



**UNIVERSIDADE FEDERAL DO OESTE DO PARÁ  
INSTITUTO DE BIODIVERSIDADE E FLORESTAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOCÊNCIAS**

**JOÃO MARCOS PEREIRA GALÚCIO**

**PLANEJAMENTO DE PROTEÍNAS IMUNOGÊNICAS MULTI-EPÍTOPO  
VISANDO O DESENVOLVIMENTO DE UMA VACINA DE NOVA GERAÇÃO  
PARA A INFECÇÃO DO VÍRUS NIPAH**

**SANTARÉM-PA  
2020**

**JOÃO MARCOS PEREIRA GALÚCIO**

**PLANEJAMENTO DE PROTEÍNAS IMUNOGÊNICAS MULTI-EPÍTOPO  
VISANDO O DESENVOLVIMENTO DE UMA VACINA DE NOVA GERAÇÃO  
PARA A INFECÇÃO DO VÍRUS NIPAH**

Dissertação apresentada ao Programa de Pós-Graduação  
em Biociências da Universidade Federal do Oeste do Pará,  
como requisito para obtenção do título de Mestre.  
Orientador: Prof. Dr. Paulo Sérgio Taube Júnior

**SANTARÉM-PA  
2020**

**Dados Internacionais de Catalogação-na-Publicação (CIP)**  
**Sistema Integrado de Bibliotecas – SIBI/UFOPA**

---

G181p Galúcio, João Marco Pereira  
Planejamento de proteínas imunogênicas multi-epítipo visando o desenvolvimento de uma vacina de nova geração para a infecção do vírus Nipah. / João Marco Pereira Galúcio – Santarém, 2020.  
107 p. : il.  
Inclui bibliografias.

Orientadora: Paulo Sérgio Taube Júnior  
Dissertação (Mestrado) – Universidade Federal do Oeste do Pará, Pró-reitoria de Pesquisa, Pós Graduação e Inovação Tecnológica, Instituto de Biodiversidade e Floresta, Programa de Pós-Graduação em Biociências.

1. Henipavírus. 2. Imunoinformática. 3. Vacina multi-epítipo. 4. Modelagem molecular. I. Taube Junior, Paulo Sérgio, *orient.* II. Título.

CDD: 23 ed. 615.372

**JOÃO MARCOS PEREIRA GALÚCIO**

**PLANEJAMENTO DE PROTEÍNAS IMUNOGÊNICAS MULTI-EPÍTOPO  
VISANDO O DESENVOLVIMENTO DE UMA VACINA DE NOVA GERAÇÃO  
PARA A INFECÇÃO DO VÍRUS NIPAH**

Dissertação apresentada ao Programa de Pós-Graduação  
em Biociências da Universidade Federal do Oeste do Pará,  
como requisito para obtenção do título de Mestre.  
Orientador: Prof. Dr. Paulo Sérgio Taube Júnior

Conceito: Aprovado

Data de aprovação: 19/11/2020



Prof. Dr. Paulo Sérgio Taube Júnior  
Orientador – Universidade Federal do Oeste do Pará



Prof. Dr. Élcio de Souza Leal  
Avaliador – Universidade Federal do Pará



Dr. Rodolpho de Campos Braga  
Avaliador – InsilicAll Company

Aos meus familiares, pelo carinho, apoio e incentivo.

## **AGRADECIMENTOS**

A Deus, ato puro de existir, pela existência e pela graça que nos alcança em todas as circunstâncias e fases da vida.

Aos meus pais, pelo amor incondicional que dedicam a mim e por todo o incentivo ao longo da trajetória acadêmica.

Ao professor Paulo Taube, por ter sido um orientador presente e flexível, além de exemplo de pessoa e de profissional.

Ao professor Kauê Santana, pelo suporte e por tantos ensinamentos repassados desde a graduação, além das oportunidades que abriu para minha trajetória acadêmica.

Ao Laboratório de Informática II do Instituto de Ciências da Educação, pelo espaço computacional cedido para que pudéssemos realizar os cálculos necessários a este trabalho.

A todos que direta ou indiretamente colaboraram para que este trabalho se tornasse possível, meus mais sinceros agradecimentos.

“Estou fascinado com a ideia de que a genética é digital. Um gene é uma longa sequência de letras codificadas, como informações do computador. A biologia moderna está se tornando muito mais um ramo da tecnologia da informação.”  
(Richard Dawkins)

## RESUMO

O Nipah (NiV) é um vírus zoonótico emergente pertencente à família Paramyxoviridae, envolvido em doenças respiratórias e neurológicas fatais em humanos. Recentemente, a Organização Mundial da Saúde (OMS) declarou o NiV como um patógeno prioritário para pesquisa e desenvolvimento de estratégias de diagnóstico, prevenção e tratamento. Devido à falta de medicamentos antivirais eficazes e a transmissibilidade em múltiplos hospedeiros, NiV representa um grande desafio para a saúde pública com o risco potencial de proliferação pandêmica ou aplicação como agente de bioterrorismo. No presente estudo, os epítomos virais foram preditos e selecionados de acordo com parâmetros físico-químicos, seletividade e afinidade a diferentes variantes do complexo principal de histocompatibilidade (MHC) classes I e II. Em seguida, esses peptídeos foram conectados usando ligantes de resíduos.  $\beta$ -defensinas também foram adicionadas (como adjuvantes) aos modelos estruturais para aumentar a imunogenicidade. A antigenicidade, imunogenicidade, alergenicidade, bem como as propriedades físico-químicas das estruturas de proteínas multi-epítopo projetadas também foram avaliadas usando métodos computacionais. Simulações de docagem molecular e de dinâmica molecular foram realizadas entre os modelos projetados com quatro receptores *toll-like* humanos, para explorar o modo de ligação e a estabilidade do complexo antígeno-TLR. Também foram analisadas as interações entre cada epítopo predito e estruturas de MHC-I e MHC-II, usando modelagem molecular. Finalmente, a adaptação de códons das sequências de cDNA projetadas e análises de expressão *in silico* usando sistemas bacterianos (*Escherichia coli*) foram realizadas para permitir um melhor desempenho para a expressão heteróloga das proteínas imunogênicas. Os modelos de proteínas propostos se mostraram como potencialmente antigênicos e não alergênicos; e contêm epítomos imunodominantes provenientes de todas as proteínas virais antigênicas. Ambos os modelos moleculares demonstraram ter afinidade e seletividade satisfatórias com o TLR3. Os epítomos conservados e não tóxicos selecionados mostraram um alto potencial para formar interação molecular estável com várias moléculas de MHC que cobrem mais de 98,0% da população humana em todo o mundo. Além disso, a análise das sequências de cDNA de ambos os modelos foi predita com expressão adequada em linhagens hospedeiras bacterianas, o que poderia facilitar ainda mais sua expressão heteróloga. Em conjunto, as análises computacionais forneceram *insights* sobre proteínas multi-epítomos biologicamente viáveis para vacinas candidatas contra o NiV que fornecem uma compreensão importante sobre a imunogenicidade das proteínas virais, constituindo-se como modelos com propriedades otimizadas para o desenvolvimento de vacinas de próxima geração.

**Palavras-chave:** Henipavírus. Imunoinformática. Vacinas multi-epítopo. Modelagem Molecular.



## ABSTRACT

Nipah (NiV) is an emerging zoonotic virus belonging to the Paramyxoviridae family, involved with fatal respiratory and neurological diseases in humans. Recently, the World Health Organization (WHO) declared NiV as a priority pathogen for research and development of diagnostic, prevention, and treatment strategies. Due to the lack of effective antiviral drugs and its potential transmissibility in humans, NiV represents a great public health challenge with the potential risk of pandemic proliferation or application as bioterrorism agent. In the present study, using immunoinformatics and molecular modeling approaches, we designed multi-epitope proteins recognized by T and B cell lymphocytes using epitopes conserved among known strains of NiV. The predicted epitopes were selected according to physicochemical parameters, selectivity, and affinity binding to major histocompatibility complex (MHC) classes I and II. Then, these polypeptide chains were bonded using residues linkers.  $\beta$ -defensin adjuvants were also added to the structural models to increase immunogenicity. The antigenicity, immunogenicity, allergenicity, as well as, the physicochemical properties of the designed multi-epitope proteins structures were also evaluated using computational methods. Molecular docking and molecular dynamics simulations were performed between models designed with four human toll-like receptors, to explore the binding mode and stability of the TLR antigen complex. The interactions between each predicted epitope and MHC-I and MHC-II structures were also analyzed, using molecular modeling. Finally, the codon adaptation of the projected cDNA sequences and *in silico* expression analyses using bacterial systems (*Escherichia coli*) were performed to allow a better performance for the heterologous expression of the immunogenic proteins. The proposed protein models were shown to be potentially antigenic and non-allergenic; and contain immunodominant epitopes from all antigenic viral proteins. Both molecular models have demonstrated satisfactory affinity and selectivity with TLR3. The selected conserved and non-toxic epitopes showed a high potential to form stable molecular interaction with various MHC molecules that cover more than 98.0% of the human population worldwide. In addition, the analysis of the cDNA sequences of both models was predicted with adequate expression in bacterial host lines, which could further facilitate their heterologous expression. Together, the computer analysis provided insights into biologically viable multi-epitope proteins for candidate NiV vaccines that provide an important understanding of the immunogenicity of viral proteins, constituting models with properties optimized for the development of next generation vaccines.

**Keywords:** Henipavirus. Immunoinformatics. Multi-epitope vaccines. Molecular modelling.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1-</b> Distribuição geográfica dos surtos henipavirais e morcegos da família Pteropodidae. ....	21
<b>Figura 2-</b> Representação dos principais constituintes da estrutura do Nipah vírus. ....	22
<b>Figura 3 -</b> Representação da patogênese do NiV. ....	25
<b>Figura 4 -</b> Pontos de acesso globais para doenças infecciosas emergentes que se originam na vida selvagem. ....	33
<b>Figura 5 -</b> Emergência de doença zoonótica pandêmica. ....	34
<b>Figura 6-</b> Linha do tempo representativa da imunidade inata e adaptativa. ....	38
<b>Figura 7 -</b> Principais vias de processamento de antígenos (MHC-I e MHC-II). ....	40
<b>Figura 8-</b> Metodologia para desenho de proteínas imunogênicas a partir do proteoma do Vírus Nipah. ....	60
<b>Figura 9 -</b> Visão geral esquemática dos dois tipos de proteínas multiepítipo projetadas. ....	65
<b>Figura 10 -</b> Estruturas secundária e terciária de proteína multi-epítipo projetada (Modelo 1). ....	68
<b>Figura 11 -</b> Estruturas secundária e terciária de proteína multi-epítipo projetada (Modelo 4). ....	69
<b>Figura 12 -</b> Estruturas das proteínas multi-epítipo projetadas mostrando a localização dos epítomos conformacionais reconhecidos pelos linfócitos B (em azul). A) Modelo 1. B) Modelo 4. ....	70
<b>Figura 13-</b> Interações intermoleculares obtidas por docagem molecular entre TLR3 e modelos da proteína multiepítipo. A) Modelo 1. B) Modelo 4. ....	73
<b>Figura 14 -</b> Mapas de potencial eletrostático. A) TLR3. B) Modelo 1. C) Modelo 2. ....	74
<b>Figura 15 -</b> Interações intermoleculares obtidas por docagem molecular entre moléculas do MHC e epítomos. preditos. A) MHC-I supertipo A24 complexado com epítipo EWISIVPNF. B) MHC-I supertipo A2 complexado com epítipo FLITRINWI. C) MHC-II molécula DRB5*01:01 complexada com epítipo MMASILLTLFRRTKK. D) MHC-II molécula DRB*07:01 em complexo com epítipo IILYVLSIASLCIGL. ....	76
<b>Figura 16 -</b> Gráfico de RMSD do receptor TLR3 no estado não ligado (livre de antígeno) e no estado ligado (complexado com os modelos multiepítipo) obtidos ao longo de 100 ns de simulação de DM. ....	77

<b>Figura 17-</b> Gráfico de RMSF do receptor TLR3 no estado não ligado (livre de antígeno) e no estado ligado (complexado com os modelos de multiepítopos) obtidos ao longo de 100 ns de simulação de DM.....	78
<b>Figura 18 -</b> Esquema de expressão <i>in silico</i> . A porção vermelha representa a inserção de proteínas de múltiplos epítopos no vetor de expressão pET28a (+) (preto). A) Modelo 1. B) Modelo 4.....	79

## LISTA DE TABELAS

<b>Tabela 1-</b> Proteínas codificadas pelo genoma NiV, identificadas por seu símbolo, código de acesso Uniprot, e suas funções moleculares envolvidas com o ciclo de vida do vírus. ....	18
<b>Tabela 2 -</b> Predição de antigenicidade das proteínas NiV usando alinhamento baseado em sequência. Cada proteína é identificada por seu respectivo índice de antigenicidade, cobertura (%), identidade (%) e homólogo humano previsto.....	61
<b>Tabela 3-</b> Epítomos virais de proteínas de NiV reconhecidas pelos linfócitos T citotóxicos e respectivos supertipos de MHC-I, <i>percentile rank</i> (PR), pontuação de imunogenicidade (IS), pontuação de processamento (PS) e cobertura populacional.. ....	62
<b>Tabela 4-</b> Epítomos virais de proteínas NiV reconhecidos pelos linfócitos T auxiliares. ....	63
<b>Tabela 5-</b> Epítomos virais reconhecidos por linfócitos B previstos a partir do proteoma do vírus. ....	64
<b>Tabela 6 -</b> <i>C-score</i> da predição estrutural obtida no servidor C-Quark, Z-Score obtido no servidor Prosa-Web, índice QMEAN obtido no servidor Swiss Model e distribuição dos ângulos residuais de acordo com as regiões do gráfico de Ramachandran. ....	66
<b>Tabela 7 -</b> Alinhamento estrutural entre epítomos previstos e estruturas de paramixovírus depositadas no <i>Protein Data Bank</i> (PDB).....	67
<b>Tabela 8:</b> Propriedades estruturais e físico-químicas de cada modelo multi-epítopo projetado. ....	71
<b>Tabela 9 -</b> Predição de alergenicidade dos seis modelos de multi-epítomos projetados usando métodos de alinhamento baseados em sequência (SDAP, AllerTOP e AlgPred Blast).....	71
<b>Tabela 10 -</b> Pontuações de docagem do HDOCK obtidas a partir do docagem molecular entre os modelos de multi-epítomos projetados (modelos 1 a 6) com a estrutura TLR3.....	72
<b>Tabela 11 -</b> A energia de ligação livre do receptor TLR3 complexado com as proteínas multi-epítopo projetadas.....	78

## LISTA DE SIGLAS

AI	Índice alifático
BCR	Receptores de células B
BHE	Barreira hematoencefálica
BLAST	<i>Basic local alignment search tool</i>
C	Proteína C
CAI	Índice de adaptação de códons
cDNA	DNA complementar
CEPI	Coalizão para Inovações em Preparação para Epidemias
CTL	Linfócitos T citotóxicos
CXCL10	C-X-C quimiocina 10
DM	Dinâmica molecular
DRX	Difração de raios X
EBV	Vírus Epstein–Barr
ER	Retículo endoplasmático
EV-A71	Enterovírus A71
ExPASy	<i>Expert protein analysis system</i>
F	Glicoproteína de fusão F0
G	Glicoproteína G
GC	Guanina-citosina
G-CSF	Fator estimulante de colônias de granulócitos
GPU	Unidade de processamento gráfico
GRAVY	Grande média de hidropatia
HeV	Hendra vírus
HLA	Antígeno leucocitário humano
HTL	Linfócito T auxiliador
HMM	Modelo oculto de Markov
IEDB	Banco de Dados de Epitopos Imunológicos
IFN- $\gamma$	Interferon-gama
IgE	Imunoglobulina E
IgG	Imunoglobulina G
II	Índice de instabilidade

JAK-STAT	Janus quinase/transdutores de sinal e ativadores de transcrição
JEV	Vírus da encefalite japonesa do vírus B
L	RNA polimerase L dependente de RNA
LRR	<i>Leucine-rich repeat</i>
M	Proteína de matriz
mAb	Anticorpo monoclonal
MHC	Complexo principal de histocompatibilidade
MMPBSA	<i>Molecular Mechanics Poisson–Boltzmann Surface Area</i>
MMPGBSA	<i>Molecular Mechanics Generalized born Surface Area</i>
MPI	Interface de transmissão de mensagens
N	Nucleoproteína
nAbs	Anticorpos neutralizantes
NCBI	<i>National Center for Biotechnology Information</i>
NiV	Vírus nipah
NK	Células <i>natural killer</i>
OMS	Organização Mundial da Saúde
P	Fosfoproteína
PDB	<i>Protein Data Bank</i>
PI	Ponto isoelétrico
PME	<i>Particle-mesh ewald</i>
PRR	Receptores de reconhecimento de padrões
PTM	Modificações pós-traducionais
R0	Número reprodutivo básico
RDT	Tecnologia de DNA recombinante
RMN	Ressonância magnética nuclear
RMSD	Desvio quadrático médio
RMSF	Variação quadrada média
SDRA	Síndrome do desconforto respiratório agudo
SNC	Sistema nervoso central
STAT1	Transdutor de sinal e ativador da transcrição 1
STAT2	Transdutor de sinal e ativador da transcrição 2
SVM	Máquina de vetores de suporte
TAP	Transportador associado à proteína de processamento de antígeno

TCR	Receptores de Células T
Th1	T helper 1
TLR	Receptor <i>toll-like</i>
TNF- $\alpha$	Fator de necrose tumoral alfa
UNIPROT	<i>Universal protein resource</i>
V	Proteína V
VLP	Partículas semelhantes a vírus
VSV	Vetores do Vírus da Estomatite Vesicular
W	Proteína W

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	16
<b>2 REVISÃO DE LITERATURA</b> .....	20
<b>2.1 Vírus Nipah</b> .....	20
2.1.1 Características biológicas .....	21
2.1.2 Patogênese viral e imunobiologia da infecção .....	23
2.1.3 Sinais clínicos e sintomas .....	27
2.1.4 Fatores de risco e mecanismos de transmissão do vírus Nipah .....	28
2.1.5 Tratamento e profilaxia .....	28
2.1.6 Importância na saúde pública e potencial de pandemia .....	31
<b>2.2 Visão geral do sistema imune</b> .....	34
2.2.1 Imunidade Inata .....	35
2.2.2 Imunidade Adaptativa.....	35
<b>2.3 Planejamento <i>in silico</i> de vacinas</b> .....	40
2.3.1 Planejamento de vacinas multi-epítipo .....	41
2.3.2 Imunoinformática .....	42
2.3.3 Metodologias para o planejamento de vacinas multi-epítipo .....	43
2.3.4 Exemplos de aplicações.....	51
<b>3 OBJETIVOS</b> .....	53
<b>3.1 Objetivo geral</b> .....	53
<b>3.2 Objetivos específicos</b> .....	53
<b>4 MATERIAL E MÉTODOS</b> .....	54
<b>4.1 Obtenção de proteomas virais e avaliação de similaridade a proteínas humanas</b> .....	54
<b>4.2 Predição de epítomos de linfócitos T e indutores de IFN-<math>\gamma</math></b> .....	54
<b>4.3 Predição de imunogenicidade</b> .....	54
<b>4.4 Predição de processamento dos epítomos com afinidade a MHC classe I</b> .....	54
<b>4.5 Predição de epítomos de células B</b> .....	54
<b>4.6 Filtragem de epítomos</b> .....	54
<b>4.7 Construção dos modelos de vacina</b> .....	54



4.8 Avaliação de antigenicidade, alergenicidade dos modelos e propriedades físico-químicas.....	54
4.9 Predição de estruturas terciárias e validação .....	54
4.10 Docagem molecular .....	54
4.11 Simulações de Dinâmica Molecular (DM).....	54
4.12 Cálculos de energia livre de ligações .....	54
4.13 Otimização de códons e expressão <i>in silico</i> em sistema vetorial .....	55
<b>5 RESULTADOS E DISCUSSÃO .....</b>	<b>61</b>
5.1 Predição de antigenicidade e avaliação de homologia de proteínas.....	61
5.2 Predição e filtragem de epítopos de células T e B.....	62
5.3 Construção dos modelos de vacina .....	64
5.4 Predição e validação de estruturas das proteínas multiepítipo .....	65
5.5 Predição de linfócitos B conformacionais/descontínuos .....	69
5.6 Propriedades físico-químicas, antigenicidade e alergenicidade .....	70
5.7 Docagem Molecular .....	72
5.8 Dinâmica Molecular .....	76
5.9 Clonagem <i>in silico</i> .....	79
<b>6 CONSIDERAÇÕES FINAIS.....</b>	<b>81</b>
<b>REFERÊNCIAS .....</b>	<b>82</b>
<b>APÊNDICES .....</b>	<b>99</b>

## 1 INTRODUÇÃO

O vírus Nipah (NiV) é um patógeno da família Paramyxoviridae, gênero *Henipavirus*, considerado como um vírus zoonótico emergente mortal, dotado de potencial para gerar graves danos à saúde humana e animal (RAMPHUL et al., 2018). O NiV foi isolado pela primeira vez na Malásia durante um surto em 1998 (CHUA et al., 1999) e, subsequentemente, em Cingapura (1999) (PATON et al., 1999), Bangladesh (2001) (HSU et al., 2004), Índia (2001) (CHADHA et al., 2006a) e nas Filipinas (2014) (CHING et al., 2015). Desde então, pequenos surtos esporádicos foram observados no Sul e Sudeste Asiático, com taxas de mortalidade variando entre 40% e 80% (ARANKALLE et al., 2011). Os surtos têm seguido a distribuição geográfica da expansão de atividades econômicas em regiões de presença gênero de morcegos frugívoros *Pteropus*, reservatório responsável pela expansão desta infecção por secreção de saliva ou urina em frutas consumidas por animais e humanos (RAHMAN et al., 2010). NiV é transmitido de animais, como porcos e outros mamíferos para humanos e também pode ser transmitido diretamente entre pessoas (HALPIN et al., 2011). Os sintomas geralmente aparecem entre 5 e 14 dias após a exposição (RAVEENDRAN et al., 2018a). As infecções humanas variam de infecção assintomática a manifestações clínicas sintomáticas com febre, dores de cabeça, mialgia, infecção respiratória aguda e encefalite, com potencial para progredir para coma em um período bastante curto (24 a 48 h) (ARUNKUMAR et al., 2018; HOSSAIN et al., 2008). Além disso, condições neurológicas de longo prazo foram relatadas em sobreviventes (CHANG et al., 2007). Atualmente, não há medicamentos ou vacinas disponíveis comercialmente para a infecção por NiV. O tratamento básico é baseado em medicamentos antivirais como favipiravir e ribavirina, que mostraram atuar contra os sintomas dos pacientes com eficácia modesta, no entanto, apresentando vários efeitos colaterais (BRODER et al., 2013; DAWES et al., 2018).

Vários hospedeiros do NiV mantêm contato com as civilizações humanas, o que favorece as adaptações virais ao genoma humano induzidas por forças evolutivas. Assim, as cepas emergentes de NiV com maior transmissibilidade humana poderiam gerar consequências devastadoras à saúde mundial (LUBY, 2013). Recentemente, um estudo usando uma abordagem baseada em aprendizado de máquina identificou espécies de morcego fora do gênero *Pteropus* como hospedeiros potenciais de NiV na Ásia, Austrália e Oceania, uma região coberta por aproximadamente 2 bilhões de pessoas (PLOWRIGHT et al., 2019). Dada sua ameaça potencial à segurança da saúde global, a infecção por NiV foi designada como doença prioritária pelo Projeto de Pesquisa e Desenvolvimento da Organização Mundial da Saúde (OMS)

(REUSKEN, 2018). O NiV também foi incluído como uma prioridade para o desenvolvimento de vacinas pela Coalizão para Inovações em Preparação para Epidemias (CEPI) e é classificado como um vírus de nível 4 de biossegurança. Além disso, seu alto potencial para atuar como agente de bioterrorismo demarca sua categoria como patógeno C, isto é, patógeno considerado potencial para gerar doenças infecciosas emergentes e que pode ser utilizado para disseminação em massa (RAVEENDRAN et al., 2018). Apesar de sua relevância para a saúde pública, ainda poucos avanços científicos foram alcançados para encontrar um tratamento adequado e prevenção da infecção por NiV. Portanto, há uma necessidade urgente de pesquisa e desenvolvimento de medicamentos e vacinas para prevenir e/ou controlar futuros surtos.

