIER, P. Heparin-like structures on respiratory syncytial virus are involved in its infectivity in vitro. **Journal of Virology**, Washington, v. 72, p. 7221-7227, Sept. 1998.
- BUENO, S. M., GONZÁLEZ, P. A.; PACHECO, R.; LEIVA, E. D.; CAUTIVO, K. M.; TOBAR, H. E.; MORA, J. E.; PRADO, C. E.; ZÚÑIGA, J. P.; JIMÉNEZ, J.; RIEDEL, C. A.; KALERGIS, A. M. Host immunity during RSV pathogenesis: review. **International Immunopharmacology**, Amsterdam, v. 8, n. 10, p. 1320-1329, Oct. 2008.

CANDEIAS, J. A. Isolation of the respiratory syncytial virus from children with acute respiratory infections. **Revista do Instituto de Medicina Tropical de São Paulo**, São Paulo v. 9, n. 1, :p. 27-30, jan/fev. 1967.

CANE, P. A.; MATTHEWS, D. A.; PRINGLE, C. R. Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. **Journal of General Virology**, London, v. 72, n. 9, p. 2091–2096, Sept. 1991.

CAPILA, I.; LINHARDT, R. J. Heparin–protein interactions. **Angewandte Chemie International**, Weinheim, v. 41, n. 3, p. 390–412, Feb. 2002.

CARDIN, A. D.; WETRAUB, H. J. R. Molecular modeling of protein-glycosaminoglycan interactions. **Arteriosclerosis**, Dallas, v. 9, n. 1, p. 21-32, Jan. 1989.

CARNEIRO, J. G. M. **Características funcionais de concentrados protéicos de soro de leite de cabras**. Ano da defesa. Total de folhas. Doutorado. (Área de Concentração)- Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas, Campinas, 1997.

CARROMEU, C. **Estudos de caracterização estrutural, localização intracelular e imunogenicidade da RNA polimerase do vírus respiratório sincicial humano** Ano da defesa. Total de folhas. ~~Tese de~~ Doutorado. (Área de Concentração)- Instituto de Ciências Biomédicas. São Paulo: Universidade de São Paulo; 2007.

CHANOCK, R.; FINBERG, L. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). II. epidemiologic aspects of infection in infants and young children. **American Journal of Hygiene**, Baltimore, v. 66, n. 3, p. 291-300, Nov. 1957b.

_____.; ROIZMAN, B.; MYERS, R. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. isolation, properties and characterization. **American Journal of Hygiene**, Baltimore, v. 66, n. 3, p. 281-290, Nov. 1957a.

CHENG, X.; PARK, H., ZHOU, H., JIN, H. Overexpression of the M2-2 protein of respiratory syncytial virus inhibits viral replication. **Journal of Virology**, Washington, v. 79, n. 22, p. 13943-13952, Nov. 2005.

CHIVIAN, D. et al.; Servidor Rosetta. **Proteins: Structure, Function, and Genetics**. New York, v. 53, p. 524–533, 2003.

_____.; CHANOCK R. M.; MURPHY, B. R. Respiratory syncytial virus. In: KNIPE, D. M.; HOWLEY, P. M.(Ed.). **Fields virology**. 4th ed. Philadelphia: Lippincott, Williams & Wilkins Raven, 2001. p. 1443-1486.

COLLINS, P. L.; BUKREYEV, A. , YANG, L.; FRICKE, J.; CHENG, L.; WARD, J. M. MURPHY, B. R. The secreted form of respirator-y syncytial vírus G glycoprotein helps the vírus evade antibody-mediated restriction of replication by acting as an antigen decoy and through effects on Fc receptor-bearing leukocytes. **Journal of Virology**, Washington, v. 82, , n. 24, p. 12191-12204, Dec. 2008.

_____.; CROWE J. R. J. E. Respiratory syncytial virus. In: KNIPE, D. M.; HOWLEY, P. M. **Fields virology**. 5th ed. Philadelphia: Lippincott Williams & Wilkins Raven. 2007. p.1602-1640.