Ferramentas modernas têm sido progressivamente aplicadas, incluindo tecnologia de DNA recombinante e design de vacina baseado em epítomos. Em comparação com as vacinas clássicas e vacinas de epítomo único, as vacinas de múltiplos epítomos têm conceitos de projeto especiais: consistem em epítomos que podem ser reconhecidos por receptores de células T (TCR) ou receptores de células B (BCR) de vários subtipos, induzem fortes respostas imunes celulares e humorais, expandem o espectro da resposta imune pela associação de vários antígenos. Além disso, esta abordagem introduz componentes estruturais validados que aumentam a imunogenicidade, tais como adjuvantes, e podem reduzir a presença de epítomos indesejados que possam desencadear respostas patológicas ou efeitos adversos (LENNERZ et al., 2014; LIN et al., 2016; LU et al., 2017a; SAADI; KARKHAH; NOURI, 2017). Com base nisso, as ferramentas de imunoinformática e de modelagem molecular podem auxiliar no planejamento de proteínas imunogênicas multi-epítomo com o objetivo de desenvolver uma vacina de próxima geração (BRENNICK et al., 2017; KUO et al., 2014). Esta estratégia foi aplicada com sucesso no estudo de diferentes vírus, como o herpesvírus humano-8 (CHAUHAN et al., 2019), vírus da dengue (ALI et al., 2017), vírus da hepatite C (IKRAM et al., 2018), vírus zika (USMAN MIRZA et al., 2016), coronavírus relacionado à síndrome respiratória do Oriente Médio (SRIVASTAVA et al., 2018), vírus chikungunya (HASAN et al., 2015) e vírus Mayaro (RODRIGUES et al., 2019). Além disso, algumas vacinas de múltiplos epítomos entraram em ensaios clínicos (LENNERZ et al., 2014; OYARZÚN; KOBE, 2016). O proteoma do NiV contém nove potenciais proteínas candidatas envolvidas na evasão imunológica e patogênese no hospedeiro humano que pode ser explorada para a produção de novas vacinas (Tabela 1) (SATTEFIELD et al., 2015; SATTEFIELD; DAWES; MILLIGAN, 2016a).

**Tabela 1-** Proteínas codificadas pelo genoma NiV, identificadas por seu símbolo, código de acesso Uniprot, e suas funções moleculares envolvidas com o ciclo de vida do vírus.

Proteína	Código de acesso (UniProt)	Funções moleculares	Referências
Glicoproteína de fusão F0 (F)	Q9IH63	Fusão de membranas virais e celulares	DIEDERICH; MAISNER, 2007; JARDETZKY; LAMB, 2014
Glicoproteína G (G)	Q9IH62	Fixação à célula hospedeira de destino	DIEDERICH; MAISNER, 2007; JARDETZKY; LAMB, 2014
Proteína de matriz (M)	Q9IK90	Montagem e brotamento de vírions	BHARAJ et al., 2016; WATKINSON; LEE, 2016
Nucleoproteína (N)	Q9IK92	Encapsidação do genoma	HUANG et al., 2011; YABUKARSKI et al., 2014
RNA polimerase L dependente de RNA (L)	Q997F0	Transcrição de mRNA viral, seu capeamento e poliadenilação	HARCOURT et al., 2000; TOMAR; KAUR; SINGH, 2019
Fosfoproteína (P)	Q9IK91	Componente do complexo de transcrição e replicação de RNA polimerase	BRUHN et al., 2014
Proteína C (C)	Q997F1	Bloqueia o sistema antiviral de interferon celular	MATHIEU et al., 2012; PARK et al., 2003
Proteína V (V)	Q997F2	Inibição da resposta imune do hospedeiro	RODRIGUEZ; PARISIEN; HORVATH, 2002; SATTERFIELD et al., 2015
Proteína W (W)	P0C1C7	Bloqueio de IFN- $\alpha$ e IFN- $\beta$	PARK et al., 2003; SATTERFIELD et al., 2015

**Fonte:** Elaborado pelo Autor.

Atualmente, vários estudos *in silico* se concentraram na análise de epítomos das proteínas F e G do NiV (ALI; MORSHED; HASSAN, 2015; SAHA et al., 2017; SAKIB et al., 2014). No entanto, para projetar proteínas multi-epítomo visando o desenvolvimento de uma vacina, é relevante desenhar a estrutura com inclusão de partes mais antigênicas de múltiplas proteínas virais para eliciar a resposta imune (STERN et al., 2017). Com base nessa premissa, no presente estudo, foram exploradas duas estratégias envolvidas com o projeto computacional de antígenos multi-epítomos do NiV: Na primeira, foram selecionados epítomos da glicoproteína de fusão (F) e glicoproteína G, ambas presentes na superfície viral e validadas experimentalmente com atividade imunomodulatória contra NiV (BOSSART et al., 2012; DEBUYSSCHER et al., 2014; MCEACHERN et al., 2008). A segunda consistiu em uma seleção de epítomos de proteínas codificadas pelo genoma de NiV com atividades

imunomodulatórias preditas (MATHIEU et al., 2012; SATTERFIELD et al., 2015). Dessa forma, foi avaliado o potencial de duas proteínas imunogênicas de múltiplos epítomos com potencial para auxiliar no desenvolvimento de uma vacina de próxima geração para combater a infecção por NiV.

## 2 REVISÃO DE LITERATURA

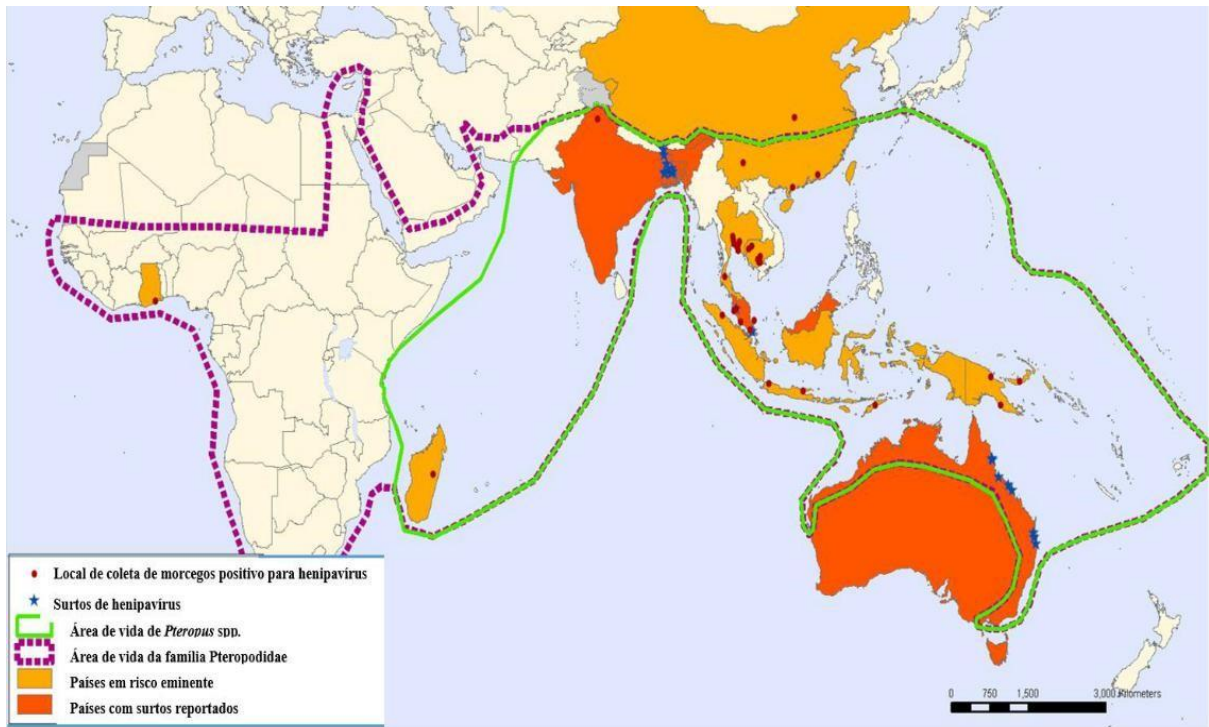
### 2.1 Vírus Nipah

O vírus Nipah (NiV) é um patógeno zoonótico emergente, pertencente à família Paramyxoviridae, constituindo-se como grande ameaça para a saúde humana e animal; É definido pela Organização Mundial da Saúde como uma doença prioritária para a pesquisa científica (REUSKEN, 2018). O vírus foi identificado pela primeira vez após um surto de doença respiratória e neurológica em porcos e subsequente encefalite em humanos na Malásia e Singapura em 1998-99 (PATON et al., 1999).

Inicialmente, pensou-se que o surto fosse causado pelo vírus da encefalite japonesa do vírus B (JEV) transmitido por mosquito devido à sua associação com porcos infectados e à detecção de IgM específica para o JEV em soros de pacientes. No entanto, a infecção pareceu ser contraída após contato direto com porcos, o que não é uma característica da infecção por JEV, assim como os principais afetados foram indivíduos adultos, diferentemente de JEV. Os isolados de vírus foram subsequentemente confirmados como sendo um paramixovírus e o vírus foi nomeado como Nipah, após surto na aldeia de Sungai Nipah, onde o vírus foi isolado pela primeira vez (CHUA, 2000; PATON et al., 1999).

A epidemiologia relacionada ao NiV não foi totalmente compreendida. No entanto, surtos limitados de NiV em humanos já foram documentados na Malásia, Cingapura, Bangladesh e Índia. O mais recente ocorreu em 2018, onde 17 fatalidades foram registradas a partir de encefalite associada a Nipah no estado de Kerala, na Índia (ARUNKUMAR et al., 2018). O gênero de morcegos *Pteropus* é o principal responsável por espalhar essa infecção secretando saliva ou urina em frutas consumidas por animais e humanos (HALPIN et al., 2011). De acordo com o relatório do *National Center for Disease Control*, após o primeiro surto significativo do (TAN et al., 2002), os surtos tem periodicamente se expandido de acordo com a localização geográfica dos morcegos (Figura 1).

**Figura 1-** Distribuição geográfica dos surtos henipavirais e morcegos da família Pteropodidae.



**Fonte:** Departamento de Alerta e Resposta Global - Organização Mundial de Saúde, 2008.

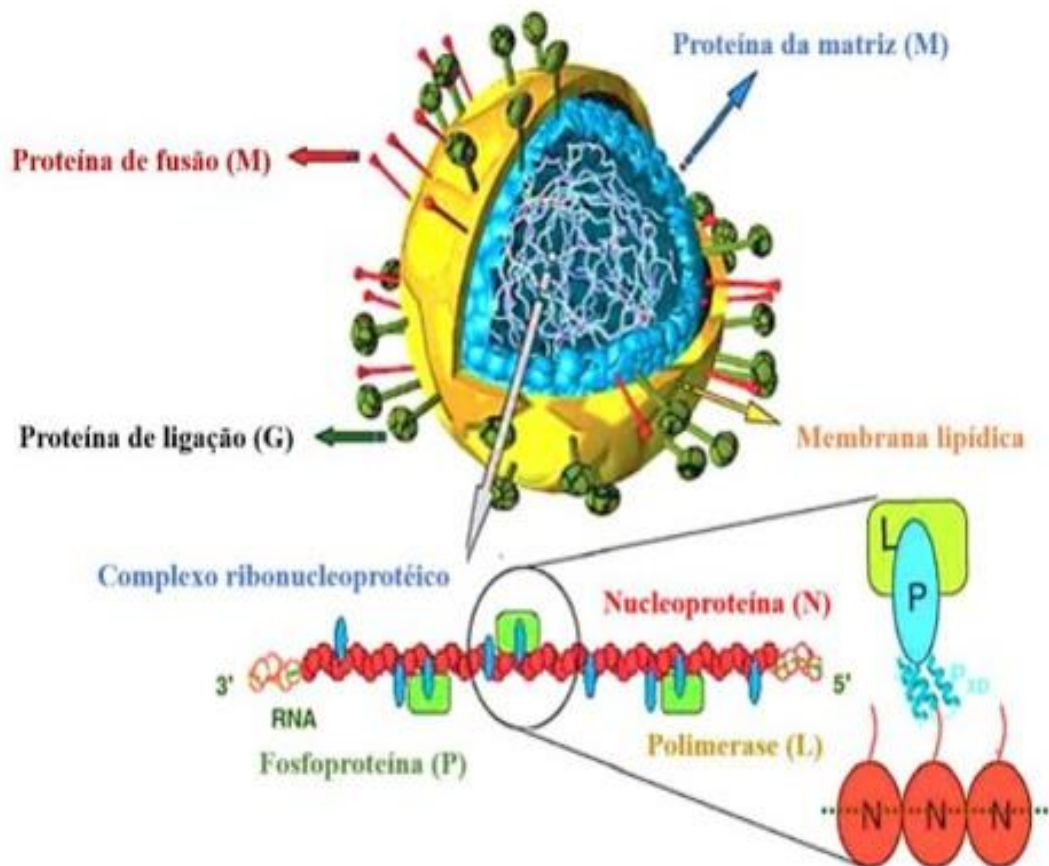
O NiV é um dos vírus zoonóticos emergentes de maior letalidade conhecidos (CLAYTON, 2017; THIBAUT et al., 2017). O potencial para expansão do contágio viral parece derivar de várias características: disponibilidade de população humana suscetível, várias cepas virais com potencial para transmissão de pessoa para pessoa e a natureza propensa a erros de replicação do vírus de RNA. Surtos de doença de NiV em regiões densamente povoadas como o Sul da Ásia podem levar a pandemias, devido à ampla conectividade global (LUBY, 2013). Muitos fatores ecológicos e moleculares estão por trás da expansão do NiV em seres humanos, embora a complicada interação entre eles não seja clara (THIBAUT et al., 2017).

### 2.1.1 Características biológicas

O vírus Nipah é envelopado, de formato pleomórfico, com genoma de RNA de sentido negativo de fita única (BRODER; WONG, 2016). Sob microscópio eletrônico, o NiV possui um padrão de estrutura morfológica semelhante ao de outros membros da família Paramyxoviridae. Como outros paramixovírus, o NiV também possui seis genes, que codificam: proteína de fusão (F), glicoproteína (G), proteína de matriz (M), nucleocapsídeo (N), fosfoproteína (P) e proteína polimerase (L). O gene da fosfoproteína (P) codifica ainda outras três proteínas acessórias importantes conhecidas como C, V e W (CLAYTON; WANG; MARSH, 2013; HARCOURT et al., 2000, 2005), totalizando nove proteínas virais.

A proteína C desempenha um papel muito importante na regulação da síntese de RNA viral e do fator de virulência. As proteínas V e W são cruciais para a virulência e atuam inibindo a ativação de um promotor indutor de interferon (HARCOURT et al., 2000, 2005). A proteína W atua ainda como inibidora da via de sinalização do receptor *toll-like 3* (TLR3), enquanto as proteínas V e C atuam na sinalização de interferon-gama (IFN- $\gamma$ ), interagindo com STAT1 e proteína STAT2 (CHUA, 2000; SHAW et al., 2005). A proteína N é responsável pela encapsidação do genoma do vírus e, portanto, protege a atividade da nuclease (YABUKARSKI et al., 2014). Além disso, a proteína da matriz, a proteína M é responsável pela integridade estrutural do vírion. A glicoproteína de fusão F mostra um papel promissor na fusão da membrana viral com a membrana da célula hospedeira, resultando na entrega de todo o conteúdo no citoplasma hospedeiro (HARCOURT et al., 2000). Por fim, a glicoproteína G contribui para a fixação do vírus para os receptores da superfície da célula hospedeira (HARCOURT et al., 2000). A proteína G tetramérica forma uma interação bidentada com a proteína F homotrimérica do NiV antes da ligação ao receptor (STONE et al., 2016).

**Figura 2-** Representação dos principais constituintes da estrutura do Nipah vírus.



**Fonte:** Adaptado de Singh et al., 2019.



Variações de cepas foram observadas entre isolados humanos de NiV coletados de surtos na Malásia, Índia e Bangladesh. Da mesma forma, isolados de vírus NiV de amostras de morcegos que foram coletados em diferentes áreas geográficas também apresentaram variação genômica (CHADHA et al., 2006b; HARCOURT et al., 2005; REYNES et al., 2005). Uma análise de similaridade pareada entre as seqüências de nucleotídeos recuperadas dos genes N e G de NiV foi realizada através de bioinformática (Tamura et al. 2013), mostrando a circulação de duas grandes cepas de NiV: NiV-B (Bangladesh) e NiV-M (Malásia). Estas representam duas linhagens com comprimento do genoma de 18.246 nucleotídeos e 18.252 nucleotídeos, respectivamente. A homologia de aminoácidos entre as proteínas expressas por NiV-M e NiV-B é >92%, enquanto há 91,8% de similaridade entre as sequências de nucleotídeos (HARCOURT et al., 2005). Como o NiV é um vírus de RNA de cadeia simples, os virions precisam empacotar sua própria RNA polimerase dependente de RNA, codificada pelo gene L, para permitir a transcrição e a replicação (AGUILAR et al., 2016).

O NiV mostra uma relação estreita com o hendra vírus (HeV), estando subseqüentemente classificados em seu próprio gênero (henipavirus) (RAVEENDRAN et al., 2018). NiV mostra 68% a 92% e 40% a 67% de similaridade com o hendra vírus (HeV) em regiões codificadas por proteínas e regiões não traduzidas, respectivamente (CLAYTON; WANG; MARSH, 2013; GOLDSMITH et al., 2003; WANG et al., 2001). Anticorpos contra o NiV também reagem de forma cruzada com HeV, mas não contra outros membros da família (ZHU et al., 2005). Em contraste com outros paramixovírus, os genomas de NiV e HeV são muito mais longos, isto é, 18,2 e 15,5 kb, respectivamente. Os genes P de ambos os vírus são mais longos que outros membros da família Paramyxoviridae (MAYO, 2002; WANG et al., 2000). Ambos os vírus não mostram as atividades de hemaglutinina e neuraminidase que são características comuns de outros paramixovírus. E também, ambos os vírus têm ampla gama de hospedeiros, enquanto outros membros do paramixovírus têm um pequeno número de hospedeiros (WANG et al., 2001; YU et al., 1998). Mais recentemente, vírus relacionados, como o vírus cedar, o vírus do morcego ganês e o vírus henipa-like mojiang também foram classificados neste gênero, embora a evidência de infecção e doença em humanos seja limitada (FERNÁNDEZ-AGUILAR et al., 2014; J.F. et al., 2009; MARSH et al., 2012).

### 2.1.2 Patogênese viral e imunobiologia da infecção

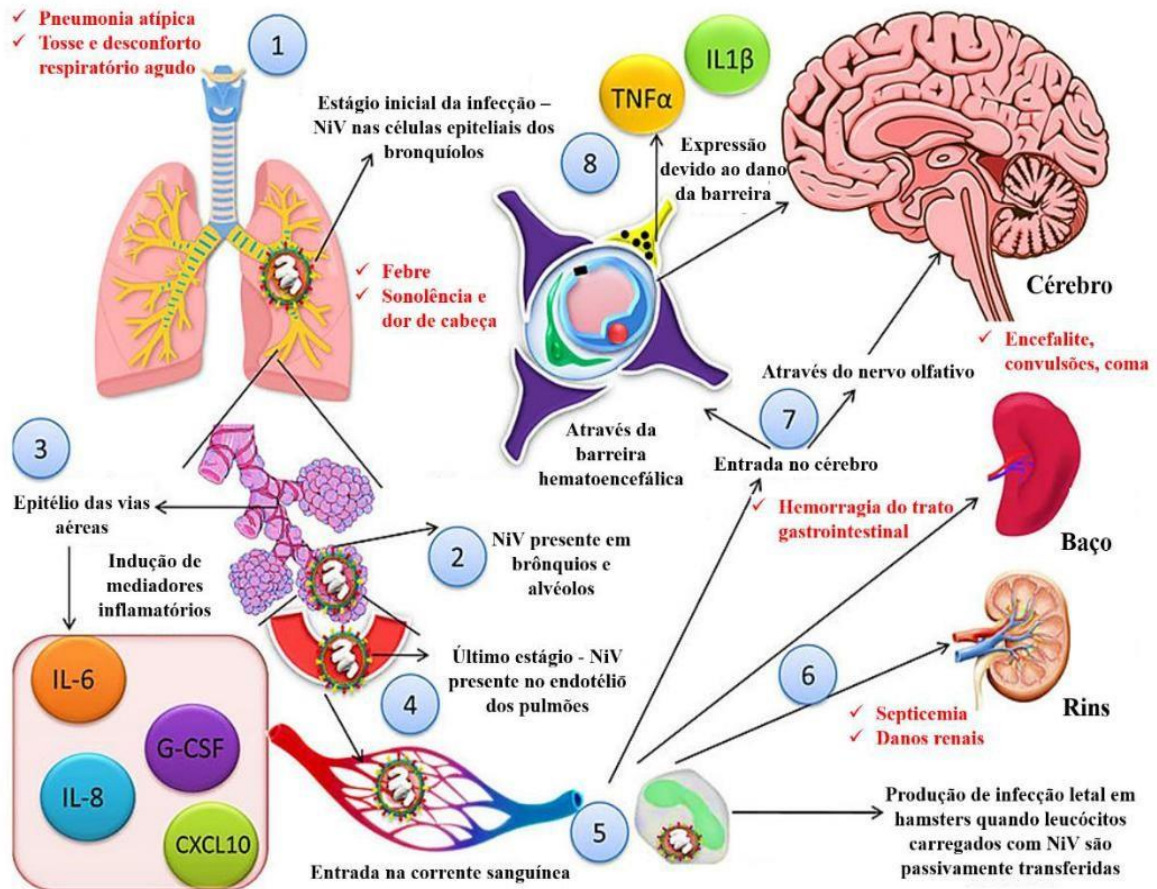
No estágio inicial da doença em humanos, a detecção de NiV pode ser feita em células epiteliais do bronquíolo (CHUA et al., 2001). Antígenos virais podem ser detectados em

brônquios e alvéolos em modelos animais experimentais; os alvos primários são epitélio dos brônquios e pneumócitos tipo II (ROCKX et al., 2010). As citocinas inflamatórias são induzidas devido à infecção do epitélio do trato respiratório; assim, o recrutamento de células do sistema imunológico e, finalmente, levando ao desenvolvimento de doença semelhante à síndrome do desconforto respiratório agudo (SDRA) (ROCKX et al., 2010). Mediadores inflamatórios significativos, isto é, as interleucina 1a, 6 e 8, o fator estimulante de colônias de granulócitos (G-CSF), o motivo C-X-C quimiocina 10 (CXCL10), dentre outros, são induzidos quando o epitélio das vias aéreas (menores) fica infectado (ESCAFFRE et al., 2013). Do epitélio respiratório, o vírus é disseminado para as células endoteliais dos pulmões na fase tardia da doença. Subsequentemente, o vírus pode entrar na corrente sanguínea, seguido de disseminação, livremente ou na forma ligada aos leucócitos do hospedeiro. Além dos pulmões, o baço e os rins, juntamente com o cérebro, podem atuar como órgãos-alvo, levando à falência múltipla de órgãos (ESCAFFRE et al., 2013; ROCKX et al., 2010).

Há desenvolvimento de infecção letal em hamsters quando leucócitos carregados com NiV são passivamente transferidos (MATHIEU et al., 2011). Em suínos, há infecção produtiva de monócitos, células *natural killer* (NK) juntamente com linfócitos T CD4+ e CD8+ (STACHOWIAK; WEINGARTL, 2012).

Duas vias estão distintamente envolvidas no processo de entrada viral no sistema nervoso central (SNC), ou seja, via hematogênica (através do plexo coroide ou vasos sanguíneos do cérebro) e/ou anterogradamente através de nervos olfatórios (WEINGARTL et al., 2005). A barreira hematoencefálica (BHE) é rompida e IL-1b juntamente com o fator de necrose tumoral alfa (TNF- $\alpha$ ) são expressas devido à infecção do SNC pelo vírus, o que acaba levando ao desenvolvimento de sinais neurológicos (ROCKX et al., 2010). Pode haver presença de corpos de inclusão no caso de um SNC infectado em humanos. Tanto nas placas cinzentas quanto nas brancas, as placas podem ficar evidentes junto com a necrose (ESCAFFRE et al., 2013). É importante destacar que o vírus pode entrar diretamente no SNC em vários modelos animais experimentais através do nervo olfativo. O epitélio olfatório da concha nasal é infectado por NiV em tais modelos animais. A infecção viral se estende posteriormente através da placa cribiforme para o bulbo olfatório. Em última análise, o vírus é disseminado por todo o córtex ventral, juntamente com o tubérculo olfativo (ESCAFFRE et al., 2013; PRESCOTT et al., 2012; WEINGARTL et al., 2005). A figura 3 ilustra, em síntese, a patogênese do vírus Nipah em humanos.

**Figura 3** - Representação da patogênese do NiV.



**Fonte:** Adaptado de Singh et al. (2019).

Estudos de resposta imunológica em relação ao vírus Nipah foram realizados por vários pesquisadores. Uma vez que o vírus exibe dois tipos de associação entre seus hospedeiros (mantendo sua persistência na natureza por meio de hospedeiros reservatórios como morcegos e causando condições clínicas fatais em animais domésticos como porcos, além dos seres humanos), as respostas imunes podem ser específicas do hospedeiro (CHUA, 2000; KULKARNI et al., 2013; NEGRETE et al., 2005). Várias proteínas dos henipavírus bloqueiam as respostas imunes inatas do hospedeiro, sobretudo as proteínas V, C e W. Em resposta a vários estímulos, a produção de interferons pode ser inibida tanto pelas proteínas V quanto W, enquanto a capacidade dos interferons para sinalização é bloqueada pelas proteínas P, V e W, levando à indução do bloqueio de um estado de resposta antiviral celular (BASLER, 2012).

O sistema imunológico inato de morcegos pteropídeos é notável por possuir um sistema de interferon Tipo 1 que pode restringir a replicação viral precoce dentro de seu corpo (ZHOU et al., 2016). Este modo de ação tem sido associado também a vários genes estimulados por interferon (ISG), particularmente aqueles envolvidos em vias não inflamatórias, de modo que

a elevação da resposta do interferon em morcegos não é aliada à inflamação crônica, como no caso de roedores e humanos (HALPIN et al., 2011). Devido a estas diferenças, as células de morcego são preparadas para reagir imediatamente ao ataque viral, mas apenas até um nível de restrição de replicação (ZHOU et al., 2011a, 2011b). Os morcegos possuem um repertório comparativamente maior de imunoglobulinas simples com mais especificidades, favorecendo a seleção clonal direta de linfócitos B para a produção de anticorpos.

Nessa condição, pode haver estágios de maturação de hipermutação e afinidade pobres ou inexistentes nas células B, levando a respostas mais fracas e produção restrita de anticorpos do que outras espécies. Estas características contribuem para o atraso na depuração viral e persistência do vírus por um período bastante longo (SCHOUNTZ et al., 2017). Estudos comparativos do vírus Nipah conduzidos em morcegos pteropídeos e hamsters reforçam esses pontos, já que o vírus apresentou menor multiplicação e desprendimento das células endoteliais do morcego, bem como com respostas de anticorpos mais fracas após os estudos (LO et al., 2010).

Outro mecanismo imune dentro dos morcegos para prevenir a completa eliminação do vírus Nipah é a modulação das respostas antivirais dos morcegos em relação à sobrevivência do vírus. No hospedeiro reservatório, o vírus emprega estratégias de evasão imune, especialmente contra o sistema imunológico inato, para escapar do ataque imune e manter a perpetuação no hospedeiro (RUPPRECHT; TURMELLE; KUZMIN, 2011).

Tais estratégias de evasão são mediadas através de proteínas acessórias codificadas no vírus (SCHOUNTZ et al., 2017). O gene P do vírus Nipah (que codifica a fosfoproteína associada à polimerase) desempenha um papel fundamental na evasão da resposta imune do hospedeiro (SHAW, 2009), inibindo as respostas antivirais através do bloqueio das vias de sinalização mediadas por interferons, especialmente a via de sinalização JAK-STAT estimulada por STAT1 (SHAW, 2009).

Embora o vírus Nipah efetivamente suprima a produção de citocinas antivirais na fase inicial da infecção, foi sugerida a liberação de alguma quantidade de citocinas inflamatórias, o que pode ser atribuído à elevação da permeabilidade vascular, favorecendo a disseminação viral (SCHOUNTZ, 2014). Presença de inclusões antígeno-positivas nos tecidos cerebrais de pacientes com Nipah e encefalite viral apontam para a inadequação de respostas inatas e adaptativas para prevenir a disseminação viral. Esses resultados sugerem a incapacidade de células dendríticas residirem no ponto de entrada primário do vírus; especialmente no trato respiratório e pulmões, tornando ineficiente a captura de antígenos e restrição de tecidos (TAN et al., 2002). Evidências também sugerem a supressão da expressão do complexo principal de

histocompatibilidade de classe I (MHC-I) em células imunes pelas proteínas virais, levando a uma repressão tanto na apresentação do antígeno pelas células apresentadoras de antígeno quanto à estimulação para a montagem de respostas adaptativas, resultando em disseminação e persistência virais em outros órgãos-alvo (DASGUPTA et al., 2014; SETO et al., 2010).

Além disso, a evasão imunológica por longo tempo também é responsável pela persistência do vírus nos tecidos cerebrais e conseqüente recaída e início tardio da encefalite induzida por citocinas e quimiocinas que destroem o tecido nervoso (TAN et al., 2002). O papel crucial das moléculas de RNA mensageiro (mRNA) presentes no genoma de NiV na inibição de vários genes hospedeiros, auxiliando assim a propagação viral e patogênese também tem sido relatados (SAINI et al., 2018). Dessa forma, a patogenicidade do vírus Nipah em suínos e em humanos pode ser correlacionada com sua capacidade e magnitude de escapar das respostas imunes no hospedeiro reservatório. Embora o vírus atinja vários hospedeiros, taxas mais altas de letalidade estão sendo associadas com surtos humanos até o momento, o que justifica um estudo abrangente para elucidar e explorar a evolução viral e adaptação em diferentes hospedeiros (KULKARNI et al., 2013).

### 2.1.3 Sinais clínicos e sintomas

A infecção por NiV em humanos produz uma síndrome encefalítica caracterizada principalmente por cefaléia, pirexia e outros sintomas neurológicos. A febre foi encontrada como sintoma universal seguido de cefaleia em 65% a 88% dos pacientes com a infecção (GOH et al., 2002). Houve grandes diferenças entre o nível de consciência dos pacientes infectados, variando entre 55% a 90% (HSU et al., 2004; LUBY et al., 2009). Além da febre, outros sintomas como vômitos, tontura, anormalidades do tronco encefálico e reflexos reduzidos ou ausentes também foram observados com frequência durante surtos de NiV (GURLEY et al., 2007; ICDDR, 2004).

Os sintomas respiratórios estão em segundo lugar entre os mais comuns, ficando atrás apenas dos sintomas neurológicos da infecção por NiV. Tosse, frio e dispnéia foram os sintomas respiratórios mais comuns relatados, enquanto dor abdominal, diarreia, gastrite e constipação foram relatados em poucos casos (CROSER; MARSH, 2013; GOH et al., 2002). Taxas mais altas de infecções respiratórias e torácicas nos surtos de Bangladesh do que durante os surtos de Malásia e Cingapura podem ser atribuídas à transmissão do vírus entre humanos no Bangladesh (CLAYTON et al., 2012; ISLAM et al., 2016).

#### 2.1.4 Fatores de risco e mecanismos de transmissão do vírus Nipah

Os estudos epidemiológicos de surtos de NiV na Malásia, Singapura, Bangladesh, Filipinas e Índia sugeriram que vários fatores desempenham um papel crucial na transmissão de NiV para humanos. O contato próximo com animais infectados com NiV, reservatórios animais e consumo de alimentos contaminados são fatores responsáveis para a transmissão (ISLAM et al., 2016). Porcos infectados com NiV foram observados como fonte principal de infecções humanas (92%) durante surtos na Malásia e em Singapura (PARASHAR et al., 2002). A infecção por NiV entre suínos e humanos provavelmente ocorreu por via respiratória. O controle e o estudo experimental também comprovam a via oral e respiratória como as principais vias de transmissão (BRODER; WONG, 2016).

No caso de surtos de Bangladesh, não foram encontradas evidências claras de transmissão por porcos, mas o consumo de alimentos contaminados por secreções de raposas voadoras foi uma das principais fontes de infecção por NiV (RAVEENDRAN et al., 2018b). O NiV foi isolado da urina e do trato respiratório de amostras de humanos infectados durante os surtos da Malásia, sugerindo a possibilidade de transmissão humana (SAZZAD et al., 2013). Em um estudo envolvendo furões como modelo animal, foi relatado que a cepa de Bangladesh de NiV exibe maior excreção oral do que outras cepas. No entanto, seu mecanismo ainda não foi bem elucidado (CLAYTON et al., 2012). A transmissão pessoa a pessoa não foi provada experimentalmente em surtos de Malásia e Singapura, mas há fortes evidências da transmissão pessoa a pessoa em Bangladesh e em surtos ocorridos na Índia (GURLEY et al., 2007; ISLAM et al., 2016).

#### 2.1.5 Tratamento e profilaxia

O tratamento para o NiV é ainda paliativo, porque não existem antivirais ou vacinas comercialmente disponíveis (SINGH et al., 2019). A ribavirina e o aciclovir chegaram a ser utilizados para tratar a infecção por NiV em surtos anteriores (THIBAULT et al., 2017). No surto da Malásia, a ribavirina foi administrada por via oral ou intravenosa em pacientes com encefalite; a taxa de mortalidade foi reduzida em 36% (CHONG et al., 2001). No surto de Singapura, o aciclovir foi administrado a todos os pacientes com encefalite por NiV, mas os mecanismos de ação do medicamento não foram elucidados (PATON et al., 1999). Em recente estudo *in vivo*, o antiviral favipiravir (T-705) mostrou resultados promissores quando testado em hamsters infectados (DAWES et al., 2018), porém sem dados consistentes em humanos.

##### 2.1.5.1 Imunização Passiva

#### 2.1.5.1.1 m102.4

Um anticorpo monoclonal humano (mAb) específico para a glicoproteína G de henipavírus foi isolado a partir de uma biblioteca de anticorpos. Este mAb, m102.4, é específico para um epítipo no sítio de ligação dos receptores Efrina-B2 e Efrina-B3 de G e mostrou forte atividade neutralizadora de reação cruzada contra HeV e NiV (ZHU et al., 2008). O m102.4 é capaz de prevenir a doença em furões e primatas não humanos inoculados com uma dose uniformemente letal de cepas nipavirais em vários momentos após a exposição (BOSSART et al., 2009; GEISBERT et al., 2014; MIRE et al., 2016). Esse tipo de imunização passiva, no entanto, seria mais adequado como tratamento pós-exposição em situações futuras de surto (MIRE et al., 2016).

#### 2.1.5.2 Imunização ativa

##### 2.1.5.2.1 Vacinas de subunidade

Uma vacina de subunidade (que utiliza somente fragmentos antigênicos de um microrganismo) baseada em uma forma solúvel e oligomérica recombinante da proteína G de hendra vírus capaz de proteger completamente os animais contra NiV sem sinais de doença clínica ou evidência de replicação ou patologia viral foi desenvolvida. Gatos e furões vacinados desenvolveram altos níveis de IgG e anticorpos naturais (nAbs) de reação cruzada (BOSSART et al., 2012; PALLISTER et al., 2013). Essa vacina foi posteriormente aprovada como uma vacina para uso em cavalos (Equivac<sup>®</sup>) na Austrália (BRODER et al., 2013). No entanto, outro estudo mostrou que porcos vacinados desenvolveram respostas de reação cruzada baixas ao NiV após teste, indicando que uma resposta de imunidade mediada por células também é necessária para conferir proteção (PICKERING et al., 2016). Foi desenvolvida uma vacina candidata à vacina baseada na proteína G recombinante ortóloga que mostrou conferir proteção completa contra NiV em gatos (MUNGALL et al., 2006). Outros modelos de vacina foram desenvolvidos, mas nenhum foi aprovado para uso em humanos (BRODER et al., 2013; SATTERFIELD; DAWES; MILLIGAN, 2016a; WALPITA et al., 2017).

##### 2.1.5.2.2 Vacinas vetorizadas

###### 2.1.5.2.2.1 Vetores de poxvírus

Um dos primeiros modelos de vacinação descritos para NiV foi realizado num modelo de hamster utilizando vírus atenuados expressando a proteína F recombinante e/ou a proteína G. Quando as glicoproteínas foram expressas individualmente ou em combinação, as vacinas recombinantes foram capazes de conferir proteção completa diante da contaminação por NiV.

Isso também ocorreu após a transferência passiva de anticorpos de hamsters vacinados com altos títulos de nAb para outros hamsters (GUILLAUME et al., 2004). Também foi desenvolvido um vetor de vacinas *canarypox* (ALVAC) que expressa F, G ou ambas as glicoproteínas, e tem sido usado para imunizar porcos em um modelo experimental sem disseminação viral ou histopatologia sendo observada em animais vacinados. A co-expressão de F e G produziu as respostas mais altas de nAb e, embora as respostas de nAb à vacina de F tenham sido baixas, todos os porcos vacinados foram protegidos contra a infecção (WEINGARTL et al., 2006).

#### 2.1.5.2.2.2 Vetores do Vírus da Estomatite Vesicular (VSV)

As plataformas do vetor VSV foram utilizadas em várias formas como vacinas candidatas. Isto inclui a associação com as proteínas F e G de NiV (DEBUYSSCHER et al., 2014; MIRE et al., 2013). Essas vacinas foram desenvolvidas usando as cepas NiV-M e NiV-B, e a administração de dose única foi capaz de proteger completamente hamsters sírios (*Mesocricetus auratus*) e furões infectados com NiV. Altas concentrações de nAb, RNA/antígenos virais em baixa e ausência de patologia foram registrados em animais vacinados com os vetores que expressam NiV F ou G (DEBUYSSCHER et al., 2014; MIRE et al., 2013). Curiosamente, em hamsters vacinados com VSV que expressam apenas NiV-N uma proteção parcial foi conferida, sugerindo um possível papel para ambas as respostas de anticorpos celulares e não neutralizantes na proteção contra a doença (DEBUYSSCHER et al., 2014).

#### 2.1.5.2.2.3 Vetores de paramixovírus

Uma vacina recombinante do vírus do sarampo que expressa as glicoproteínas do NiV mostrou proteger completamente hamsters contra infecção de NiV. No entanto, os animais vacinados com duas doses da mesma vacina estavam apenas parcialmente protegidos exibindo sinais clínicos de doença neurológica (YONEDA et al., 2013). Outra plataforma vetorizada de paramixovírus usando uma estirpe de vírus da doença de Newcastle recombinante que expressa as proteínas G ou F foi também descrita. Nesses estudos, a imunogenicidade foi testada em camundongos e suínos, onde respostas nAb e respostas de células T CD8 + foram induzidas quando usadas isoladamente ou em combinação (KONG et al., 2012).

#### 2.1.5.2.2.4 Outros vetores virais

Mostrou-se que as partículas de replicação do vírus da encefalite equina venezuelana que codificam as proteínas F ou G de NiV, como candidatas a vacina, induzem respostas nAb



altamente potentes em murganhos após 3 doses (DEFANG et al., 2010). Um vetor adeno-associado também foi descrito e testado num modelo de hamster. Uma dose única desta vacina foi suficiente para proteger os hamsters contra doses letais de NiV e induziu potentes respostas de nAb (PLOQUIN et al., 2013).

#### 2.1.5.2.3 Partículas semelhantes a vírus

As proteínas F, G e M de NiV foram simultaneamente expressas em células de mamíferos para criar com sucesso partículas semelhantes a vírus (VLP) (VERA-VELASCO et al., 2018). Estes foram testados em um modelo de camundongo e mostraram induzir uma resposta nAb, e proteger hamsters vacinados após um regime de 3 doses, bem como com apenas uma dose (WALPITA et al., 2017).

Embora, atualmente, haja diversidade de modelos explorados para buscar a profilaxia de NiV, não há vacinas candidatas que tenham avançado para testes clínicos em seres humanos, em parte devido à natureza altamente complicada dos estudos de eficácia de fase 3 em doenças relativamente raras como o Nipah. Um candidato ideal forneceria bons níveis de proteção após uma vacinação de dose única e, portanto, seria econômico e adequado em situações de emergência. O sucesso da vacina licenciada Equivac HeV em cavalos é um exemplo altamente relevante de como isso pode ser alcançado.

#### 2.1.6 Importância na saúde pública e potencial de pandemia

Atualmente, Nipah é um agente viral de estágio III: pode se espalhar para os seres humanos e causar surtos limitados de transmissão (WOLFE; DUNAVAN; DIAMOND, 2007). Foi sugerido que os estágios zoonóticos são melhor compreendidos como aumentos progressivos no número reprodutivo básico ( $R_0$ ) do agente infeccioso (LLOYD-SMITH et al., 2009). A transição de uma zoonose no estágio III para um estágio IV ocorre quando o  $R_0$  de um patógeno muda de  $<1$  para  $>1$ . A infecção pelo vírus Nipah possui  $R_0$  médio de 0,48. No entanto, com a emergência de novas cepas virais que conduzam a  $R_0 > 1$ , em nosso mundo globalmente conectado, a humanidade poderá enfrentar uma pandemia devastadora (LUBY et al., 2009).

O vírus do sarampo, da família Paramyxoviridae, que foi uma das principais causas de mortalidade humana por séculos, aparentemente evoluiu entre os séculos XI e XII a partir da mutação do progenitor da peste bovina, um paramixovírus cujos hospedeiros são ruminantes (FURUSE; SUZUKI; OSHITANI, 2010). À medida que os humanos domesticavam o gado, cepas variantes desse paramixovírus associado a ruminantes desenvolveram  $R_0 > 1$  em humanos

e mataram dezenas de milhões de pessoas nos séculos subsequentes. No mesmo sentido, existe heterogeneidade genética substancial associada a alterações fenotípicas entre as cepas do vírus Nipah observadas em isolados humanos (LO et al., 2012).

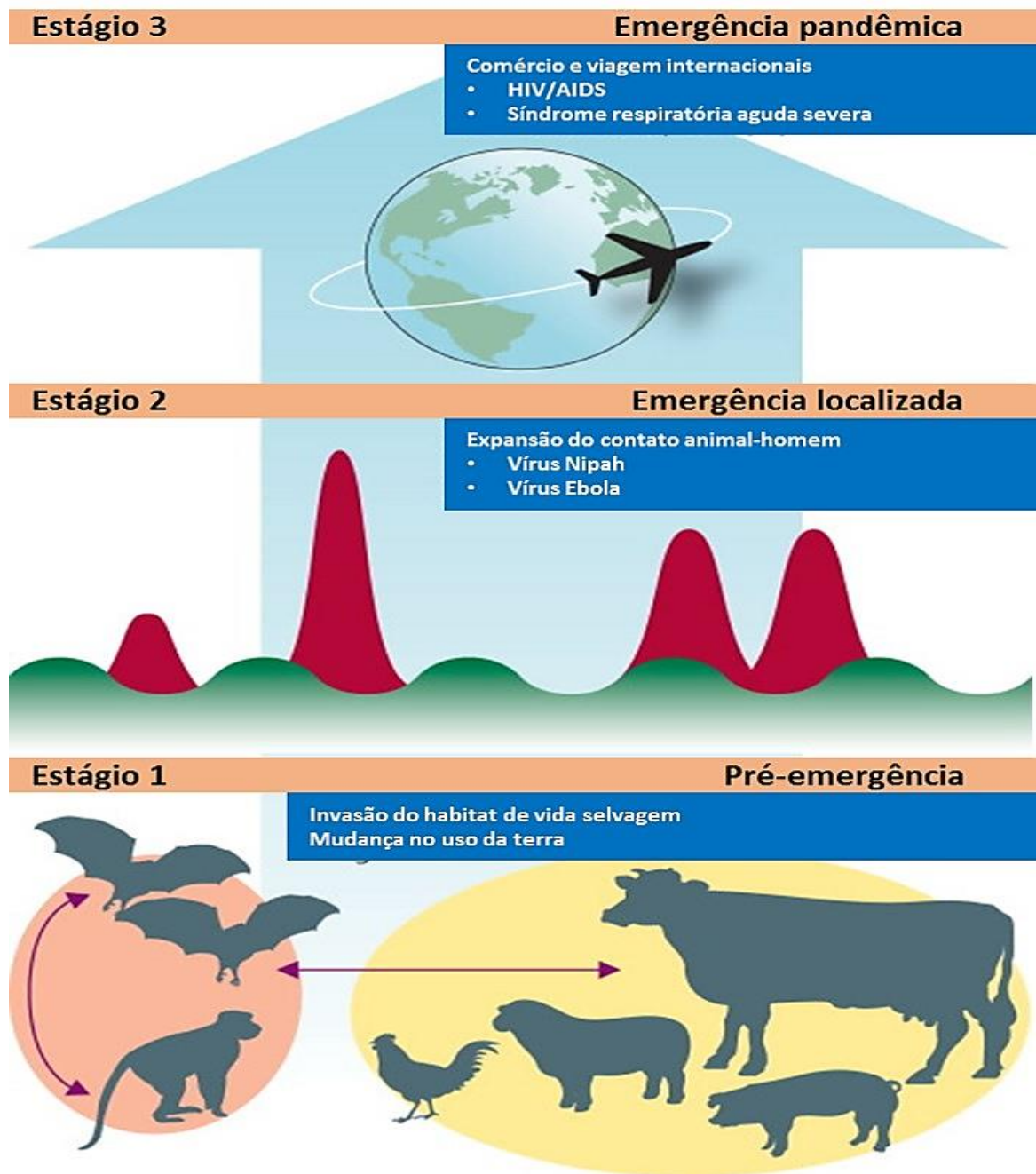
É perfeitamente concebível, portanto, que em um futuro surto possa haver um valor de  $R_0 > 1$ , que conduza a uma transmissão mais eficiente e sustentada. Embora a infecção humana não faça parte do ciclo de vida viral e não seja necessária para sua sobrevivência e propagação contínuas nos morcegos, quando as mudanças no ambiente proporcionam a um agente infeccioso novas oportunidades para atingir uma faixa ampliada de hospedeiros, a seleção natural favorece os agentes que produzem mais variantes, como os vírus de RNA (LUBY, 2013).

De acordo com Morse et al. (2012), a formação de uma pandemia segue uma dinâmica de infecção e não as propriedades do patógeno, em três fases progressivas. Mudanças ecológicas, sociais ou socioeconômicas (por exemplo, mudança no uso da terra) alteram a dinâmica da transmissão de patógenos no hospedeiro ou entre hospedeiros e permitem que o patógeno se expanda na população hospedeira, se espalhe para uma nova região, ou seja, transmitido para outra população ou espécie. Cada uma dessas mudanças aumenta a probabilidade de o patógeno entrar em contato e se espalhar para os seres humanos (progredir para o estágio 2) (KARESH et al., 2012; MCARTHUR, 2019).