COWTON, V. M.; MCGIVERN, D. R.; FEARNNS, R. Unravelling the complexities of respiratory syncytial virus RNA synthesis. **Journal of General Virology**, London, v. 87, n. 7, p. 1805-1821, July 2006.

DUHOVNY, D.; NUSSINOV, R.; WOLFSON, H. J. Efficient unbound docking of rigid molecules. In: GUSFIELD, et al.(Ed.), **Proceedings of the 2nd Workshop on Algorithms in Bioinformatics(WABI)** Rome: Springer Verlag ,~~Italy~~–2002. p. 185-200.(Lecture Notes in Computer Science, v. 2452) .

ESTEVES, A.; PARREIRA. R.; VENENNO. T. et al. Molecular epidemiology of HIV type 1 infection in Portugal: high prevalence of non-B subtypes. **AIDS Research and Human Retroviruses**, New York, v. 18, n. 5, p. 313-325, Mar. 2002.

FAÇANHA, M. C.; & PINHEIRO, A. C. Doenças respiratórias agudas em serviços de saúde entre 1996 e 2001, Fortaleza, CE. **Revista de Saúde Pública**, São Paulo, v. 38, n. 3, p. 346-350, jun. 2004.

FALSEY, A. R.; WASH, E. E. Respiratory syncytial virus infections in adults. **Clinical Microbiology Reviews**, Washington, v. 13, n. 3, p. 371-84, July 2000.

FEARNNS, R.; COLLINS, P. L. Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. **Journal of Virology**, Washington, v. 73, n. 7, p. 5852-5864, July 1999.

FELDMAN, S. A.; HENDRY, M .R.; BEELER, J. A. Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. **Journal of Virology**, Washington, v. 73, n. 8, p. 6610-6617, Aug. 1999.

GARCIA, O., MARTIN, M.; DOPAZO, ARBIZA, J.; FRABASILE, S.; RUSSI, J.; HORTAL, M; PEREZ-BREÑA, P.; MARTINEZ, I.; GARCIA-BARRENO, B. J. Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. **Journal of Virology**, Washington, v. 68, n. 9, p. 5448–5459, Sept. 1994.

GASTEIGER, C. et al. In: WALKER, J. M.(Ed.). **The proteomics protocols handbook**. Totowa: Humana Press, 2005. p. 571–607.

GEYER, M.; FACKLER, O. T.; PETERLIN, B. M. Structure-function relationships in HIV-1 Nef. **EMBO Reports**, Oxford, v. 2, n. 7, p. 580-585, July 2001.

GHILDYA, L. R.; MILLS, J.; MURRAY, M.; VARDAXIS, N.; MEANGER, J. Respiratory syncytial virus matrix protein associates with nucleocapsids in infected cells. **Journal of General Virology**, London, v. 83, n. 4, p. 753-757, Apr. 2002.

GORMAN, J. J.; FERGUSON, B. L.; SPEELMAN, D.; MILLS, J. Determination of the disulfide bond arrangement of human respiratory syncytial virus attachment (G) protein by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. **Protein Science**, New York, v. 6, n. 6, p. 1308-1315, Nov. 1997.

HALL, T. A. **BioEdit**: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Oxford: Oxford University Press, 1999. p. 95-98. (Nucleic Acids Symposium Series, n. 41)

JOHNSON, P. R.; SPRIGGS, M. K.; OLMSTED, R. A.; COLLINS, P. L. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 84, n. 16, p.5625–5629, Aug. 1987.

KIM, H. H.; LEE, M. H.; LEE, J. S. Eosinophil cationic protein and chemokines in nasopharyngeal secretions of infants with respiratory syncytial virus (RSV) bronchiolitis and non-RSV bronchiolitis. **Journal of Korean Journal of Medical Sciences**, Seoul, v. 22, n. 1, p. 37-42, Feb. 2007.

LANGEDIJK, J. P. M.; DE GROOT, B. L.; BERENDSEN, J. C.; VAN OIRSCHOT, J. T. Structural homology of the central conserved region of the attachment protein G of respiratory syncytial virus with the fourth subdomain of 55-kDa tumor necrosis factor receptor. **Virology**, New York, v. 243, n. 2, p. 293-302, Apr. 1998.