Os fatores que causam o surgimento do estágio 1 tendem a ser ambientais, agrícolas ou mudanças demográficas, por exemplo, mudança de gado para uma região pela primeira vez ou transporte de animais selvagens para outros locais. Tais eventos ocorreram antes do surgimento do vírus Nipah na Malásia em 1997, após a criação intensiva de fazendas de suínos e pomares de frutas em uma região frequentada por morcegos. Esses morcegos começaram a se alimentar de árvores frutíferas ao redor dos porcos, permitindo a transmissão viral a porcos e a emergência do estágio 1 (PULLIAM et al., 2012). Da mesma forma, a mudança no uso da terra na África trouxe gado, pessoas e vida selvagem para o mesmo habitat, levando à transmissão multidirecional de patógenos entre animais e primatas não humanos (estágio 1), e algumas repercussões nas pessoas (estágio 2) (GOLDBERG, 2008).

O estágio 3 (emergência pandêmica total), por sua vez, é caracterizado pela transmissão sustentada de pessoa a pessoa e disseminação em larga escala, frequentemente auxiliada por viagens aéreas globais (por exemplo, HIV/AIDS, síndrome respiratória aguda grave, COVID-19) ou pelo movimento internacional de hospedeiros ou vetores de reservatórios através do comércio (por exemplo, vírus do Nilo Ocidental) (KARESH et al., 2012; MCARTHUR, 2019).

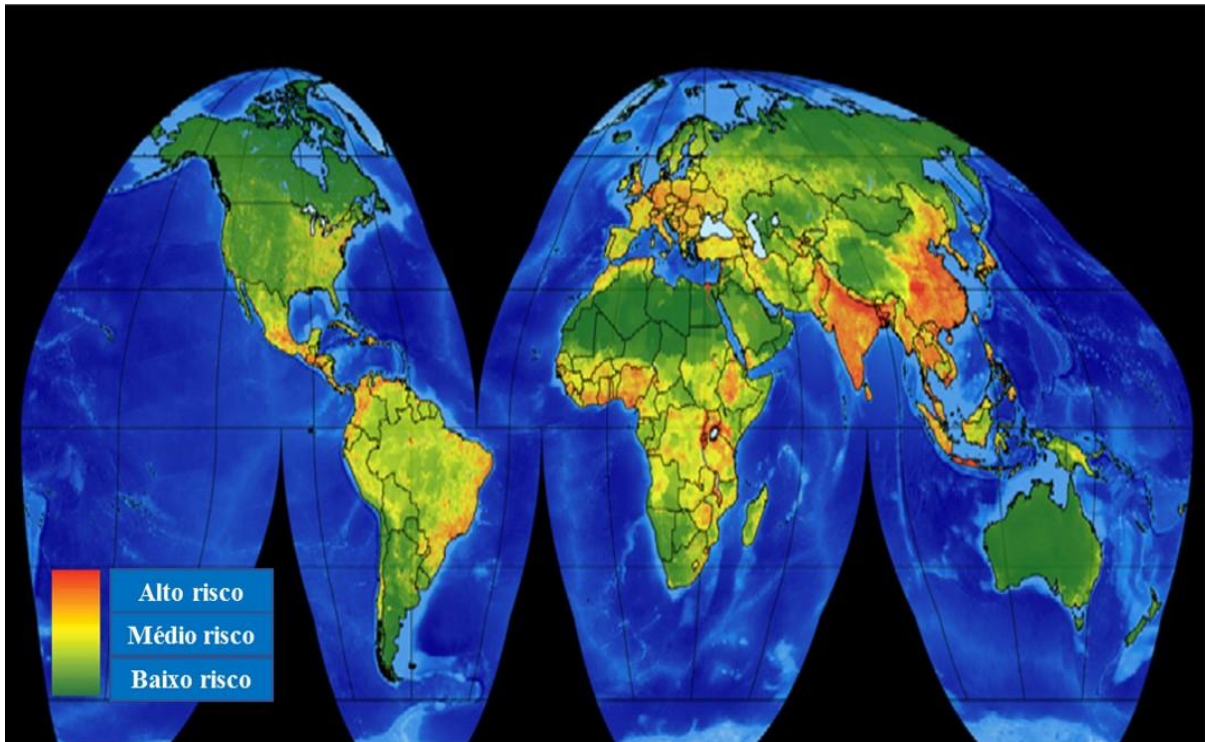
**Figura 4** - Pontos de acesso globais para doenças infecciosas emergentes que se originam na vida selvagem.



**Fonte:** Adaptado de Morse et al., 2012.

Ainda, um banco de dados de todas as doenças infecciosas emergentes conhecidas desde 1940, assim como a distribuição global das mesmas, fornece uma medida da probabilidade de uma região gerar uma nova doença infecciosa emergente zoonótica. No presente caso, a região de surtos do vírus Nipah é considerada de alto risco (Figura 5) (JONES et al., 2008; WIETHOELTER et al., 2015).

**Figura 5** - Emergência de doença zoonótica pandêmica.



**Fonte:** Adaptado de Morse et al., 2012.

Devido à a indisponibilidade de tratamentos efetivos e do potencial risco de pandemia, portanto, existe uma indispensável necessidade de desenvolvimento de um regime eficaz contra doenças infecciosas, que são causa grave de mortalidade no mundo todo. A era pós-genômica favorece a progressão de abordagens computacionais para dar suporte à vacinação e ao combate das doenças infecciosas (BACKERT; KOHLBACHER, 2015).

## 2.2 Visão geral do sistema imune

O sistema de defesa consiste em uma ampla variedade de células e moléculas que protegem os animais contra microrganismos patogênicos invasores e câncer. Reconhecimento e resposta são as duas atividades principais do sistema imunológico, que atua tanto de maneira inespecífica como específica. Além disso, é capaz de discriminar entre moléculas estranhas e as próprias células e proteínas do corpo (HOWELL; SHEPHERD, 2018). Após o reconhecimento de um organismo estranho, é desencadeada uma resposta efetora através do recrutamento de uma variedade de células e moléculas para eliminar o organismo invasor. A exposição posterior ao mesmo organismo estranho induz uma resposta de memória, caracterizada por uma reação imune mais rápida e elevada que serve para eliminar o patógeno e prevenir a doença. Existem duas subdivisões no sistema imunológico: imunidade inata e

imunidade adaptativa. Ambas diferem no tempo e duração da resposta, tipos de células efetoras e sua especificidade para diferentes classes de microrganismos exógenos. O sistema imune inato representa uma resposta não específica a uma partícula estranha potencialmente perigosa. O sistema imune adaptativo exibe um alto grau de memória e especificidade (HOWELL; SHEPHERD, 2018).

### 2.2.1 Imunidade Inata

A imunidade inata é um sistema de defesa presente em plantas, fungos, insetos e demais organismos multicelulares e representa a primeira linha de defesa contra um patógeno invasor (MAGRATH, 2013). Tal sistema é composto por quatro tipos de barreiras defensivas, notavelmente anatômicas (por exemplo, pele e membranas mucosas), fisiológicas (por exemplo, temperatura, pH baixo), fagocíticas (por exemplo, monócitos, neutrófilos, macrófagos teciduais) e inflamatórias, como as proteínas séricas. Células do sistema imune inato: fagócitos, neutrófilos, macrófagos, células NK (*natural killer*), mastócitos, basófilos, células dendríticas, eosinófilos. A resposta desenvolvida é, portanto, rápida, inespecífica e incapaz de induzir memória imunológica (MAGRATH, 2013). Respostas inatas são desencadeadas por padrões moleculares associados a patógenos, que sinalizam através de receptores de reconhecimento de padrões (PRR), nas células respondedoras para iniciar uma cascata de eventos para estimular respostas mediadas por anticorpos e células a fim de eliminar o patógeno (KAUR; SECORD, 2019; KUMAR; KAWAI; AKIRA, 2011).

### 2.2.2 Imunidade Adaptativa

O sistema imune adaptativo é ativado pela imunidade inata. Os componentes desse possuem uma dinâmica temporal mais lenta, com alto grau de especificidade e uma resposta secundária mais potente. Constitui-se de dois tipos: humoral (sistema mediado por anticorpos) e sistema mediado por células (VANGURI, 2014). Suas principais células são os linfócitos T e B. O sistema imune adaptativo frequentemente incorpora células e moléculas do sistema inato no combate a corpos estranhos nocivos. Por exemplo, o sistema complemento (moléculas do sistema inato) pode ser ativado por anticorpos (moléculas do sistema adaptativo) (VANGURI, 2014).

#### 2.2.2.1 Sistema Imunológico Humoral

Esse tipo de imunidade é mediado por macromoléculas encontradas em fluidos extracelulares, como anticorpos secretados, proteínas do complemento e certos peptídeos antimicrobianos.

#### 2.2.2.1.1 Sistema complemento

Sistema composto por proteínas da membrana plasmática e solúveis no sangue. Assim chamado porque ajuda ou “complementa” a capacidade dos anticorpos e das células fagocitárias de remover patógenos de um organismo. A ativação desse sistema leva à citólise, quimiotaxia, opsonização, liberação imunológica e inflamação. Três vias bioquímicas ativam o sistema do complemento: a via clássica do complemento, a via alternativa do complemento e a via da lectina ligadora de manose (MAGRATH, 2013).

#### 2.2.2.1.2 Células B

Pertencem a um grupo de glóbulos brancos conhecidos como linfócitos. Nos mamíferos, células B imaturas são formadas na medula óssea (VANGURI, 2014). Há um rearranjo gênico aleatório durante a maturação das células B na medula óssea que gera mais de mil células B com diferentes especificidades antigênicas. Posteriormente, há um processo de seleção para eliminar quaisquer células B com anticorpos de base que reconheçam os componentes próprios do organismo. Isso garante que os anticorpos auto-reativos (autoanticorpos) não sejam produzidos (VANGURI, 2014).

#### 2.2.2.1.3 Hipermutação somática

Quando uma célula B reconhece um antígeno, inicia-se um processo maciço de proliferação. Durante tal ato, o locus do receptor de células B sofre mutação somática nas regiões hipervariáveis, em ritmo até cem vezes maior que a taxa normal de mutação em todo o genoma (LI et al., 2004). A hipermutação aumenta a capacidade dos receptores de imunoglobulina presentes nas células B reconhecerem e se ligarem a um antígeno específico (HOWELL; SHEPHERD, 2018).

#### 2.2.2.1.4 Anticorpos

A produção de anticorpos é a principal função do sistema imunológico humoral (KARUPIAH; CHAUDHRI, 2004). Os anticorpos são moléculas heterodiméricas, que se dobram para formar uma cavidade tridimensional de ligação ao antígeno. O sítio de ligação é formado pela combinação de regiões determinantes de complementaridade hiper variáveis

(CDRs) contribuídas por cadeias leves e pesadas da molécula de anticorpo. As cadeias laterais de aminoácidos no CDR contatam porções carregadas no antígeno para formar uma estrutura de ajuste complementar do antígeno com o anticorpo. As sequências de aminoácidos que participam da interação formam epítipo no antígeno e um parátipo no anticorpo (BLUM; WEARSCH; CRESSWELL, 2013). Anticorpos ou imunoglobulinas (Ig) são secretados por células plasmáticas e reconhecem uma parte única do organismo estranho, chamada antígeno (LITMAN; CANNON; DISHAW, 2005; MAGRATH, 2013). O anticorpo tem uma ponta estruturada em “Y” para um epítipo específico, conhecido como parátipo.

#### 2.2.2.2 Sistema Imunológico Mediado por Células

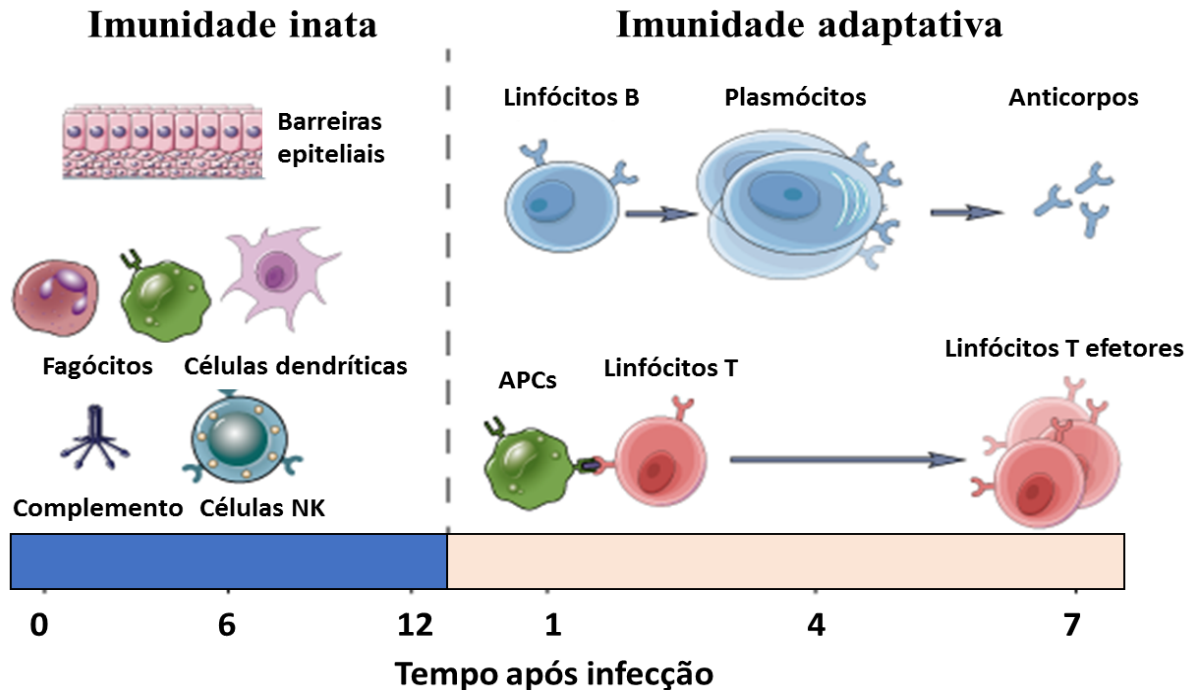
Não envolve anticorpos, mas ativa os fagócitos e os linfócitos T citotóxicos específicos para o antígeno e libera várias citocinas em resposta a um ataque de antígeno.

##### 2.2.2.2.1 Linfócitos T

Surgem na medula óssea e migram para o timo para amadurecer, ao contrário das células B, que migram para o baço e outros órgãos linfóides secundários (KINDT; GOLDSBY; OSBORNE, 2008). Dentro do timo, expressam uma molécula única de ligação ao antígeno em sua membrana, chamada de receptor de célula T (TCR). Os TCRs podem reconhecer apenas o antígeno que está ligado às proteínas da membrana celular, chamadas moléculas do complexo principal de histocompatibilidade, diferentemente das células B. Existem duas subpopulações bem definidas de células T: T auxiliaadoras e T citotóxicas. O linfócito se torna uma célula efetora (ativada) que secreta vários fatores de crescimento conhecidos coletivamente como citocinas, depois que uma célula Th reconhece e interage com um complexo de moléculas do MHC classe II. As citocinas secretadas desempenham um papel importante na ativação de células B, células Tc, macrófagos e várias outras células que participam da resposta imune. Sob a influência de citocinas derivadas de Th, uma célula Tc reconhece um antígeno de classe I do MHC, prolifera e se diferencia em um a linfócito T citotóxico (CTL) (KINDT; GOLDSBY; OSBORNE, 2008), que geralmente não secreta citocinas. CTL tem uma função vital na eliminação de células antigênicas, tais como células infectadas por vírus e células tumorais. A maturação das células T também inclui rearranjos aleatórios de uma série de segmentos gênicos que codificam o receptor de ligação ao antígeno da célula, como a maturação das células B. O rearranjo aleatório dos genes do TCR é capaz de gerar especificidades antigênicas únicas. Cada célula de linfócito T expressa inúmeros receptores, cuja descendência clonal têm especificidade idêntica para o antígeno. No entanto, essa diversidade é posteriormente reduzida através de um

processo de seleção para garantir que apenas células T com receptores capazes de reconhecer antígenos associados a moléculas do complexo principal de histocompatibilidade sejam capazes de amadurecer (KINDT; GOLDSBY; OSBORNE, 2008).

**Figura 6-** Linha do tempo representativa da imunidade inata e adaptativa. Os mecanismos da imunidade inata fornecem a defesa inicial contra infecções. As respostas imunes adaptativas se desenvolvem mais tarde e requerem a ativação dos linfócitos. A cinética das respostas imunes inatas e adaptativas são aproximações e podem variar em diferentes infecções.



**Fonte:** Adaptado de Abbas et al. (2012).

#### 2.2.2.2.2 Complexo principal de histocompatibilidade

Também chamado de antígeno leucocitário humano (HLA), o MHC constitui-se de um grande complexo genético subdividido em três classes principais de glicoproteínas ligadas à membrana: moléculas de MHC classe I, classe II e classe III. Estas moléculas não possuem especificidade para antígenos; em vez disso, ligam-se a um espectro de peptídeos antigênicos derivados da degradação intracelular de moléculas de antígeno. Tanto na classe I como na classe II, as moléculas de MHC possuem regiões variáveis, um sítio no qual o peptídeo antigênico se liga e é apresentado aos linfócitos T. As células Th geralmente reconhecem o antígeno combinado com moléculas de classe II, enquanto as células Tc geralmente reconhecem o antígeno combinado com moléculas de classe I (RAJALINGAM et al., 2015). As principais diferenças entre essas três classes são: (1) Os genes do MHC da classe I codificam glicoproteínas



expressas na superfície de quase todas as células nucleadas; a principal função dos produtos do gene de classe I é a apresentação de antígenos peptídicos para células T CD8+.

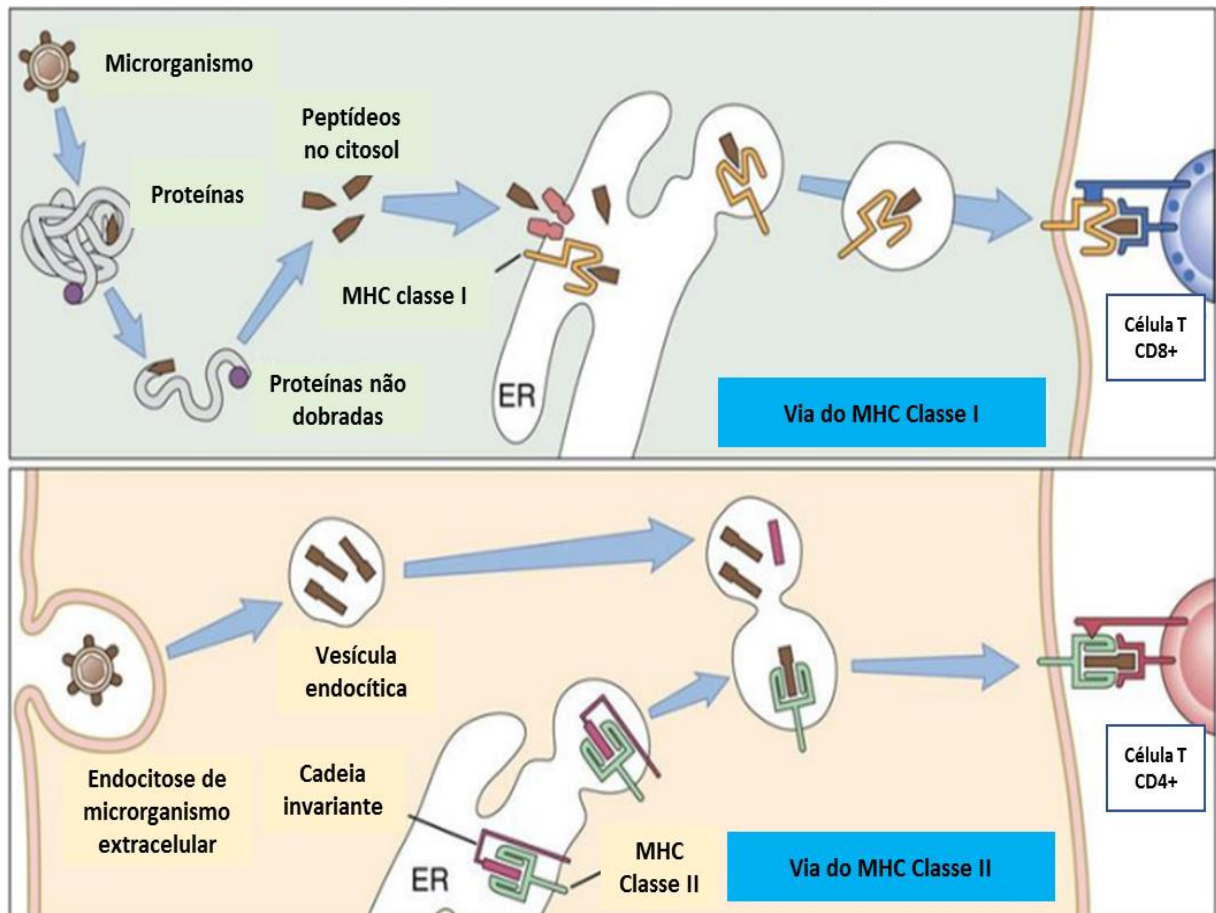
(2) Genes de MHC Classe II codificam glicoproteínas expressas principalmente em células que apresentam antígenos (macrófagos, células dendríticas e células B), onde apresentam peptídeos antigênicos processados para células T CD4+.

(3) Os genes do MHC de Classe III codificam várias proteínas relacionadas ao sistema imune, incluindo componentes do sistema complemento e moléculas envolvidas na inflamação (DEAKIN et al., 2006). Outro aspecto importante está relacionado as suas características estruturais, onde as moléculas de MHC classe I e classe II têm características estruturais comuns e ambas têm papéis no processamento de antígenos. No entanto, a região do MHC de classe III codifica moléculas que têm pouco em comum com as moléculas de classe I ou II (RAJALINGAM et al., 2015). Enquanto as moléculas de MHC de classe I são encontradas na superfície das células nucleadas, as moléculas de MHC de classe II estão presentes apenas nas células e linfócitos que apresentam antígenos.

A via de processamento de antígenos pelo MHC classe I geralmente ocorre em quatro etapas principais: 1) geração de peptídeos antigênicos por hidrólise do antígeno protéico por proteossomo constitutivo ou imunoproteossomo, 2) transporte de peptídeos do citosol para o retículo endoplasmático (ER) pelo transportador associado à proteína de processamento de antígeno (TAP), 3) ligação de peptídeos a moléculas de classe I do antígeno leucocitário humano (HLA), 4) transporte do complexo peptídeo-MHC através dos complexos de Golgi para a superfície celular para apresentação às células T CD8+ (Figura 6) (ABBAS, A; LICHTMAN, A.; PILLAI, 2012).

A via de processamento de antígenos pelo MHC classe II para apresentação a células T CD4+ envolve a fragmentação, por proteases, de proteínas exógenas, incorporadas por endocitose, em um endossomo. As cadeias alfa e beta do MHC classe II, junto com uma cadeia invariante, são sintetizadas, montadas no retículo endoplasmático e transportadas através do aparelho de Golgi e trans-Golgi para chegar no endossomo, onde a cadeia invariante é digerida, e os fragmentos de peptídios da proteína exógena são capazes de se associar com moléculas de MHC classe II, que finalmente são transportadas para a superfície da célula. Esta é a principal via para antígenos exógenos (Figura 6) (ABBAS, A; LICHTMAN, A.; PILLAI, 2012).

**Figura 7** - Principais vias de processamento de antígenos (MHC-I e MHC-II).



Fonte: Adaptado de Abbas et al. (2012).

### 2.3 Planejamento *in silico* de vacinas

As vacinas têm sido uma das melhores estratégias intervencionistas para o controle de doenças infecciosas (PLOTKIN; PLOTKIN, 2012). Desde a primeira demonstração de proteção contra a varíola, numerosas vacinas foram desenvolvidas e um grande número de outros candidatos foi testado experimentalmente em modelos animais e em espécies-alvo.

O planejamento de vacinas pode ser dividido em duas grandes categorias: a abordagem tradicional e a moderna. As formas tradicionais de desenvolver vacinas têm sido produzir em massa os patógenos, inativá-los por meios físicos ou químicos e purificar e formular o antígeno (RAPPUOLI, 2007). Em muitos casos, esta provou ser uma estratégia muito útil, especialmente na erradicação ou eliminação de algumas doenças infecciosas em nível regional. No entanto, vacinas vivas atenuadas geralmente não são ideais devido à natureza viva do antígeno e aos riscos potenciais para seu uso em indivíduos em que o sistema imunológico está

subdesenvolvido (indivíduos pediátricos), em declínio (indivíduos geriátricos) ou suprimido (por exemplo, pacientes transplantados, ou devido a agentes infecciosos, como o vírus da imunodeficiência humana, e em gestantes) (PLOTKIN; PLOTKIN, 2012).

Adicionalmente, o planejamento de vacinas tradicionais é caro, demorado e pouco eficiente para patógenos que possuem alta diversidade antigênica (WOLF et al., 2010). Isso ocorre devido à diversidade genética, informações insuficientes sobre a interação entre patógeno e hospedeiro, ausência de uma linha celular permissiva e falta de modelos animais adequados (ARORA; MISRA; SAJID, 2017). Observou-se que o desenvolvimento de vacinas para doenças graves, tais como o vírus da imunodeficiência humana (HIV) e tuberculose também foi afetado por esses inconvenientes (THOMAS; DILBAROVA; RAPPUOLI, 2016). Devido às limitações da tecnologia convencional, surgiram tecnologias modernas, incluindo tecnologia de DNA recombinante, vacinologia reversa, biologia estrutural, vacinas conjugadas, tecnologia de próxima geração e desenho de vacinas baseadas em epítopos (THOMAS; DILBAROVA; RAPPUOLI, 2016).

Por outro lado, o sistema imunológico humano é muito complexo e opera em múltiplos níveis, isto é, moléculas, células, órgãos e organismos. Cada indivíduo tem um sistema imunológico único e responderá de maneira diferente aos desafios imunológicos (GARDY et al., 2009). Considerando a combinação de estruturas biológicas e processos dentro de um organismo, grande quantidade de dados relacionados ao sistema imunológico está sendo gerada.

### 2.3.1 Planejamento de vacinas multi-epítipo

Um epítipo antigênico é uma unidade básica que provoca uma resposta imunitária celular ou humoral. Logo, uma vacina multi-epítipo é composta de uma série de peptídeos sobrepostos e é, portanto, uma abordagem ideal para a prevenção e tratamento de tumores ou infecções virais complexas (KUO et al., 2014; LU et al., 2017b). Tais modelos podem, ainda, ser acrescidos de um adjuvante para potencializar a resposta imune contra um patógeno específico (AGUILAR; RODRÍGUEZ, 2007).

Embora algumas vacinas multi-epítipo tenham entrado em ensaios clínicos fase I, por exemplo, EMD640744 em pacientes com tumores sólidos avançados (LENNERZ et al., 2014), o desenvolvimento de vacinas multi-epítipo eficazes continua a ser um grande desafio. Uma vacina multi-epítipo ideal deve ser projetada para incluir epítopos que possam induzir células T auxiliares e citotóxicas e células B para estimular respostas eficazes contra um tumor ou vírus alvo (HE et al., 2018; LENNERZ et al., 2014; LU et al., 2017).

Em comparação com vacinas clássicas, os modelos multi-epítipo têm conceitos de design com as seguintes propriedades: (I) consistem em múltiplos epítipos restritos a MHC que podem ser reconhecidos por TCR de múltiplos clones de vários subconjuntos; (II) consistem em epítipos de CTL, HTL e linfócitos B que podem induzir simultaneamente respostas imunitárias celulares e humorais fortes; (III) consistem em múltiplos epítipos de diferentes antígenos de tumor ou vírus que podem expandir os espectros de tumores ou vírus alvo; (IV) eles introduzem alguns componentes com capacidade adjuvante que podem aumentar a imunogenicidade e as respostas imunológicas de longa duração; e (V) reduzem componentes indesejados que podem desencadear respostas imunológicas patológicas ou efeitos adversos (HE et al., 2018; LENNERZ et al., 2014; LU et al., 2017b; SAADI; KARKHAH; NOURI, 2017). Vacinas multi-epítipo bem projetadas com tais vantagens devem se tornar agentes profiláticos e terapêuticos poderosos contra tumores e infecções virais.

### 2.3.2 Imunoinformática

Os imunologistas vêm empregando técnicas experimentais de alta produtividade o que tem gerado uma grande quantidade de dados funcionais, clínicos e epidemiológicos. Portanto, o desenvolvimento de abordagens computacionais para armazenar e analisar esses dados é uma solução viável e vantajosa. Isso deu origem ao campo denominado imunoinformática, que consiste na utilização de ferramentas computacionais para fornecer *insights* e investigar questões relacionadas às respostas imunes. Imunogenômica, imunoproteômica, predição de epítipos e vacinologia reversa *in silico* são diferentes abordagens da pesquisa imunológica computacional.

Nesse sentido, nas duas décadas anteriores, juntamente com a descoberta de medicamentos baseados em peptídeos, várias ferramentas foram projetadas para auxiliar no planejamento de imunoterapias. O início da análise de sequências e da tecnologia de DNA recombinante (RDT) abriram o caminho para o desenvolvimento de metodologias como o desenho de vacinas baseado em epítipos através da análise genômica de patógenos (BACKERT; KOHLBACHER, 2015). Tal estratégia foi pioneira em *Neisseria meningitidis*, responsável pela sepse e meningite meningocócica, com a produção de vacina conjugada com base em polissacarídeos capsulares. Estas vacinas estão disponíveis para as cepas patogênicas C, Y e W135 (PIZZA et al., 2000).

De modo geral, a imunoinformática revolucionou as abordagens racionais para o *design* de vacinas. Métodos computacionais desenvolvidos e refinados com base em imunologia experimental e grandes conjuntos de dados economizam, de maneira exponencial, tempo e

custos (MOISE; MOSS; DE GROOT, 2012; VAISHNAV et al., 2015). Assim, a aplicação de bioinformática fornece uma abordagem objetiva, na qual um número limitado de candidatos em potencial pode ser investigado por seu envolvimento na resposta imune por meio de exploração experimental direcionada, inclusive possibilitando a descoberta de antígenos anteriormente desconhecidos e não descritos como candidatos a vacinas. Essas ferramentas são baseadas em sistemas estatísticos e de aprendizado de máquina e estão bem estabelecidas na análise e modelagem de interações moleculares (BACKERT; KOHLBACHER, 2015).

Como o sistema imunológico é inerentemente um sistema de múltiplos componentes, a modelagem de respostas imunes requer o uso de múltiplas tecnologias de modelagem computacional e alavancagem de redes curadas celulares e moleculares. As técnicas de modelagem computacional podem capturar conhecimento processual e mecanicista existente em modelos, combinar teoria com dados e descobrir novos conhecimentos por meio de análises e simulações de modelos (BACKERT; KOHLBACHER, 2015; TOMAR; DE, 2014).

O interesse no planejamento computacional de vacinas pode, ainda, beneficiar-se das abordagens de modelagem molecular, um campo de estudo que explora o comportamento de moléculas e átomos, analisando suas propriedades físico-químicas e moleculares através de algoritmos computacionais (CRAGG; GROTHAUS; NEWMAN, 2014; ŚLEDŹ; CAFLISCH, 2018). Esse campo progrediu juntamente com os avanços nos métodos espectroscópicos como a cristalografia por difração de raios X (DRX) e RMN, que permitiram a resolução de milhares de estruturas proteicas tridimensionais, fornecendo informações estruturais vitais sobre alvos moleculares de drogas (BERMAN et al., 2014). A pesquisa imunológica incorporou uma grande variedade de métodos de modelagem molecular para estudar sistemas biológicos e químicos complexos, tais como predição estrutural de proteínas e ácidos nucleicos, simulações de docagem e de dinâmica molecular (ŚLEDŹ; CAFLISCH, 2018).

Em um processo contínuo, os resultados experimentais são usados para refinar e aprimorar modelos computacionais desenvolvidos. Esse processo assegura que o modelo aborde a biologia observada (CARBO et al., 2014).

### 2.3.3 Metodologias para o planejamento de vacinas multi-epítipo

#### 2.3.3.1 Obtenção e alinhamento de sequências

Atualmente, uma grande quantidade de informações biológicas geradas em genética, biotecnologia e biologia molecular são organizadas e armazenadas com a ajuda de ferramentas de bioinformática. Esses dados podem ser efetivamente usados para extrair informações funcionais para diversos fins médicos e biológicos (CHEN; HUANG; WU, 2017).

O banco de dados do *National Center for Biotechnology Information* (NCBI) é a fonte de sequências biológicas mais utilizada, contendo os registros mais abrangentes de sequências de RNA, DNA e proteínas (SAYERS et al., 2019). Uma quantidade abrangente de informações sobre as seqüências de proteínas e suas anotações é também fornecida no *Universal Protein Resource* (Uniprot) (BATEMAN et al., 2017). Os bancos de dados UniProt incluem o *UniProt Knowledgebase* (UniProtKB), o *UniProt Archive* (UniParc) e o *UniProt Reference Clusters* (UniRef). As seqüências de proteínas são armazenadas nesse banco de dados como registros contendo várias informações celulares, moleculares, biológicas, funcionais e estruturais, além de links externos para obter mais informações (BATEMAN et al., 2017).

O *Protein Data Bank* (PDB) é outro banco de dados altamente utilizado que deposita dados sobre as estruturas tridimensionais de proteínas, ácidos nucléicos e seus complexos. As informações obtidas nesse banco de dados ajudam os pesquisadores a entender melhor vários aspectos de macromoléculas de interesse na biomedicina e na agricultura (KLEYWEGT; VELANKAR; PATWARDHAN, 2018).

Ainda, o *InnateDB* (<http://www.innatedb.ca>) é um banco de dados dedicado a facilitar investigações em nível de sistema da resposta imune inata para mamíferos (humanos, camundongos e bovinos). Foi desenvolvido para proporcionar uma percepção mais sofisticada sobre as respostas do sistema imunológico inato. Até o momento, mais de 27.172 interações curadas, 5237 publicações revisadas e 9460 genes curados são depositados no *InnateDB* (LYNN et al., 2008). Bancos de dados científicos e ferramentas de softwares também podem ser acessados pelo ExPASy (<https://www.expasy.org/>). Esta ferramenta fornece aos cientistas acesso a vários dados, incluindo genômica, transcriptômica, proteômica, filogenia, genética de populações, biologia de sistemas e etc. (ARTIMO et al., 2012a).

### 2.3.3.2 Predição de estruturas

#### 2.3.3.2.1 Estrutura secundária proteica

A predição da estrutura secundária da proteína é uma das análises mais importantes para entender as propriedades estruturais de uma proteína. O significado dessas análises se torna mais evidente ao lidar com as sequências de proteínas que não compartilham sequências comuns com proteínas de estrutura resolvida. Para prever várias características da estrutura secundária de proteínas, bem como a formação de pontes dissulfeto, diferentes métodos de predição foram desenvolvidos, como SOPMA (<https://npsa-prabi.ibcp.fr/cgi-bin/>) (FERRON et al., 2006; PIROVANO; HERINGA, 2010).

### 2.3.3.2.2 Predição e refinamento da estrutura terciária proteica

A lacuna entre o número de sequências de proteínas identificadas e o número de proteínas em suas estruturas experimentalmente resolvidas acentuam a necessidade de desenvolver métodos precisos de predição de estruturas 3D de proteínas. Além disso, obter uma compreensão realista do mecanismo de interação proteína-proteína em complexos de proteínas requer *insights* estruturais sobre a relação entre duas moléculas a nível atômico (CHOU, 2012). Embora a estrutura de alguns desses complexos seja determinada experimentalmente por cristalografia de raios-X e seus dados estejam disponíveis no PDB, a estrutura de muitos complexos conhecidos ainda não está disponível. Os métodos de predição da estrutura 3D de proteínas são úteis para preencher tais lacunas (CHAKRABORTY et al., 2017). De fato, métodos teóricos, juntamente com os avanços nos métodos laboratoriais, estão acelerando a compreensão da interação de moléculas. As principais abordagens para prever a estrutura terciária da proteína são modelagem por homologia (PEITSCH; SCHWEDE, 2011), *threading* (PENG; XU, 2011) e predição *ab initio* (LEE; FREDDOLINO; ZHANG, 2017).

A primeira abordagem prediz a estrutura com base no fato de que sequências evolutivamente semelhantes se enovelariam em estruturas semelhantes. De modo geral, quanto maior a porcentagem de resíduos idênticos nas sequências de proteínas comparadas e alinhadas, maiores as relações evolutivas. Assim, duas proteínas que compartilham apenas 25% de identidade de sequência geralmente têm o mesmo enovelamento geral (PEITSCH; SCHWEDE, 2011). Entre os métodos de modelagem, a modelagem por homologia é concedida como a mais precisa. No entanto, encontrar uma estrutura-molde adequada com alta cobertura e identidade de sequência é um desafio recorrente. Geralmente, as estruturas-molde com menos de 35% de cobertura são considerados modelos não confiáveis. Se a modelagem da homologia não for aplicável, o método *threading* é o método preferível, baseando-se na comparação de uma sequência proteica de interesse com um banco de dobramentos estruturais conhecidos. Um conjunto de pontuações geradas pelos algoritmos de análise determina o enovelamento de proteína que é assumido pela sequência de interesse (PENG; XU, 2011). *Ab initio* é outra abordagem de predição de estruturas com base em princípios físicos, que incluem a identificação do menor estado de energia livre e o cálculo de todos os parâmetros energéticos do enovelamento de proteínas (LEE; FREDDOLINO; ZHANG, 2017). De acordo com a "hipótese termodinâmica", uma estrutura proteica está em seu estado nativo quando atinge o mínimo de energia global. A abordagem *ab initio* é baseada nesta hipótese e nos fundamentos da otimização global determinística (LEE; FREDDOLINO; ZHANG, 2017). Embora essa abordagem possa gerar resultados precisos de modelagem, altos requisitos computacionais

limitam sua utilização. As ferramentas mais conhecidas têm o limite de predição de 200 aminoácidos, o que a torna imprópria para sequências de proteínas mais longas (LEE; FREDDOLINO; ZHANG, 2017).

Os modelos tridimensionais preditos geralmente contêm distorções estruturais. Portanto, tais modelos são submetidos à minimização de energia e a refinamentos completos da estrutura atômica para obter melhor qualidade tridimensional. Nesse sentido, vários algoritmos foram desenvolvidos para refinar e minimizar a energia dos modelos 3D construídos. A avaliação da qualidade dos modelos preditos antes e após as execuções de refinamento é realizada principalmente por ferramentas de avaliação da qualidade (DAPKŪNAS; OLECHNOVIČ; VENCLOVAS, 2019; ZHANG; SKOLNICK, 2007). O gráfico de Ramachandran avalia a qualidade do modelo por meio do cálculo dos ângulos de torção *phi-psi* para cada resíduo da proteína. Finalmente, os resíduos são classificados como favorecidos, permitidos e não permitidos (CARUGO; DJINOVIĆ-CARUGO, 2013). Existem outras ferramentas de avaliação da qualidade do modelo capazes de estimar a qualidade do modelo e descrever seus principais aspectos geométricos. Usando a função de pontuação composta, erros estruturais globais e locais dos modelos podem ser estimados.

#### 2.3.3.3 Predição de epítomos de células T

Para desenvolver uma vacina baseada em epítomos, o passo mais importante é a seleção de epítomos de células T com forte afinidade de ligação a moléculas apropriadas de MHC (PATRONOV; DOYTCHINOVA, 2013). A ligação molecular entre peptídeos antigênicos e moléculas de MHC é necessária para o reconhecimento pelas células T citotóxicas. Portanto, qualquer algoritmo de predição de epítomos de células T precisa identificar peptídeos de ligação ao MHC. A ligação do peptídeo ao MHC e a interação desse complexo com um receptor específico de células T são os requisitos para uma resposta funcional das células T (PATRONOV; DOYTCHINOVA, 2013). Desde a primeira identificação de um epítomo mínimo de células T, um grande número de epítomos de células T CD4 + e CD8 + foi definido. Vários bancos de dados depositaram epítomos determinados experimentalmente, possibilitando que ferramentas fossem criadas para a predição em sequências diversificadas (PATRONOV; DOYTCHINOVA, 2013).

Os peptídeos com capacidade de ligação às moléculas do MHC podem ser preditos por diferentes metodologias, incluindo o modelo oculto de Markov (HMM), redes neurais artificiais (RNA), máquina de vetores de suporte (SVM) e matrizes quantitativas (KONSTANTINOU, 2017). Os principais métodos desenvolvidos e utilizados para identificar epítomos estão



disponíveis no Banco de Dados de Epítopos Imunológicos (IEDB, do inglês *Immune Epitope Database and Analysis Resource*), agregando dados devidamente curados da literatura (VITA et al., 2015).

#### 2.3.3.4 Predição de epítopos de células B

Os epítopos de células B são classificados em lineares (contínuos) e conformacionais (descontínuos). Epítopos lineares são peptídeos curtos que são compostos de aminoácidos sequenciais no antígeno (EL-MANZALAWY; HONAVAR, 2013). No entanto, a maioria dos epítopos de células B é descontínua (90% dos epítopos de células B), onde os aminoácidos do epítopo podem não ser contíguos na sequência primária e apenas o dobramento de proteínas os leva à proximidade espacial (YAO et al., 2013). O mapeamento de epítopos de células B dentro de sequências proteicas de antígenos desempenha um papel crucial no design de vacinas, no design de imunodiagnósticos e no design de imunógenos para a produção de anticorpos (EL-MANZALAWY; HONAVAR, 2013).

As principais metodologias de predição contínua de epítopos de células B envolvem métodos baseados em sequência, métodos de aprendizado de máquina e métodos baseados na escala de propensão de aminoácidos (SINGH; ANSARI; RAGHAVA, 2013). Recentemente, há um interesse crescente no desenvolvimento de métodos de predição de epítopos de células B conformacionais. Esses métodos dependem principalmente do conhecimento sobre a estrutura 3D da proteína (MANAVALAN et al., 2018; RUBINSTEIN; MAYROSE; PUPKO, 2009). Embora avanços significativos tenham sido feitos no campo da predição de epítopos de células B, a cristalografia por difração de raios X continua sendo a maneira mais precisa de identificar os epítopos de células B. Vários bancos de dados depositaram tais epítopos determinados experimentalmente e diferentes ferramentas foram desenvolvidas para prever as sequências epitópicas de células B.

#### 2.3.3.5 Seleção de ligantes e adjuvantes

Ligantes têm sido mostrados como componentes indispensáveis no desenho de proteínas. Tal grupo pode ser utilizado para ligar diferentes epítopos ou domínios proteicos e é classificado em três categorias estruturais: flexíveis, rígidos e cliváveis. Um alto conteúdo de aminoácidos pequenos e polares, como glicina e serina, resulta em ligantes flexíveis e solúveis (CHEN; ZARO; SHEN, 2013). Os ligantes rígidos contêm vários resíduos prolina, adotando estruturas helicoidais que lhes permitem exibir estruturas relativamente rígidas. Comparados aos ligantes flexíveis, os ligantes rígidos são mais eficientes em separar os domínios funcionais

dos antígenos projetados, para manter sua estabilidade ou bioatividade. Desse modo esses ligantes poderiam potencialmente ser utilizados para melhorar a bioatividade das proteínas projetadas (CHEN; ZARO; SHEN, 2013a; CHUNG et al., 2009).

Um adjuvante pode ser definido como um componente da vacina que melhora a resposta imune do hospedeiro (TOM et al., 2019). Os adjuvantes podem ser projetados incorporando vários componentes, como agonistas dos receptores *toll-like*, que identificam o patógeno e são responsáveis por induzir imunidade inata (KANZLER et al., 2007). Muitos bancos de dados estão disponíveis para encontrar o adjuvante, como o *Vaxijen* (<http://www.violinet.org/vaxj>). Este banco de dados incorpora aproximadamente 400 vacinas que usam um adjuvante e contém mais de 100 adjuvantes de vacina.

### 2.3.3.6 Avaliação das propriedades do antígeno projetado

#### 2.3.3.6.1 Antigenicidade

A antigenicidade de vacinas ou imunógenos projetados está entre os critérios mais importantes para o design eficiente de proteínas. Uma proteína com alta pontuação de antigenicidade traria melhores resultados de imunização na fase empírica dos estudos imunológicos (VAISHNAV et al., 2015).

#### 2.3.3.6.2 Análise das propriedades físico-químicas e alergenicidade

Diferentes propriedades físico-químicas da proteína alvo, incluindo ponto isoelétrico, peso molecular, composição de aminoácidos, índice de instabilidade, meia-vida *in vitro* e *in vivo* e grande média de hidropaticidade podem ser calculadas pelo servidor ProtParam (<http://web.expasy.org/protparam/>) (GASTEIGER et al., 2009). O ponto isoelétrico calculado pode ser útil para a purificação de proteínas através do método de focalização isoelétrica, desenvolvendo sistemas tampão. O volume relativo dos aminoácidos alanina, valina, isoleucina e leucina de uma proteína é chamado índice alifático. O alto índice alifático significa a estabilidade térmica das proteínas globulares a várias temperaturas. Quanto menor o GRAVY calculado, melhor a interação com a água (GASTEIGER et al., 2009).

Outro aspecto observado é a alergenicidade. Um alérgeno é uma proteína ou glicoproteína que produz uma resposta imune anormalmente vigorosa e causa uma reação alérgica. A localização dos epítomos das células B e T no alérgeno pode não coincidir (AALBERSE; STADLER, 2006). Dessa forma, epítomos de células B que estão localizados na superfície do antígeno podem ser reconhecidos por anticorpos IgE. Além disso, é difícil prever a alergenicidade de uma molécula com base em seus epítomos conformacionais conhecidos.

Várias ferramentas de banco de dados e predição já foram desenvolvidas para analisar a reatividade cruzada de alérgenos (AALBERSE; STADLER, 2006).

### 2.3.3.7 Otimização do projeto

#### 2.3.3.7.1 Avaliação de modificações pós-traducionais (PTMs)

O PTM é o processo de modificar as proteínas sintetizadas, exercendo uma mudança geralmente enzimática em uma sequência de proteínas alvo. Os PTMs podem ocorrer em quase todas as proteínas em suas cadeias laterais de aminoácidos ou nos terminais C ou N da proteína, o que leva a efeitos significativos na estrutura da proteína, propriedades dinâmicas e atividade biológica. Várias abordagens computacionais foram desenvolvidas para prever locais de PTM na sequência de antígenos. Glicosilação, fosforilação e palmitoilação são alguns dos PTMs mais comuns com precisão na expressão de proteínas (BYRUM; TAVERNA; TACKETT, 2013; EIDHAMMER et al., 2013).

#### 2.3.3.7.2 Otimização de códons

Uma vez que a utilização de códons é diferente para os hospedeiros de expressão, deve-se levar em consideração o *design* do gene recombinante. O Índice de Adaptação de Códons (CAI), que varia de 0 a 1, é o principal fator considerado para a otimização de genes. Para otimizar a expressão da proteína recombinante, várias ferramentas de otimização de códons foram desenvolvidas. As tabelas de uso de códons já foram definidas para cada hospedeiro de expressão que pode ser selecionado para a tradução reversa da sequência e expressão ideal do modelo desenvolvido. A otimização do códon favorece a expressão de uma construção sintética respectivo vetor (CHIN; CHUNG; LEE, 2014; KAHL, 2015).