LASKOWSKI, R. A.; MACARTHUR, M. W.; MOSS D. S.; THORNTON, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, Copenhagen, v. 26, n. 2, p. 283-291, Apr. 1993.

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

LEMANSKE, R. F. Viral infections and asthma inception. **Journal of Allergy and Clinical Immunology**, Saint Louis, v. 114, n. 5, p. 1023-1026, Nov. 2004.

LEHNINGER, A. L.; NELSON, D. L.; COX, M. M. **Principles of biochemistry**. 2nd ed. São Paulo: Ed. Savier, 1999. 839 p.

LEVINE S.; KLAIBER-FRANCO, R.; PARADISO, P. R. Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. **Journal of General Virology**, London, v. 68, n. 9, p. 2521-2524, Sept. 1987.

LING, Z.; TRAN, K. C.; TENG, M. N. Human respiratory syncytial virus nonstructural protein NS2 antagonizes the activation of beta interferon transcription by interacting with RIG-I. **Journal of Virology**, Washington, v. 83, n. 08, p. 3734-3742, Apr. 2009.

LO, M. S.; BRAZAS, R. M.; HOLTZMAN, M. J. Respiratory syncytial virus nonstructural proteins NS1 and NS2 mediate inhibition of Stat2 expression and alpha/beta interferon responsiveness. **Journal of Virology**, Washington, v. 79, n. 14, p. 9315-9319, July 2005.

LOLE, K. S.; BOLLINGER R. C.; PARANJAPE, R. S.; et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. **Journal of Virology**, Washington, v. 73, n. 1, p. 152-160, Jan. 1999.

MANOHA, C.; ESPINOSA, S.; AHO, S. L.; HUET, F.; POTHIER, P. Epidemiological and clinical features of HMPV, RSV and RVs infections in young children. **Journal of Clinical Virology**, 38:221-6. 2007.

MASON, S. W.; LAWETZ, C.; GAUDETTE, Y.; DÔ, F.; SCOUTEN, E.; LAGACÉ, L.; SIMONEAU, B.; LIUZZI, M. Polyadenylation-dependent screening assay for respiratory syncytial virus RNA transcriptase activity and identification of an inhibitor. **Nucleic Acids Research**, Oxford, v. 32, n. 16, p. 4758-4767, Sept.; 2004.

MCINTOSH, K.; & CHANOCK, R. M. Respiratory syncytial virus, ~~p. 1045-1072~~. In. FIELDS, B. N. (Ed.), **Virology**. New York: Raven Press, 1990. p. 1045-1072.

MEISSNER, H. C. et al. Revised indications for the use of palivizumab and respiratory syncytial virus immune globulin intravenous for the prevention of respiratory syncytial virus infections. **Pediatrics**, v. 112, p. 1447-1452, Dec. 2003.

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**, Washington, [S-1], v. 78, p. 3524-3532, Apr. 2004.

MORRIS, J. A.; BLOUNT, R. E.; SAVAGE, R. E. Recovery of cytopathogenic agent from chimpanzees with coryza. **Proceedings of the Society for Experimental Biology and Medicine**, New York, v. 92, n. 3, p. 544-549, July 1956.

MORRIS, S. K.; DZOLGANOVSKI, B.; BEYENE, J. et al. A meta-analysis of the effect of antibody therapy for the prevention of severe respiratory syncytial virus infection. **BMC Infectious Diseases**, Rockville Pike, v. 9, n. 106, July 2009. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2720977/pdf/1471-2334-9-106.pdf>>. Acesso em: 25 jul. 2011.

MOUNT, D. W. **Bioinformatics: sequences and genome analysis**. 2nd ed. New York: Cold Spring Harbor, 2004.

NAKAMURA, M.; ITOKAZU, K.; TAIRA, K. et al. Genotypic and phylogenetic of the G gene of respiratory syncytial virus isolates in Okinawa, Japan, 2008. **Japan Journal of Infectious Diseases**, Tokyo, v. 62, n. 4, p. 326-327, 2009.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. Disponível em: <http://www.ncbi.nlm.nih.gov/>. Acesso em: 10 jan. 2010.