### 2.3.3.8 Simulações por mecânica molecular

#### 2.3.3.8.1 Docagem Molecular

Em linhas gerais, docagem molecular ou atracamento molecular é um dos métodos amplamente utilizados para prever o modo de ligação entre moléculas, seja de um ligante em um sítio de ligação de um receptor ou o modo de interação entre macromoléculas (DE MOLFETTA et al., 2009). A disponibilidade da estrutura tridimensional do alvo molecular é uma condição necessária; seja estrutura resolvida experimentalmente (por cristalografia de raios X ou RMN) ou uma estrutura obtida por técnicas computacionais (como modelagem por homologia) (SALMASO et al., 2018).

O método de docagem é composto, essencialmente, por dois módulos: um mecanismo de amostragem de conformações/orientações (método de busca) e uma função de pontuação, baseada em diferentes algoritmos, que associa uma pontuação a cada complexo obtido (MOAL et al., 2013). Genericamente, o processo de amostragem deve buscar efetivamente o espaço conformacional descrito pela energia livre, onde a energia, na docagem, é mensurada pela função de pontuação. A função de pontuação, portanto, deve associar a conformação de ligação nativa ao mínimo global de energia. As funções de pontuação desempenham o papel de seletor de *poses*, usado para discriminar modos de ligação corretos e incorretos no conjunto de resultados gerados pelo mecanismo de amostragem. Além disso, novas funções de pontuação foram desenvolvidas, por exemplo, com base em tecnologias de aprendizado de máquina, impressões digitais de interação e pontuações de mecânica quântica (LIONTA et al., 2014).

#### 2.3.3.8.2 Dinâmica Molecular

Dinâmica Molecular (DM) é um método que utiliza as equações de movimento de Newton para simular computacionalmente a evolução temporal de um conjunto de átomos interagindo em meio geralmente aquoso, descrevendo a variação do comportamento molecular em função do tempo (DE VIVO et al., 2016). Isto é, as simulações de DM fornecem informações sobre alterações conformacionais locais e não locais dependentes do tempo em uma molécula ou complexo (HEINECKE et al., 2015). Essas mudanças podem ser usadas para interpretar processos biológicos em nível molecular, como análise de interfaces de ligação, identificação de resíduos importantes de aminoácidos e epítomos de ligação, modelagem de energias termodinâmicas e análise de novas moléculas de importância imunológica, incluindo drogas e vacinas (HEINECKE et al., 2015). Os principais alvos das simulações de DM em proteínas do sistema imunológico estão relacionados a anticorpos e moléculas de MHC nos estados livre ou ligados a antígeno, assim como os demais receptores de células T. Isso ocorre porque anticorpos e moléculas de MHC possuem, independentemente, domínios estruturais conservados notáveis, mas ligam um número inesgotável de motivos peptídicos curtos de fontes patogênicas ilimitadas (CRESSWELL et al., 2005; NEEFJES et al., 2011)

As simulações de DM fornecem informações efetivas sobre as conformações do sistema e dos resíduos de aminoácidos no nível atômico que eventualmente auxiliam no projeto de medicamentos e vacinas (HEINECKE et al., 2015). Conformações de motivos peptídicos antigênicos também e suas interações com os receptores *toll-like* também podem ser exploradas através de estudo por simulações de DM (CHAUHAN et al., 2019; LATA et al., 2018). Além disso, as proteínas do complemento envolvidas na imunidade inata também estão recebendo

atenção devido ao fato de a ativação não regulamentada do complemento estar implicada em um grande número de doenças (BIRYUKOV; STOUTE, 2018).

A atual geração de computadores se beneficia do paralelismo e dos aceleradores para acelerar o processo de cálculo. Os códigos de simulação mais populares são compatíveis há muito tempo com a interface de transmissão de mensagens (MPI) e unidades de processamento gráfico (GPU). Quando um grande número de núcleos de computadores pode ser usado simultaneamente, distribui-se o sistema entre os processadores, estratégia conhecida como decomposição espacial.

#### 2.3.4 Exemplos de aplicações

O primeiro exemplo de uso de imunoinformática em um planejamento de vacinas foi o caso da meningite meningocócica do grupo B, que foi desenvolvida após todo o sequenciamento do genoma do patógeno *Neisseria meningitidis* e subsequente série de análises imunômicas e de estrutura-atividade (DELANY; RAPPUOLI; SEIB, 2013; GIULIANI et al., 2006; PIZZA et al., 2000; SCARSELLI et al., 2011).

Uma abordagem usando genômica e proteômica também levou à identificação de antígenos dos estreptococos do grupo A e do grupo B (NUCCITELLI et al., 2011; RODRÍGUEZ-ORTEGA et al., 2006). Genômica e proteômica seguidas de análises de secretoma também levaram à descoberta de antígenos protetores conservados contra *Chlamydia pneumoniae* (STEPHENS, 2000; THORPE et al., 2007). As descobertas computacionais também foram fundamentais no desenho de novas candidatas a vacinas. A partir de antígenos difíceis de purificar, fracos e recombinantes, foi produzida uma glicoproteína do vírus sincicial respiratório altamente estável, capaz de provocar fortes respostas funcionais de anticorpos (ENGLUND; CHU, 2017; MCLELLAN et al., 2013).

Da mesma forma, uma vacina peptídica sintética foi desenvolvida para aumentar a eficiência da apresentação de antígenos e estimular a resposta imune humoral e, assim, ajudar no tratamento da asma. Os pesquisadores descreveram que um peptídeo de interleucina-13 foi expresso e aumentado na asma, e isso poderia servir como um antígeno crucial a ser usado no rastreamento da asma por métodos experimentais e computacionais (HAYMAN; HERMANS; PAINTER, 2017).

Recentemente, foi desenvolvida a vacina multi-epítipo denominada EBV LMP2m, que codifica múltiplos epítipos CTL, HTL e B do gene EBV-LMP2 (LIN et al., 2016). A expressão da proteína recombinante desenhada foi otimizada para maximizar sua produção em *E. coli*. As reações específicas de células CTL e Th e a ativação de células B (IgG específica do soro e

anticorpos IgA da mucosa) foram analisadas em ratinhos BALB/c imunizados com a vacina multi-epítipo (LIN et al., 2016). Os anticorpos séricos específicos para múltiplos epítipos em pacientes com tumores relacionados ao EBV (carcinoma nasofaríngeo) ou infecção por EBV foram analisados usando ELISA e *western blotting*. Todas as análises mostraram que a vacina multi-epitópica EBV LMP2m tinha uma imunogenicidade muito elevada. Portanto, o EBV LMP2m pode ser considerado como um potencial candidato a vacina e agente de diagnóstico (LIN et al., 2016). Foi efetivada também a construção de uma vacina de subunidades multi-epítipo contra as doenças transmitidas por vetores como malária e visceral leishmaniose, explorando a proteína salivar dos vetores (mosquito) (PANDEY; PRAJAPATI, 2019).

O avanço dos métodos computacionais também tem levado à identificação de novos e potenciais candidatos a componentes de vacina. Evidências mostram que uma vacina contra o vírus Zika pode ser racionalmente projetada com alta eficiência, através do domínio da proteína do envelope III (EDIII) (TAI et al., 2019). No mesmo sentido, o modelo intitulado EV-A71 foi descoberto como um candidato à vacina de próxima geração contra o Enterovírus A71, responsável por causar febre aftosa e complicações neurológicas graves (YEE et al., 2019). Para combater a infecção hospitalar causada principalmente por *Pseudomonas aeruginosa* em pacientes imunocomprometidos, as proteínas PSE17-1, PSE41-5 e PSE54 foram identificados como antígenos potenciais, assim como para *Acinetobacter baumannii*, patógeno em evolução encontrado principalmente em unidades de terapia intensiva, para o qual Ren et al., (2019) desenvolveram uma vacina multi-epítipo usando ferramentas bioinformáticas e avaliaram sua resposta imune em camundongos. Eles observaram uma alta produção de anticorpos IgG que protegem contra doses letais de *Acinetobacter baumannii* (REN et al., 2019).

Em resumo, as vacinas multi-epítipo podem ser consideradas como uma estratégia promissora contra tumores e infecções virais. Os métodos de imunoinformática podem ajudar a prever e rastrear epítipos apropriados para conceber uma vacina multi-epítipo eficaz. Dada a relevância da infecção do vírus nipah e a existência de poucos estudos que exploram o proteoma viral para o desenho de proteínas multi-epítipo, as ferramentas de imunoinformática podem ser aplicadas para, potencialmente, produzir componentes para uma solução profilática viável (KSIAZEK et al., 2015).

### 3 OBJETIVOS

#### 3.1 Objetivo geral

Identificar, através da aplicação de ferramentas de imunoinformática e modelagem molecular, modelos multi-epítipo de proteínas imunogênicas baseadas no proteoma do vírus Nipah.

#### 3.2 Objetivos específicos

3.2.1 Analisar a presença e distribuição de epítopos de linfócitos B lineares e conformacionais, assim como epítopos de linfócitos T citotóxicos e linfócitos T auxiliares nas proteínas do vírus Nipah.

3.2.2 Analisar a conservação de epítopos dentro de diferentes proteomas virais obtidos em banco de dados.

3.2.3 Construir modelos de vacina multi-epítipo com propriedades físico-químicas otimizadas, potencial antigenicidade e não alergenicidade.

3.2.4 Analisar o modo de interação e a estabilidade dos complexos formados entre receptores *toll-like* e os modelos de vacina propostos.

3.2.5. Analisar o modo de ligação entre as estruturas de epítopos preditos de células T e variantes do complexo principal de histocompatibilidade de classes I e II.

3.2.6. Predizer e otimizar a expressão em sistema bacteriano das sequências de cDNA das vacinas.

## 4 MATERIAL E MÉTODOS

### 4.1 Obtenção de proteomas virais e avaliação de similaridade a proteínas humanas

A base de dados Uniprot (BATEMAN et al., 2017) foi utilizada para obtenção das sequências de aminoácidos, em formato FASTA, do proteoma de referência do vírus Nipah (ID: UP000170143) e outros 14 proteomas publicamente disponíveis do vírus (Apêndice A). Dados os proteomas obtidos, as respectivas proteínas foram analisadas com a ferramenta BLASTp (JOHNSON et al., 2008) para identificar potencial similaridade de sequências (identidade) com proteínas humanas.

### 4.2 Predição de epítomos de linfócitos T e indutores de IFN- $\gamma$

Epítomos de linfócitos T citotóxicos (CTL) e auxiliares (HTL) foram preditos e analisados utilizando-se a ferramenta TepiTool, derivada do *Immune Epitope Database and Analysis Resource* (IEDB) (PAUL et al., 2016), sendo submetidas, para análise, as sequências proteicas das proteínas virais em formato FASTA. Entre os diferentes parâmetros de predição, doze alelos representativos de supertipos imunitários do complexo principal de histocompatibilidade (MHC) classe I (SIDNEY et al., 2008) foram selecionados para predição de epítomos de 9mers de CTL e vinte e seis tipos mais frequentes/representativos na população mundial, abrangendo sete supertipos de MHC classe II (GREENBAUM et al., 2011) foram selecionados para predição de epítomos de 15 resíduos de HTL (Apêndice B). O parâmetro de predição utilizado foi o método de consenso *IEDB recommended*, com *percentile rank* de 1,0 como valor de corte para seleção de epítomos CTL e 10,0 para epítomos HTL. De modo geral, quanto mais próximo de zero for o *percentile rank* obtido, maior a probabilidade de um peptídeo ser um epítomo.

Para o desenho de peptídeos indutores de IFN- $\gamma$ , os epítomos HTL obtidos foram submetidos ao servidor IFNepitope (<http://crdd.osdd.net/raghava/ifnepitope/>), utilizando-se a metodologia de modelo híbrido entre a seleção baseada em motivos e máquina de vetores de suporte (SVM), diferenciando-se indutores de não indutores. Portanto, ao fim do processo, foram mantidos epítomos CTL com afinidade a MHC-I e epítomos HTL com afinidade a MHC-II e potencial indução de IFN- $\gamma$ .

### 4.3 Predição de imunogenicidade

A imunogenicidade dos epítomos selecionados preditos para interagir com MHC classes I e II foi ainda analisada na ferramenta de imunogenicidade disponível no IEDB (CALIS et al.,



2013; DHANDA et al., 2018). As pontuações de imunogenicidade quanto a MHC-I variam entre 0 e 1; a pontuação mais alta indica uma maior probabilidade de desencadear uma resposta imune. As pontuações para a predição de ligação ao MHC-II variam de 0 a 100, com valores baixos identificando peptídeos mais imunogênicos e valores altos de peptídeos não imunogênicos.

#### **4.4 Predição de processamento dos epítomos com afinidade a MHC classe I**

Os epítomos de CTL foram analisados usando a ferramenta de processamento IEDB MHC-I, que fornece acesso a previsões de processamento de epítomos por meio da via de apresentação de antígeno MHC de classe I, que inclui clivagem proteossomal e interação com o transportador associado ao processamento de antígenos (TAP). Essas pontuações e a predição de ligação ao MHC classe I (*percentile rank*) são combinadas no servidor para gerar uma pontuação de processamento (*processing score*). O objetivo desta predição é identificar ligantes de MHC-I, ou seja, peptídeos que são naturalmente processados a partir de suas proteínas de origem e apresentados por moléculas MHC de classe I (TENZER et al., 2005). Todas as pontuações associam valores mais altos com maior eficiência prevista.

#### **4.5 Predição de epítomos de células B**

Epítomos lineares (contínuos) de células B foram preditos por meio da ferramenta BepiPred 2.0, através da análise das sequências proteicas virais por um algoritmo floresta aleatória baseado em epítomos anotados (JESPERSEN et al., 2017a). Os resíduos foram analisados sendo considerado o limiar padrão para seleção ( $>0,5$ ).

Epítomos conformacionais (descontínuos) de linfócitos B foram preditos usando o servidor ElliPro (PONOMARENKO et al., 2008), que implementa três algoritmos realizando as seguintes tarefas: (1) aproximação da forma da proteína como um elipsóide; (2) cálculo do índice de protrusão do resíduo (IP); e (3) agrupamento de resíduos vizinhos com base em seus valores de IP.

#### **4.6 Filtragem dos epítomos**

Foram utilizados três critérios de filtragem para selecionar os epítomos virais para o desenho da proteína: (1) imunogenicidade e processamento previstos, (2) conservação estrutural entre as cepas virais e (3) toxicidade potencial. A conservação estrutural de epítomos entre as diferentes sequências analisadas foi verificada usando a ferramenta de análise de conservação do epítomos do IEDB (BUI et al., 2007). Adicionalmente, os resultados obtidos

foram comparados através de alinhamento múltiplo usando MUSCLE (EDGAR, 2004) e regiões com as maiores sobreposições foram analisadas, sendo retidos somente os epítomos com 100% de conservação em todas as cepas virais. Apenas epítomos conservados em todas as quinze cepas analisadas foram retidos. Para verificar a potencial natureza não tóxica dos epítomos obtidos, a ferramenta ToxinPred será utilizada com a abordagem híbrida baseada em motivos ( $e\text{-value} = 10$ ) e máquina de vetores de suporte ( $threshold = 10$ ) (GUPTA et al., 2013).

#### 4.7 Construção dos modelos de vacina

Seis modelos de proteínas multi-epítomo compostas de até 400 resíduos foram desenhados a partir das análises anteriores de B e linfócitos T epitomos. O primeiro modelo consistiu nos melhores epítomos obtidos da proteína de fusão e glicoproteína G, validados experimentalmente com atividade imunomoduladora contra infecção por NiV (BOSSART et al., 2012; DEBUYSSCHER et al., 2014; MCEACHERN et al., 2008). O segundo modelo inclui epítomos selecionados de todas as proteínas virais codificadas pelo genoma de NiV. As subunidades das proteínas de múltiplos epítomos projetadas foram iniciadas pela sequência do adjuvante  $\beta$ -defensina 4A, seguida por diferentes combinações dos epítomos virais selecionados que foram adicionados sequencialmente. Em seguida, a sequência do adjuvante  $\beta$ -defensina 103 foi adicionada à região C-terminal dos modelos projetados. Por fim, o ligante EAAAK foi alocado entre epítomos e adjuvantes e os ligantes GPGPG, AAY, e KK foram fundidos com epítomos reconhecidos por HTL, CTL e linfócitos B, respectivamente (CHEN; ZARO; SHEN, 2013).

#### 4.8 Avaliação de antigenicidade, alergenicidade dos modelos e propriedades físico-químicas

A potencial antigenicidade da sequência final das vacinas multi-epítomo foi analisada por ANTIGENpro (CHENG et al., 2005), através de um *pipeline* baseado em máquina de vetores de suporte. Para avaliar a alergenicidade dos modelos projetados foram utilizados dois servidores: SDAP, que efetiva análise de alinhamento com proteínas alergênicas presentes em vários bancos de dados (IVANCIUC; SCHEIN; BRAUN, 2003). Adicionalmente, a natureza alérgica ou não alérgica dos modelos foi predita pelo servidor AlgPred através de alinhamento com peptídeos representativos de alérgenos (ARPs) e predição por um módulo de máquina de vetores de suporte (SAHA; RAGHAVA, 2006). As propriedades físico-químicas incluindo peso molecular, meia-vida, índice de instabilidade, índice alifático e grande média de hidropatia

(GRAVY) foram calculadas pela ferramenta ExPASy-ProtParam, através da sequência primária das proteínas (ARTIMO et al., 2012b).

#### 4.9 Predição de estruturas terciárias e validação

As estruturas tridimensionais das proteínas foram modeladas usando o algoritmo C-QUARK, que combina o método de predição *ab initio* à predição de mapa de contatos baseada em aprendizado profundo (ZHENG et al., 2019). Em seguida, os modelos foram submetidos a 25ns de simulação de dinâmica molecular, seguida por procedimento de refinamento no programa *GalaxyRefine*, através da reconstrução de cadeias laterais seguida pelo subsequente relaxamento geral da estrutura por simulação de dinâmica molecular (HEO; PARK; SEOK, 2013).

A fim de reconhecer os possíveis erros nas estruturas terciárias iniciais e comparar os modelos antes e depois dos processos de refinamento, foi utilizado o servidor *ProSA-web* (WIEDERSTEIN; SIPPL, 2007), para gerar o *Z-score* em comparação a estruturas experimentalmente determinadas depositadas no *Protein Data Bank* (PDB). Gráficos de Ramachandran (programa Procheck) e função QMEAN disponíveis no servidor SwissModel (BIASINI et al., 2014) foram aplicados para avaliar a qualidade estereoquímica dos modelos e para estimar a energia qualidade dos modelos de proteína, respectivamente.

Em uma terceira etapa de validação de predição estrutural, cada epítipo presente nas proteínas projetadas foi submetido à comparação com estruturas terciárias de proteínas da família Paramyxoviridae depositadas no PDB. Tais estruturas foram selecionadas com base na identidade e cobertura em relação aos epítipos. Em seguida, os epítipos e respectivos modelos foram estruturalmente alinhados com base no desvio quadrático médio (RMSD) das posições atômicas.

#### 4.10 Docagem molecular

Foram realizados dois tipos de simulações para analisar a interação de epítipos virais com diferentes receptores imunológicos. A primeira estratégia consistiu na docagem molecular entre peptídeos virais reconhecidos por CTL e HTL e variantes dos principais complexos de histocompatibilidade classes I e II, respectivamente, conforme predição prévia obtida no servidor IEDB. As estruturas de todos os receptores foram obtidas no Protein Data Bank ou obtidas por modelagem de homologia usando o programa Modeller (FISER; SALI, 2003) (Apêndice C). A segunda estratégia consistiu na docagem molecular entre modelos de proteínas imunogênicas desenhadas a partir dos epítipos e o domínio de ligação extracelular do receptor

*toll-like 3* (PDB ID: 1ZIW). Além disso, as proteínas projetadas também foram docadas nos sítios de ligação dos receptores *toll-like 1* (PDB ID: 6NIH), 2 (ID: 3A7B) e 4 (ID: 3W3L). As cavidades de ligação dos receptores foram analisadas usando o servidor CASTp (TIAN et al., 2018), cujos algoritmos localizam, delineam e medem propriedades geométricas e topológicas de estruturas de proteínas. As simulações foram produzidas usando o servidor HDOCK, que aplica um protocolo híbrido entre modelagem baseada em modelos e simulações *free docagem* (YAN et al., 2017). Em seguida, as principais interações não covalentes dos complexos proteína-proteína foram analisadas pela ferramenta *Generate* disponível no banco de dados PDBsum (LASKOWSKI et al., 2018). Por fim, foram selecionados modelos baseados nas melhores pontuações de docagem para serem analisados por simulações de DM.

#### 4.11 Simulações de dinâmica molecular (DM)

Simulações de DM foram realizadas no pacote Amber16 (SALOMON-FERRER; CASE; WALKER, 2013) entre proteínas multi-epítopo selecionadas e o receptor *toll-like 3*. Foram analisados dois sistemas: o receptor no estado não ligado (livre de antígeno) e em estado complexado com os modelos multi-epítopo. As estruturas das proteínas foram tratadas com mecânica molecular usando o campo de força Amberff14SB. Inicialmente, os complexos foram protonados e solvatados em uma caixa d'água octaédrica truncada com o modelo de solvatação explícita TIP3P (MARK; NILSSON, 2001) usando o módulo Tleap do pacote Amber16. A distância entre a parede da caixa d'água e os átomos do sistema solvatado foi fixada em 15,0Å. Contraíons Cl<sup>-</sup> foram adicionados para manter a eletroneutralidade do sistema, de acordo com a carga total calculada pelo módulo Tleap. Inicialmente, um protocolo de minimização de energia de nove etapas para os sistemas foi realizado, envolvendo os métodos *steepest descent* e gradiente conjugado (BRERETON, 2010). O algoritmo SHAKE foi aplicado para todas as ligações de hidrogênio dos sistemas analisados (KRAUTLER; VAN GUNSTEREN; HANENBERGER, 2001). Os sistemas de proteína foram gradualmente aquecidos durante 5ns de simulação para aumentar a temperatura para 300 K. A temperatura foi mantida em 300 K (temperatura final) por acoplamento ao termostato de Langevin (LZAGUIRRE et al., 2001) usando uma frequência de colisão de 2 ps<sup>-1</sup> e a pressão isotrópica foi mantida constante em 1 bar usando o barostato de Berendsen. Em seguida, 10 ns de dinâmica foram realizados para a etapa de equilíbrio. Finalmente, 100ns de simulação de DM foram aplicados para cada sistema de proteína em *ensemble* isotérmico-isobárico (NPT) usando temperatura de 300K e pressão de 1 bar. Um ciclo de integração de 2fs foi aplicado para analisar as trajetórias e um *cut-off* de 12,0 Å foi aplicado para as interações não ligadas e o método de malha de partículas

Ewald (PME)(SHAN et al., 2005) foi aplicado para calcular as interações eletrostáticas de longo alcance. Os gráficos de RMSD e RMSF foram obtidos para cada sistema durante os 100 ns da trajetória MD usando como referência os átomos pesados (C $\alpha$ , N e O) do esqueleto da proteína.