NOKES, J. D.; CANE, P. A. New strategies for control of respiratory syncytial virus infection: review. **Current Opinion of Infectious Diseases**, Philadelphia, v. 21, n. 6, p. 639-643, Dec. 2008.

OHUCHI, M.; OHUCHI, R.; FELDMANN, A.; KLENK, H. D. Regulation of receptor binding affinity of influenza virus hemagglutinin by its carbohydrate moiety. **Journal of Virology**, Washington, v. 71, n. 11, p. 8377–8384, Nov. 1997.

OLSZEWSKA, W.; OPENSHAW, P. Emerging drugs for respiratory syncytial virus infection. **Expert Opinion on Emerging Drugs**, London, v. 14, n. 2, p.207-217, June 2009.

ÖLUND, M. R.; LINDELL, A. T. ; STENLER, S.; RIEDEL, H. M; WIRGART, B. Z.; GRILLNER. L. Molecular epidemiology and genetic variability of respiratory syncytial virus (RSV) in Stockholm, 2002–2003. **Journal of Medical Virology**, New York, v. 80, n. 1, p. 159–167, Jan. 2008.

PARVEEN, S. et al. Genetic variability in the G protein gene of group A and B respiratory syncytial viruses from India. **Journal of Clinical Microbiology**, Washington, v. 44, n. 9, p. 3055-3064, Sept. 2006.

PAVLOVA, S.; HADZHILOVA, T.; ABADJIEVA, P. et al. Application of RT-PCR for diagnosis of respiratory syncytial virus and human metapneumovirus infection in Bulgaria, 2006-7 and 2007-8. **Eurosurveillance**, Stockholm, v. 14, n. 23, June 2009. Disponível em: <<http://www.eurosurveillance.org/images/dynamic/EE/V14N23/art19233.pdf>>. Acesso em: 15 ago. 2011.

PERET, T. C.; HALL, C. B.; SCHNABEL, K. C.; GOLUB, J. A.;-ANDERSON, L. J. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. **Journal of General Virology**, London, v. 79, n. 9, p. 2221–2229, Sept. 1998.

PITREZ, P. M.; STEIN, R.; STUERMER, L.; MACEDO, I.S.; SCHMITT, V. M.; JONES M. H.; et al. Bronquiolite aguda por rinovírus em lactentes jovens. **Jornal de Pediatria**, Rio de Janeiro, v. 81, n. 5, p. 417-420, out. 2005.

POLLARD, T.D.; EARNSHAW. W. Cell Biology. Philadelphia: Saunders, 2007.

PUKALA, T. L.; BERTOZZI, T.; DONNELLAN, S. C.; BOWIE, J. H.; SURINYA-JOHNSON, K. H.; LIU, Y.; JACKWAY, R. J.; DOYLE, J. R.; LLEWELLYN, L. E.; TYLER, M. J. Host-defense peptide profiles of the skin secretions of interspecific hybrid tree frogs and their parents, female *Litoria splendida* and male *Litoria caerulea*. **The FEBS Journal**, Oxford, v. 273, n. 15, p. 3511-3519, Aug. 2006.

RAMACHANDRAN, G. N.; SASISEKHARAN, V. Conformation of polypeptides and proteins. **Advances in Protein Chemistry**, New York, v. 23, p. 283-487, 1968.

RAMAN, R.; RAGURAM, S.; VENKATARAMAN, G.; PAULSON, J.C.; SASISEKHARAN, R. Glycomics: an integrated systems approach to structurefunction relationships of glycans. **Nature Methods**, New York, v. 2, n. 11, p. 817–824, Nov. 2005.

RAWLING, J.; MELERO, J. A. The use of monoclonal antibodies and lectins to identify changes in viral glycoproteins that are influenced by glycosylation: the case of human

respiratory syncytial virus attachment (G) glycoprotein: review. **Methods in Molecular Biology**, New York, v. 379, p. 109-125, 2007.