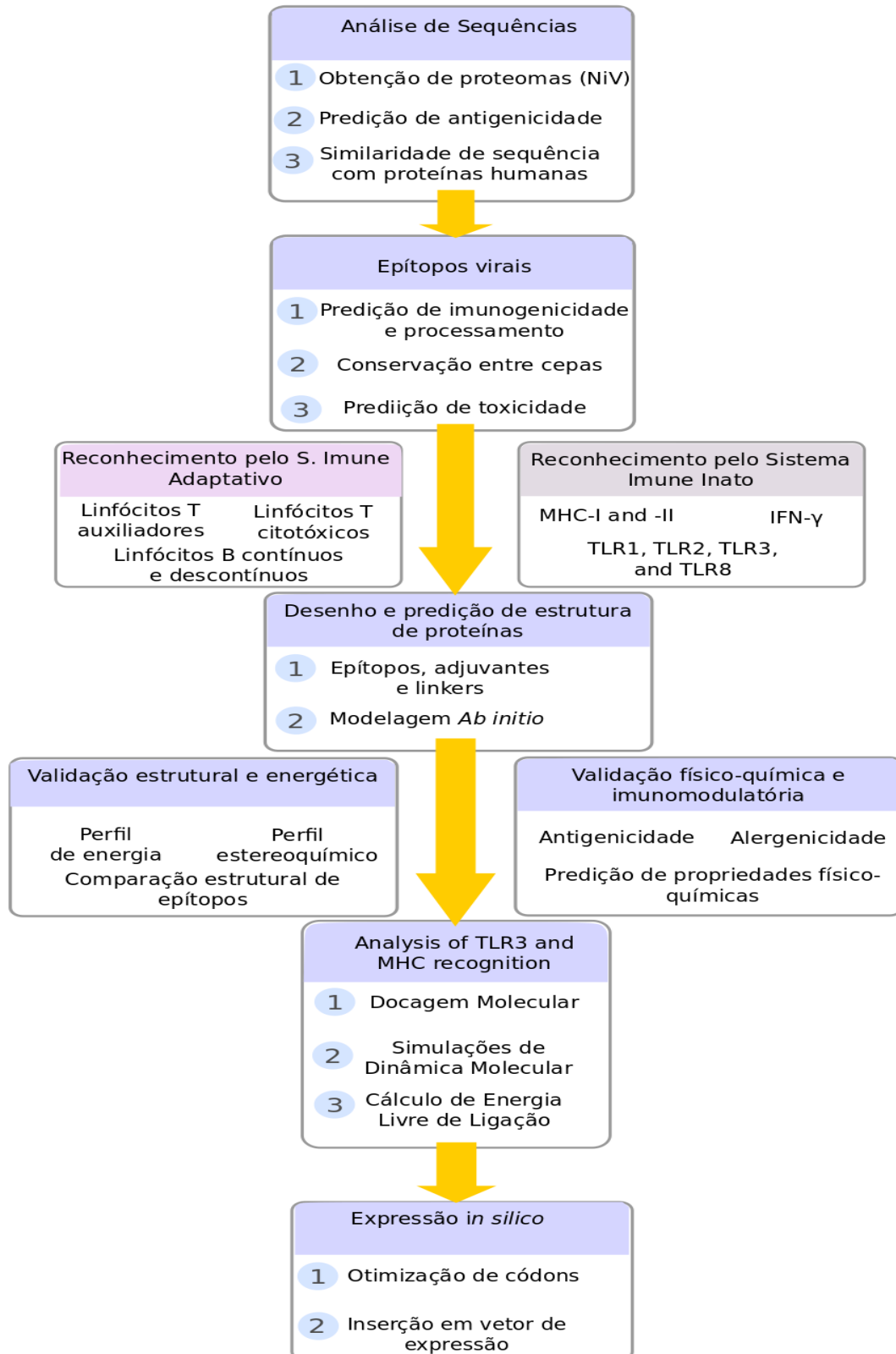
#### 4.12 Cálculos de Energia Livre de Ligação

Para avaliar as afinidades de ligação dos modelos de proteínas multi-epítipo projetados complexados com TLR3, foram realizados cálculos de energia livre de ligação dos complexos usando os métodos MM/PBSA (*Molecular Mechanics Poisson–Boltzmann Surface Area*) e MM/GBSA (*Molecular Mechanics Generalized Born Surface Area*). Os frames finais da trajetória de DM foram obtidos usando o módulo Cpptraj, então as energias livres de ligação foram calculadas usando o módulo MMPBSA.py disponível no pacote Amber16 (GENHEDEN; RYDE, 2015).

#### 4.13 Otimização de códons e expressão *in silico* em sistema vetorial

A otimização do uso de códons foi realizada para aumentar a eficiência da expressão de cDNA em vetores de clonagem com a finalidade de obter uma expressão heteróloga ótima nos hospedeiros bacterianos. Em primeiro lugar, as sequências de cDNAs das proteínas desenhadas foram geradas pelo servidor JCat (GROTE et al., 2005) usando o hospedeiro bacteriano *Escherichia coli* como sistema de expressão. Além disso, a ferramenta GenScript Rare Codon Analysis (<https://www.genscript.com/tools/rare-codon-analysis>) foi usada para realizar a otimização do uso de códons e para gerar o índice de conteúdo de guanina-citosina (GC), o índice de adaptação de códons e medir a frequência de códons em tandem (GIBSON, 2012). Em seguida, o vetor pET28a+ foi usado como um vetor de expressão para o cDNA dos modelos multi-epítipo usando o software SnapGene (<https://www.snapgene.com/>). Às sequências dos modelos selecionados foram inseridos sítios de restrição BamHI/HindIII; o códon de iniciação ATG e sequência GGATCC (sítio de restrição BamHI) foram introduzidos na região inicial da sequência e o sítio de restrição HindIII (AAGCTT) foi adicionado ao final da sequência. Além disso, os modelos de cDNA foram marcados com uma etiqueta de polihistidina na extremidade 3'. Um esquema das etapas de metodologia abordadas neste trabalho está ilustrado na Figura 8.

**Figura 8-** Metodologia para desenho de proteínas imunogênicas a partir do proteoma do Vírus Nipah.



**Fonte:** Elaborado pelo Autor.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Predição de antigenicidade e avaliação de homologia de proteínas

As sequências de aminoácidos das nove proteínas do NiV, baseadas nos proteomas obtidos no banco de dados Uniprot (Apêndice A), foram preditas com antigenicidade pelo servidor AntigenPRO, com valor limite de predição igual a 0,4. Todas as proteínas investigadas foram identificadas com potencial antigenicidade e as proteínas V, W e P apresentaram maior potencial antigênico (0,95, 0,94 e 0,94, respectivamente).

A similaridade de sequência entre proteínas do hospedeiro e antígenos provenientes do patógeno podem levar a doenças autoimunes (KANDUC, 2012) e mimetismo molecular (ROJAS et al., 2018). Com base nessa suposição, o alinhamento de sequência foi produzido através de BLASTp usando o proteoma humano como um parâmetro padrão. As análises mostraram que as proteínas W, N e C não apresentaram grau suficiente de similaridade com nenhuma outra proteína humana. Estudos anteriores mostram que as proteínas de NiV apresentam menos de 40% de similaridade de sequência com proteínas codificadas pelo genoma humano (OJHA et al., 2019). No entanto, as proteínas F e M mostram uma identidade maior do que 40% com AngRem52. Além disso, as proteínas P, V e W têm uma identidade de 30,51% com AngRem104 (Tabela 2).

**Tabela 2** - Predição de antigenicidade das proteínas NiV usando alinhamento baseado em sequência. Cada proteína é identificada por seu respectivo índice de antigenicidade, cobertura (%), identidade (%) e homólogo humano previsto.

Proteína	Antigenicidade	Cobertura (%)	Identidade (%)	Potencial Homólogo
Glicoproteína de fusão F0 (F)	0,55	40	41.96	AngRem52
Glicoproteína G (G)	0,89	13	27.17	Mel transforming oncogene variant
Proteína V (V)	0,94	25	30.51	AngRem104
Proteína W (W)	0,94	25	30.51	AngRem104
Fosfoproteína (F)	0,95	16	30.51	AngRem104
Proteína de matriz (M)	0,50	95	53.12	AngRem52
Proteína C (C)	0,70	Não disponível	Não disponível	Não disponível
Nucleoproteína (N)	0,74	Não disponível	Não disponível	Não disponível
Polimerase de RNA (L)	0,41	Não disponível	Não disponível	Não disponível

**.Fonte:** Elaborado pelo Autor.

Investigações experimentais previamente revelaram que as sequências AngRem52 e AngRem104, derivadas de cDNA originalmente descritas como produtos de genes humanos, são sequências virais de um membro da família Paramyxoviridae (LI et al., 2006; SCHOMACKER; COLLINS; SCHMIDT, 2004). Portanto, considerando a falta de

similaridade a proteínas humanas, foram exploradas todas as proteínas virais para predição de epítomos visando o desenho de proteínas imunogênicas.

## 5.2 Predição e filtragem de epítomos de células T e B

Para prever e analisar os epítomos reconhecidos por CTL e HTL, foram utilizadas as ferramentas disponíveis no *Immune Epitope Database and Analysis Resource* (IEDB) (VITA et al., 2015). A análise inicial foi baseada na afinidade das proteínas virais (classificação pelo *percentile rank*) com 12 moléculas representativas de diferentes supertipos do MHC-I, com a finalidade de obtenção de epítomos reconhecidos por linfócitos T citotóxicos. Foram selecionados epítomos com pontuações de imunogenicidade superiores a 0 e com conservação nos 15 proteomas de NiV analisados. Foram obtidos 78 epítomos potencialmente imunogênicos que foram submetidos à análise *in silico* quanto ao processamento intracelular com base na via de apresentação do antígeno do MHC-I (pontuação de processamento), promiscuidade por afinidade (supertipos MHC-I) e sua potencial toxicidade. Com base nesses filtros, foram selecionados um total de 14 epítomos de CTL para o desenho de proteínas (Tabela 3).

**Tabela 3-** Epítomos virais de proteínas de NiV reconhecidas pelos linfócitos T citotóxicos e respectivos supertipos de MHC-I, *percentile rank* (PR), pontuação de imunogenicidade (IS), pontuação de processamento (PS) e cobertura populacional..

Proteína	Epítomo (Sequência)	MHC-I (supertipos)	PR	IS	PS	Cobertura
L	SHMWRELAL	B39, B8	0,06	0,32	1,96	13,15%
C	WTQWLQTTY	A1, A26	0,11	0,08	2,5	22,62%
F	AQITAGVAL	B44	0,12	0,21	1,29	18,43%
G	LMMTRLAVK	A3	0,12	0,08	2,52	17,34%
F	ETLLRTLGY	A26, A1	0,14	0,08	2,41	22,62%
G	AMDEGYFAY	A1	0,17	0,26	1,26	3,42%
F	VLFANCISV	A2	0,2	0,04	2,18	39,08%
G	FLIDRINWI	A2	0,2	0,37	1,94	39,08%
P, W, V	KSRGIPIKK	A3	0,35	0,16	1,26	16,81%
M	YMYLICYGF	A24, B62	0,37	0,08	2,52	28,02%
G	TVNPLVVNW	B58	0,4	0,05	1,37	16,81%
F	EWISIVPNF	A24, A26	0,49	0,03	2,57	26,52%
N	RASAATATL	B7, B58	0,6	0,12	2,01	15,97%
F	IVRVYFPIL	B8, B39, B62	0,9	0,23	1,98	10,55%

**Fonte:** Elaborado pelo Autor.

Da mesma forma, os epítomos reconhecidos pelo HTL foram identificados usando critérios semelhantes, exceto a predição de processamento intracelular. A predição da ligação dos peptídeos a 26 variantes do MHC-II foi explorada, cobrindo 6 supertipos (DR, DP, DP2,



DRB3, DR4 e DQ). Os epítomos previstos como indutores de MHC-II foram filtrados em relação à indução de IFN- $\gamma$ . Finalmente, um total de 17 epítomos HTL foram obtidos para estudos adicionais (Tabela 4).

**Tabela 4-** Epítomos virais de proteínas NiV reconhecidos pelos linfócitos T auxiliares.

Proteínas	Epítomo (Sequência)	MHC- II (supertipos)	Percentile rank	Imunogenicidade	Cobertura populacional
C	MMASILLTLFRRTKK	DR, DP, DP2	0,17	39,42	37,50%
N	AVEIISDIGNYVEET	DRB3, DQ, DP, DP2	0,21	39,71	23,90%
F	NSEWISIVPNFILVR	DR, DRB3, DR4, DQ	0,64	28,08	41,94%
L	ADKVLEYAPIMKAHA	DR, DR4, DRB3, DP, DP2	0,7	36,1	59,94%
F	FISFIIVEKKRNTYS	DR	0,91	44,94	14,77%
F	VPNFILVRNTLISNI	DR, DR4, DRB3	1,2	35,09	36,03%
M	KRKKIRTIAAYPLGV	DR, DRB3	1,4	34,98	29,83%
F	IILYVLSIASLCIGL	DR, DRB3	1,6	40,41	43,78%
F	TAVLGNVIISLGKYL	DR, DRB3, DQ	1,6	44,72	24,27%
G	DSKILSAFNTVIALL	DR, DRB3, DR4, DP, DP2, DQ	1,6	34,48	17,82%
G	SAVYNNEFYVLCVAV	DP2, DRB3, DQ	1,65	75,54	12%
G	SKILSAFNTVIALLG	DR, DR4	2,1	35,05	32,82%
G	GVYNDAFLIDRINWI	DRB3, DP2	2,4	41,84	23,90%
P, V, W	EFEDFAGSSSEVIV	DR, DQ	2,45	78,23	18,23%
F	NIEIGFCLITKRSVI	DR	6,4	46,43	6,40%

**Fonte:** Elaborado pelo Autor.

De acordo com as análises *in silico*, o epítomo FLIDRINWI reconhecido por linfócitos T citotóxicos, e o epítomo ADKVLEYAPIMKAHA reconhecido por linfócitos T auxiliares atingem 39,08% e 59,94% da população mundial, respectivamente. Todos os epítomos selecionados juntos mostraram cobertura acima de 98% da população humana, um índice superior ao relatado em investigações prévias de análise de epítomos virais (ALI; MORSHED; HASSAN, 2015; SAHA et al., 2017; SAKIB et al., 2014).

Também foi realizada uma predição baseada em sequência dos epítomos lineares reconhecidos pelos linfócitos B usando o servidor BepiPred 2.0 (JESPERSEN et al., 2017) e apenas epítomos não tóxicos e conservados entre as cepas de NiV conhecidas foram mantidos na biblioteca de peptídeos. Dos 85 peptídeos inicialmente previstos, apenas 12 epítomos com o comprimento entre 5 e 11 resíduos foram selecionados para o projeto de proteína de múltiplos epítomos (Tabela 5).

**Tabela 5-** Epítomos virais reconhecidos por linfócitos B previstos a partir do proteoma do vírus.

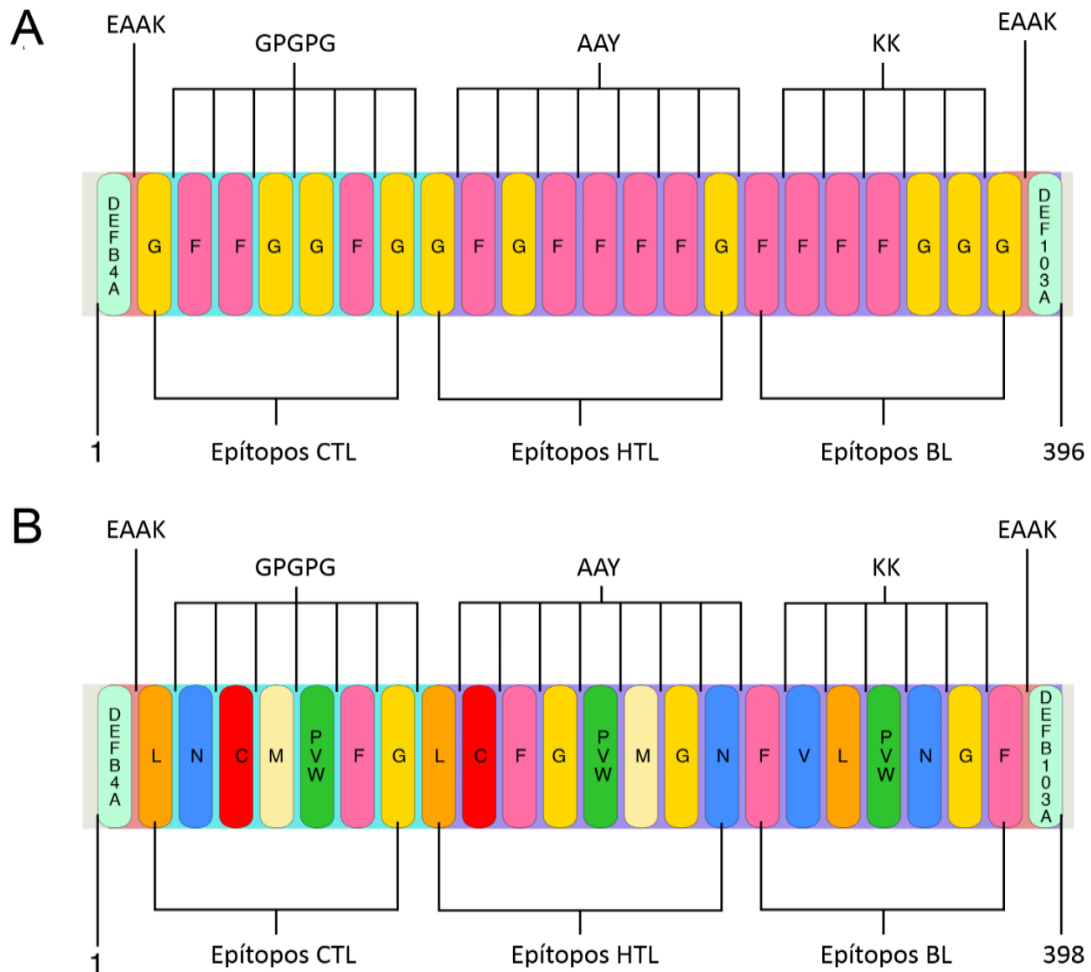
<b>Proteína</b>	<b>Epítopo viral</b>
Glicoproteína de fusão F0	KIGLVKG
Glicoproteína de fusão F0	ISIVPNFI
Glicoproteína de fusão F0	IQELLPVS
Glicoproteína de fusão F0	TAGVALY
Glicoproteína G	VYHCSAV
Glicoproteína G	YFPAVGFL
Glicoproteína G	PSKVIKS
Glicoproteína G	EGVYN
Proteína V	YHDHGGE
RNA polimerase L	LYILRQSKQGD
Nucleoproteína	VLLEESIQT
Proteínas W, P and V	ELVNDGL

**Fonte:** Elaborado pelo Autor.

### 5.3 Construção dos modelos de vacina

Duas estratégias de desenho foram utilizadas para obter estruturas de proteínas multiepítopo (Figura 9). Usando diferentes combinações de epítomos, foram obtidos seis modelos finais compostos por um total de 395-398 resíduos usando 8 epítomos CTL e 7 HTL; e 7 epítomos contínuos reconhecidos por linfócitos B. Os modelos 1, 2 e 3 foram baseados na primeira estratégia (epítomos das proteínas F e G), enquanto os modelos 4, 5 e 6 contêm epítomos de todas as proteínas virais (Apêndice D).

**Figura 9** - Visão geral esquemática dos dois tipos de proteínas multiepítipo projetadas. Os modelos finais incluíram as seguintes sequências: epítomos virais reconhecidos por CTL, HTL e linfócitos B lineares (BL), que foram fundidos com ligantes AAY, GPGPG e KK, respectivamente. O ligante EAAAK foi usado para conectar as  $\beta$ -defensinas como adjuvantes nas regiões N e C terminais de ambas as estruturas.



**Fonte:** Elaborado pelo Autor.

Diferentes estudos relataram sucesso no projeto de modelos de proteínas multiepítipo para NiV usando proteínas F e G, encontradas na superfície viral (ALI; MORSHED; HASSAN, 2015; SAHA et al., 2017; SAKIB et al., 2014). No entanto, estudos experimentais mostraram que as proteínas V (RODRIGUEZ; PARISIEN; HORVATH, 2002), W (SATTERFIELD; DAWES; MILLIGAN, 2016b; SHAW et al., 2005), P (JENSEN et al., 2020), M (BHARAJ et al., 2016), C (MATHIEU et al., 2012) e N (SUGAI et al., 2017) desempenham um papel importante na evasão da resposta imune humana. Dessa forma, foram incluídos epítomos provenientes de todas as proteínas virais com o fim de potencializar a resposta imune e maximizar a diversidade antigênica.

#### 5.4 Predição e validação de estruturas das proteínas multiepítipo

Embora grande parte das pesquisas envolvendo desenho de proteínas imunogênicas empregue as abordagens de modelagem por homologia e *threading* para predição estrutural (CHAUHAN et al., 2019; IKRAM et al., 2018; ALI et al., 2017; USMAN MIRZA et al., 2016), a modelagem *ab initio* já foi aplicada com sucesso para explorar o espaço conformacional de epítomos que não possuem estruturas de referência com cobertura e identidade adequadas para utilização como modelos para predição estrutural (ZHENG et al., 2019), como nos casos da proteína do nucleocapsídeo do sangue hemorrágico da Crimeia-Congo vírus da febre (SRINIVASAN et al., 2011), do fator de infectividade viral do vírus da imunodeficiência humana (DA COSTA et al., 2014) e da proteína capsidial do vírus do mosaico do Cardamomo (KUMAR et al., 2016). No presente estudo, aproximadamente 12% e 20% de cobertura e de identidade de sequências foram encontradas, respectivamente, para as proteínas multi-epítomo projetadas, em relação às estruturas depositadas no *Protein Data Bank*. Com base nisso, foi aplicada a abordagem *ab initio* de C-QUARK desenvolvida por Zhang et al (2019).

Com base no *C-score* obtido, os modelos 2 e 4 mostraram melhores pontuações de predição estrutural, com valores iguais a -22,44 e -17,81, respectivamente (Tabela 6). Ainda, os resultados de validação obtidos mostraram que a predição estrutural realizada foi consistente. Todas as proteínas obtiveram mais de 85% de resíduos nas regiões mais favoráveis do gráfico de Ramachandran e menos de 3% de resíduos em regiões não permitidas (Apêndice E). Da mesma forma, os *Z-Scores* obtidos no servidor Prosa-Web (WIEDERSTEIN; SIPPL, 2007), que compara a qualidade energética das estruturas preditas com as estruturas depositadas experimentalmente no *Protein Data Bank*, mostraram que os modelos apresentam índices de qualidade comparativamente semelhantes ao de estruturas cristalográficas. Além disso, os gráficos do QMEAN apresentaram pontuações positivas para ambas as estruturas, o que indica a qualidade energética satisfatória dos modelos (Tabela 6).

**Tabela 6** - *C-score* da predição estrutural obtida no servidor C-Quark, *Z-Score* obtido no servidor Prosa-Web, índice QMEAN obtido no servidor Swiss Model e distribuição dos ângulos residuais de acordo com as regiões do gráfico de Ramachandran.

Model ID	<i>C-Score</i>	<i>Z-Score</i>	QMEAN	Regiões do Gráfico de Ramachandran			
				Regiões mais favoráveis	Regiões adicionalmente permitidas	Regiões generosamente permitidas	Regiões não permitidas
4	-22,44	-9,89	0,83	92,30%	5,80%	0,30%	1,50%
2	-17,81	-6,75	0,75	88,90%	9,80%	0,90%	0,30%
3	-15,43	-8,11	0,85	87,30%	9,30%	1,20%	2,20%
6	-15,42	-7,45	0,71	89,80%	8,00%	0%	2,20%
1	-13,99	-11,9	0,81	90,50%	8,30%	0%	1,20%
5	-13,31	-7,11	0,78	86,50%	10,10%	1,20%	2,10%

**Fonte:** Elaborado pelo Autor.

Adicionalmente, a comparação das estruturas dos epítomos potenciais presentes nas proteínas desenhadas com fragmentos de estruturas cristalográficas de diversos membros da família Paramyxoviridae fornece evidências da conservação estrutural desses peptídeos (Tabela 7) e indica que a presença de ligantes entre epítomos não gerou impactos estruturais de grande impacto. Os RMSDs obtidos variaram entre 0,079 e 1,567.