RICCETTO, A. G. L. et al. Respiratory syncytial virus (RSV) in infants hospitalized for acute lower respiratory tract disease: incidence and associated risks. **The Brazilian Journal of Infectious Diseases**, Salvador, v. 10, n. 5, p. 357-361, out. 2006.

RIXON, H. W.; BROWN, G.; AITKEN, J.; MCDONALD, T.; GRAHAM, S.; SUGRUE, R. J. The small hydrophobic (SH) protein accumulates within lipid-raft structures of the Golgi complex during respiratory syncytial virus infection. **Journal of General Virology**, London, v. 85, n. 5, p. 1153-1165, May 2004.

ROST, B.; YACHDAV, G.; LIU, J. The predict protein server. **Nucleic Acids Research**, Oxford p. W321-326, July 2004. Disponível em: http://nar.oxfordjournals.org/content/32/suppl_2/W321.full.pdf+html>. Acesso em: 21 mai. 2011.

RUBIO, M. L.; TOLEDO, K. A.; CALÇA, J.; BONFIM, C. M.; GOMES, D. E.; DURIGON, E. L.; FOSSEY, M. A.; SOUZA, F. P. Heparin and dextran sulfate: their role on RSV infectivity. **Virus: Reviews and Research**, São Paulo, v. 15, n. 2, p. 1-16, ~~X-z~~, 2010. Disponível em: < <http://www.virusreviewsandresearch.com/volume15n2pdf/93.pdf>>. Acesso em: 26 nov. 2011.

SANTOS, F. C.; ORTEGA, J. M. Bioinformática aplicada à genômica. In: BOREM, A. ; GIÚDICE, M. del.; SEDYAMA, T. **Melhoramento genômico**. Viçosa, MG UFV, 2003.

SCHWEDE, T.; PEITSCH, M. **Computational structural biology: methods and applications**. 4^{ed.} Singapore: World Scientific Publishing, 2008.

SHARMA, A.; TENDULKAR, A. V.; WANGIKAR, P. P. Drug discovery against H1N1 virus (influenza A virus) via computational virtual screening approach. **Medicinal Chemistry Research**, v. 20, n. 9, p. 1445-1449, 2011.

SHORTLE, D.; SIMONS, K. T.; BAKER, D.-Clustering of low-energy conformations near the native structures of small proteins. **Proceedings of the National Academy of Sciences of the United States of America**, Washington , v. 95, n. 19, p. 11158-11162, Sept. 1998.

SIEW, N.; ELOFSSON, A.; RYCHLEWSKI, L.; FISCHER, D. MaxSub: an automated measure for the assessment of protein structure prediction quality. **Bioinformatics**, Oxford, v. 16, n. 9, p. 776-785, Sept. 2000.

SILVA, V. B. **Estudos de modelagem molecular e relação estrutura atividade de oncoproteína hnRNP K e ligantes**. Ano da defesa. Total de folhas. Dissertação (Mestrado em Física Biológica). Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, São Paulo, 2008.

STRYER, L. **Biochemistry**. 3rd. ed. New York: W. H. Freeman, 1988.

SULLENDER, W. M. Respiratory syncytial virus genetic and antigenic diversity. **Clinical Microbiology Review**, Washington, v. 13, n. 1, p. 1-15, Jan. 2000.

SWEDAN, S.; MUSIYENKO, A.; BARIK, S. Respiratory syncytial virus nonstructural proteins decrease multiple members of the cellular interferon pathways. **Journal of Virology**, Washington, v. 83, n. 19, p. 9682-9693, Oct. 2009.

TANG, R. S.; NGUYEN, N.; CHENG, X.; JIN, H. Requirement of cysteines and length of the human respiratory syncytial virus M2-1 protein for protein function and virus viability. **Journal of Virology**, Washington, v. 75, n. 23, p. 11328-11335, Dec. 2001.

_____.; WHITEHEAD, S. S.; COLLINS, P. L. Contribution of the respiratory syncytial virus G glycoprotein and its secreted and membrane-bound forms to virus replication in vitro and in vivo. **Virology**, New York, v. 289, n. 2, p. 283-296, Oct. 2001.

TENG, M. N.; COLLINS, P. L. The central conserved cystine noose of the attachment G protein of human respiratory syncytial virus is not required for efficient viral infection in vitro or in vivo. **Journal of Virology**, Washington, v. 76, n. 12, p. 6164-6171, June 2002.

THOMAZELLI, L. M. et al. Vigilância de 8 vírus respiratórios em amostras clínicas de pacientes pediátricos no sudeste do Brasil. **Jornal de Pediatria**, Rio de Janeiro, v. 83, n. 5, p. 422-428, out. 2007.

TRAN, T. L.; CASTAGNÉ, N.; BHELLA, D.; VARELA, P. F.; BERNARD, J.; CHILMONCZYK, S.; et al. The nine C-terminal amino acids of the respiratory syncytial virus protein P are necessary and sufficient for binding to ribonucleoprotein complexes in which six ribonucleotides are contacted per N protein protomer. **Journal of General Virology**, London, 2007 Jan; v. 88, n. 1, p.196-206, Jan. 2007.

TRIFONOV, E. N. Earliest pages in bioinformatics., **Bioinformatics**, Oxford, v. 16, n. 1, p. 5-9, Jan. 2000.

VIEIRA, S. E.; GILIO, A. E.; MIYAO, C. R.; PAHL, M. M. C.; LOTUFO, J. P. B.; HEIN, N.; et al. Infecção nosocomial pelo vírus respiratório sincicial em enfermaria de pediatria. **Pediatria**, São Paulo, v. 24 n. 1/2, p. 17-24, 2002. Disponível em: <<http://www.pediatriasaopaulo.usp.br/upload/pdf/549.pdf>>. Acesso em: 11 set. 2011.

VOET, D.; VOET, J. G. **Biochemistry**. 2nd ed. New York: John Wiley, Sons, INC, 1995.

_____.;_____.; PRATT, C.; **Fundamentos de bioquímica**. Porto Alegre: Editora ARTMED, 2000.

VOJDANI, F. Solubility. In: HALL, G. M. **Methods of testing protein functionality**.; London: Blackie Academic & Professional, 1996. chapt. 2, p. 11-60.;

WALSH, E. E.; HRUSKA, J. Monoclonal antibodies to respiratory syncytial virus proteins: identification of the fusion protein. **Journal of Virology**, Washington, v. 47, n. 1, p. 171-177, July 1983.

WERTZ, G. W.; COLLINS, P. L.; HUANG, Y.; GRUBER, C.; LEVINE, S.; BALL, L. A. Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an

unusual type of viral membrane protein. **Proceedings of the National Academy of Sciences of the United States of America**, v. 82, n. 12, p. 4075-4079, June 1985.

WHEELER, D. L.; BARRETT, T.; BENSON, D. A.; BRYANT, S. H.; CANESE, K.; CHETVERNIN, V.; et al.; Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**, v. 35, p. 5-12, 2007.

WORLD HEALTH ORGANIZATION. 2008. World Health statistics 2008: [updated August 2008]. Disponível em: www.who.int/infectious-disease-report . Acesso em: 20 out. 2011.

WU, H. et al. Development of motavizumab, an ultra-potent antibody for the prevention of respiratory syncytial virus infection in the upper and lower respiratory tract. **Journal of Molecular Biology**, London, v. 368, n. 3, p. 652-665, May 2007.

ZIMMER M.; DANKO, J. P.; PENNINGS, S. C.; DANFORD, A. R.; ZIEGLER, A.; UGLOW, R. F.; CAREFOOT, T. H. Hepatopancreatic endosymbionts in intertidal isopods (Crustacea: Isopoda), and their contribution to digestion. **Marine Biology**,. New York, v. 138, n. 5,-p.-955–963,- May 2001.

ZLATEVA, K.T. et al. Subgroup prevalence and genotype circulation patterns of human respiratory syncytial virus in belgium during ten successive epidemic seasons. **Journal of Clinical Microbiology**, Washington , v. 45, n. 09, p. 3022-3030, Sept. 2007.