**Tabela 7** - Alinhamento estrutural entre epítomos previstos e estruturas de paramixovírus depositadas no *Protein Data Bank* (PDB).

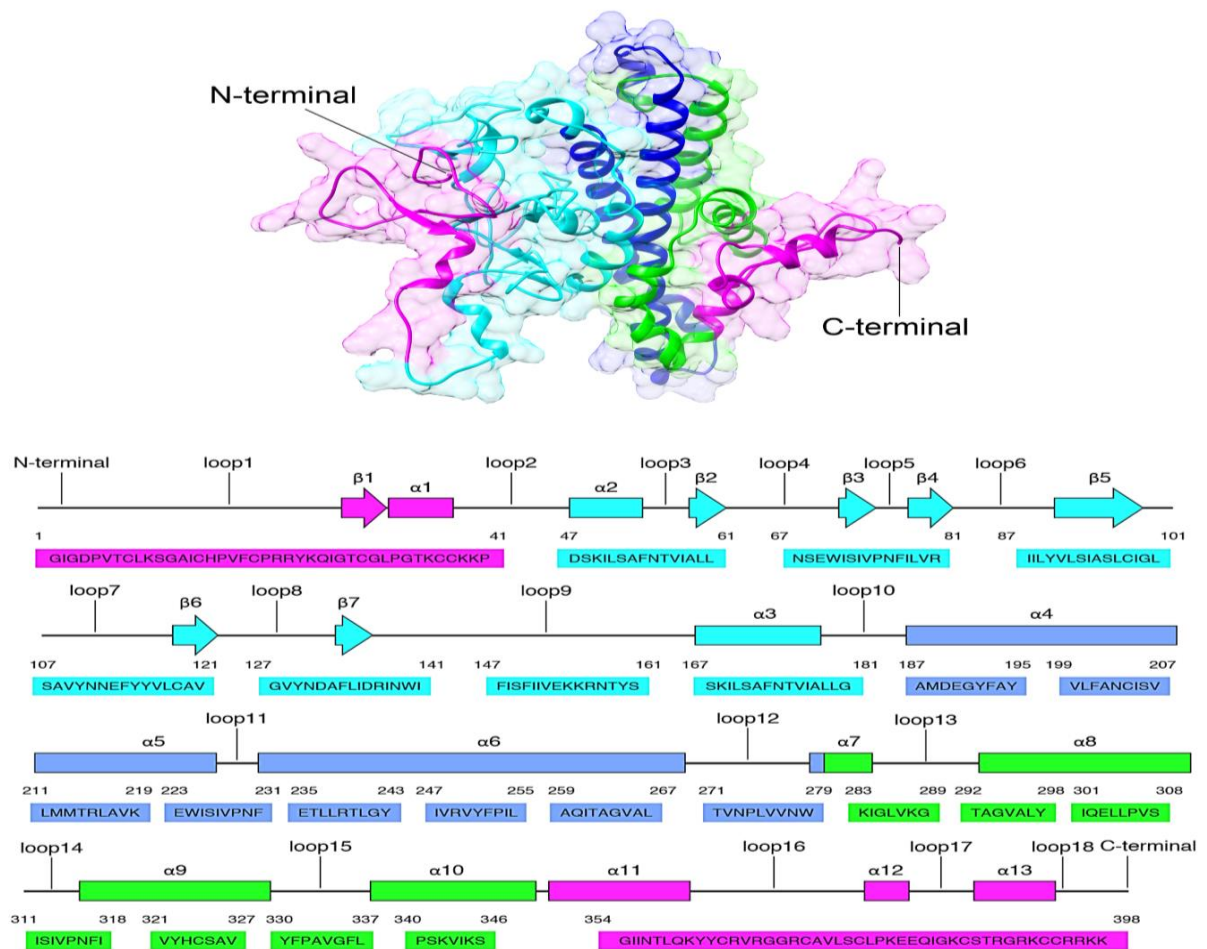
Epítopo	Proteína	PDB ID (Referência)	RMSD	Cobertura	Identidade
VLFANCISV	F	6TYS	0,079	100%	100%
YMYLICYGF	M	6BK6	0,215	100%	88.89%
IVRVYFPIL	F	6TYS	0,216	100%	100%
ETLLRTLGY	F	6TYS	0,236	100%	100%
TAGVALY	F	6TYS	0,241	100%	100%
ADKVLEYAPIMKAHA	L	6V85	0,254	80%	58.3%
ELVNDGL	W, P, V	4CO6_B	0,261	100%	100%
RASAATATL	N	1EJ6	0,297	77%	100%
VPNFILVRNTLISNI	F	6TYS	0,388	100%	100%
AQITAGVAL	F	6TYS	0,420	100%	100%
WTQWLQTLY	C	2O2C	0,508	100%	77.78%
MMASILLTLFRRTKK	C	3WV4	0,589	80%	50%
EGVYN	G	3D11	0,573	100%	100%
EFEDDEFAGSSSEVIV	P,V,W	3CER	0,607	93%	52.94%
DSKILSAFNTVIALL	G	5MRW	0,618	100%	64.29%
VYHCSAV	G	3D11	0,674	100%	100%
GVYNDAFLIDRINWI	G	3D11	0,772	100%	100%
TVNPLVNVW	G	3D11	0,794	100%	100%
AMDEGYFAY	G	3D11	0,835	100%	100%
YHDHGGE	V	6KSI	0,847	100%	75%
NSEWISIVPNFILVR	F	6TYS	0,860	100%	100%
NSEWISIVPNFILVR	F	6TYS	0,921	100%	100%
LYILRQSKQGD	L	5J3U	0,933	90%	100%
TAVLGNVIISLGKYL	F	6TYS	0,994	100%	100%
ISIVPNFI	F	6TYS	0,995	100%	100%
SKILSAFNTVIALLG	G	4U4C	1,09	86%	61.54%
IILYVLSIASLCIGL	F	4G91	1,11	86%	50%
AVEIISDIGNYVEET	N	4C06_A	1,032	100%	100%
KSRGIPIKK	P, W, V	4ZWQ	1,034	88%	75%
FISFIVEKKRNTYS	F	5ULD	1,044	93%	64.29%
YFPAVGFL	G	6PD4	1,078	100%	100%
NIEIGFCLITKRSVI	F	6TYS	1,088	100%	100%
KIGLVKG	F	6TYS	1,096	100%	100%

PSKVIKS	G	5F4L	1,112	85%	83.33%
SAVYNNEFYVLCVAV	G	3D11	1,157	100%	100%
SHMWRELAL	L	4GUZ	1,186	100%	63.64%
LMMTRLAVK	G	3D11	1,198	100%	100%
EWISIVPNF	F	6TYS	1,253	100%	100%
IQELLPVS	F	6TYS	1,255	100%	100%
KRKKIRTIAAYPLGV	M	6BK6	1,299	100%	100%
FLIDRINWI	G	3D11	1,494	100%	100%
VLLEESIQT	N	4C06_A	1,567	100%	100%

**Fonte:** Elaborado pelo Autor.

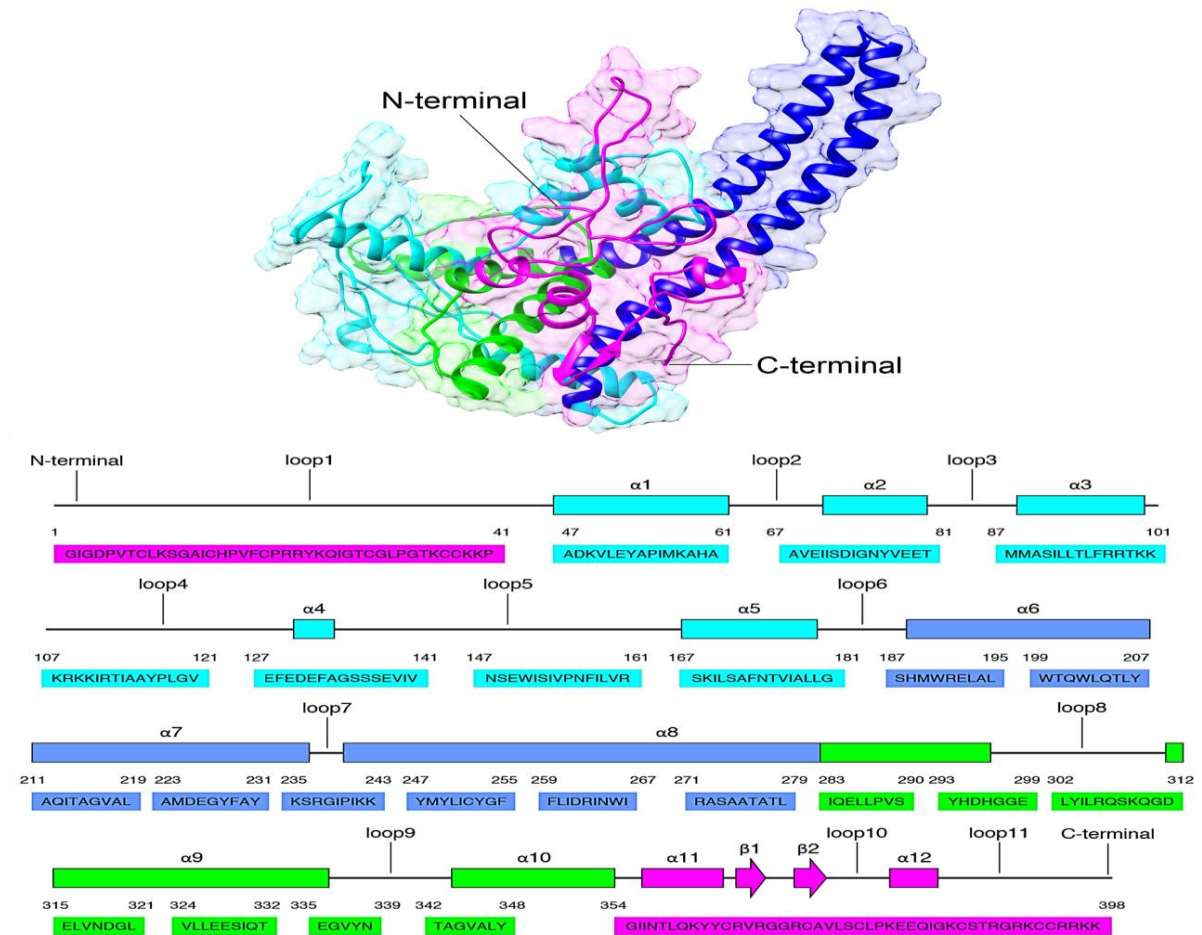
As estruturas tridimensionais dos modelos 1 e 4 de proteínas com múltiplos epítomos podem ser observadas nas Figura 10 e 11, que ilustram a distribuição espacial dos resíduos contendo regiões adjuvantes, epítomos reconhecidos por linfócitos T citotóxicos, T auxiliares e B e seus respectivos *linkers*.

**Figura 10** - Estruturas secundária e terciária de proteínas multi-epítomo projetadas (Modelo 1).



**Fonte:** Elaborado pelo Autor.

**Figura 11** - Estruturas secundária e terciária de proteínas multi-epítipo projetada (Modelo 4).

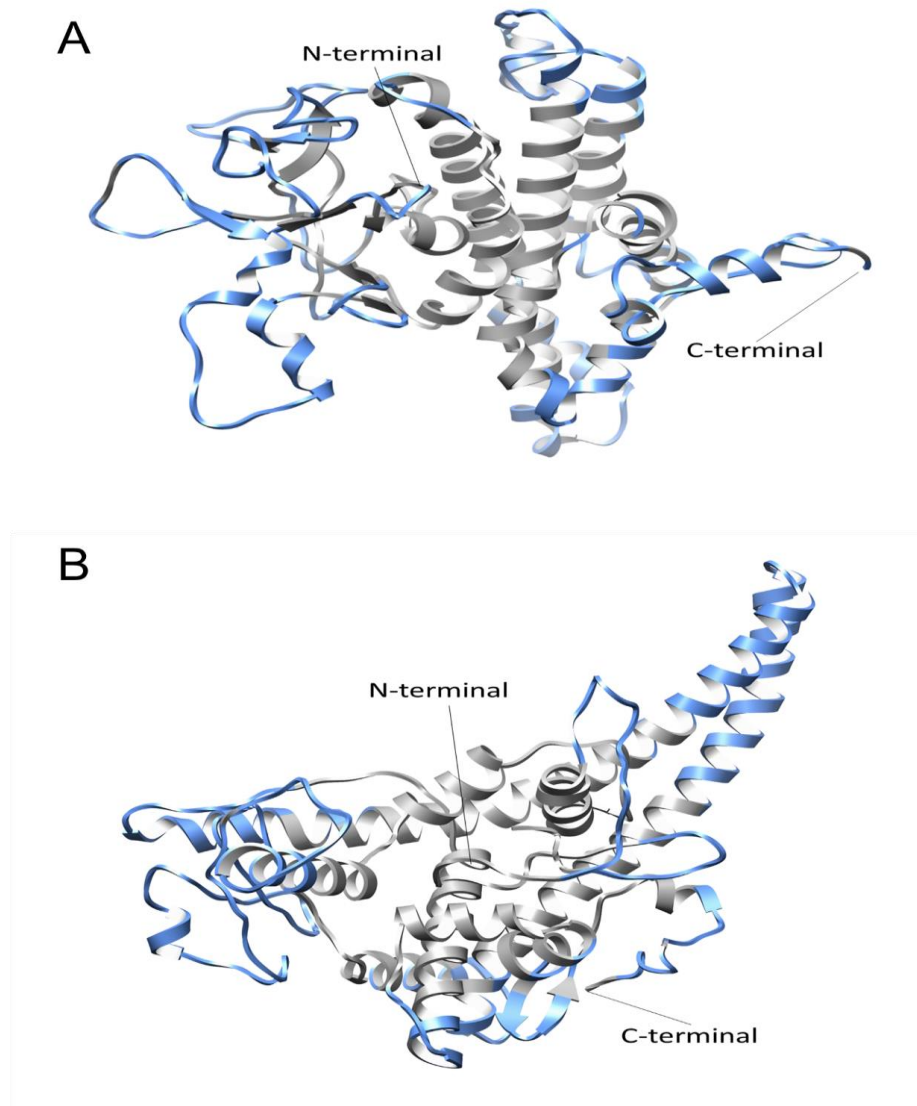


Fonte: Elaborado pelo Autor.

### 5.5 Predição de linfócitos B conformacionais/descontínuos

A aplicação de um algoritmo de aprendizado de máquina para prever epítomos conformacionais/descontínuos reconhecidos por linfócitos B apontou para 10 epítomos potenciais no modelo 1 e 6 epítomos no modelo 4 reconhecidos por linfócitos B. Observa-se que tais regiões de epítomos abrangem regiões de superfície em proteínas e também coincidem com a localização de epítomos reconhecidos por linfócitos T, o que se mostra como favorável à múltipla estimulação de reconhecimento pelo sistema imune ( Figura 12).

**Figura 12** - Estruturas das proteínas multi-epítipo projetadas mostrando a localização dos epítomos conformacionais reconhecidos pelos linfócitos B (em azul). A) Modelo 1. B) Modelo 4.



**Fonte:** Elaborado pelo Autor.

### 5.6 Propriedades físico-químicas, antigenicidade e alergenicidade

Diferentes propriedades físico-químicas das proteínas desenhadas foram calculadas de acordo com os parâmetros do servidor ProtParam (GASTEIGER et al., 2003, 2009). O potencial isoelétrico (PI) dos modelos variou entre 9,51 e 9,68, o que indica uma natureza básica, enquanto o peso molecular variou entre 42.374,36 (modelo 3) e 43,007,21 (modelo 4). Os índices de instabilidade obtidos foram menores do que o valor padrão (40), o que indica que as



proteínas são potencialmente estáveis. O índice alifático variou entre 84,16 e 100,03, o que indica que ambas as proteínas são termoestáveis, um fator interessante para a estabilidade estrutural à temperatura humana. Os valores de meia-vida de cada modelo foram preditos em 30 h para modelos *in vitro* (reticulócitos de mamíferos) e este tempo é consistente para ser reconhecido pelo sistema imunológico. A grande média de hidropaticidade (GRAVY) variou entre 0,132 e 0,311; e essas pontuações positivas referem-se à sua natureza hidrofóbica. Finalmente, todos os modelos de multiepítotos projetados foram preditos como potencialmente antigênicos, obtendo taxas mais altas do que o valor de corte do servidor AntigenPRO (0,4).

Em conjunto, as propriedades analisadas mostraram que as proteínas analisadas são potencialmente estáveis e atendem a várias características desejadas para a formulação de vacinas de próxima geração (BATEMAN et al., 2017) (Tabela 8).

**Tabela 8:** Propriedades estruturais e físico-químicas de cada modelo multi-epítoto projetado.

Propriedades	Modelos (ID)					
	1	2	3	4	5	6
Peso molecular	42,423.35	42,423.35	42,374.36	43,007.21	43,007.21	42,958.23
Potencial isoelétrico	9,68	9,68	9,73	9,51	9,51	9,57
Índice de instabilidade	21,68	21,68	21,61	32,12	32,5	32,15
Índice alifático	100,28	100,28	100,03	84,45	84,45	84,16
Meia vida	30 horas	30 horas	30 horas	30 horas	30 horas	30 horas
GRAVY	0,311	0,311	0,286	0,132	0,132	0,157
Antigenicidade	0,5693	0,5827	0,5809	0,5148	0,5128	0,5331

**Fonte:** Elaborado pelo Autor.

A alergenicidade dos modelos foi também predita usando diferentes métodos computacionais. Considerando os métodos de alinhamento baseados em sequência (SDAP, AllerTOP, e AlgPred Blast), todas as proteínas multi-epítoto obtiveram valores satisfatórios, mostrando assim potencialmente não induzir alergenicidade (Tabela 9). Os modelos 5 e 6 foram previstos para induzir alergenicidade, portanto, ambos foram excluídos das análises posteriores de interação com MHC-I, MHC-II e receptores *toll-like*.

**Tabela 9 -** Predição de alergenicidade dos seis modelos de multi-epítotos projetados usando métodos de alinhamento baseados em sequência (SDAP, AllerTOP e AlgPred Blast).

Modelo (ID)	SDAP	AllerTOP	AlgPred	AlgPred (BLAST)
1	Não alérgeno	Não alérgeno	Não alérgeno	Não alérgeno
2	Não alérgeno	Não alérgeno	Não alérgeno	Não alérgeno
3	Não alérgeno	Não alérgeno	Não alérgeno	Não alérgeno
4	Não alérgeno	Não alérgeno	Não alérgeno	Não alérgeno
5	Não alérgeno	Não alérgeno	Alérgeno	Não alérgeno
6	Não alérgeno	Não alérgeno	Alérgeno	Não alérgeno

**Fonte:** Elaborado pelo Autor.

## 5.7 Docagem Molecular

### 5.7.2 Interação entre proteínas desenhadas e TLR3

No presente estudo, foi selecionado o TLR3 para analisar a afinidade de ligação e seletividade das proteínas multi-epítopo devido à evidência de seu importante papel no reconhecimento da infecção por NiV e sua inibição mediada, principalmente, pela proteína viral W (SHAW et al., 2005). O TLR3 também é relevante por estar envolvido com a ativação da sinalização Th1 mediada por TRIF/TRAM, induzindo o sistema imunológico inato contra infecções virais (KAISER et al., 2013).

Inicialmente, as análises espaciais mostraram apenas uma única cavidade central com um volume acessível ao solvente de aproximadamente  $18.867.592 \text{ \AA}^3$ , onde os modelos foram docados para analisar as interações intermoleculares. A docagem molecular revelou que os complexos apresentaram valores de energias favoráveis de interação com o TLR3. Os complexos 4 (TLR3 complexado com o modelo 4) e 1 (TLR3 e modelo 1) obtiveram a melhor afinidade apresentando energias de ligação iguais a  $-318,45 \text{ kcal mol}^{-1}$  e  $-305,44 \text{ kcal mol}^{-1}$ , respectivamente, mostrando assim uma ligação satisfatória. Em seguida, selecionamos esses dois modelos para outras simulações de MD. Os receptores TLR1, TLR2 e TLR4 também foram analisados com os dois modelos de multi-epítopos projetados, mas obtiveram afinidades mais baixas (Tabela 10).

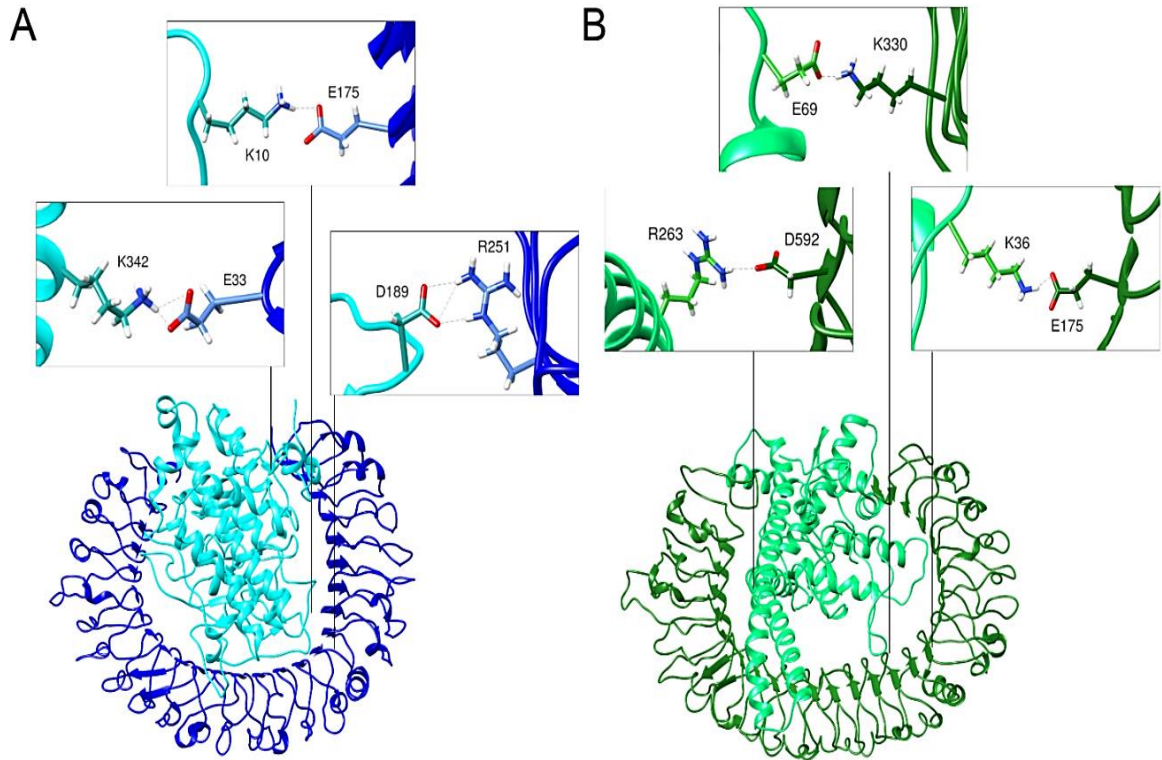
**Tabela 10** - Pontuações de docagem do HDock obtidas a partir do docagem molecular entre os modelos de multi-epítopos projetados (modelos 1 a 6) com a estrutura TLR3.

Complexos (ID)	HDock score ( $\text{kcal}\cdot\text{mol}^{-1}$ )			
	TLR3	TLR1	TLR2	TLR4
4	-318,45	-296,32	-306,37	-247,67
1	-305,44	-289,22	-299,32	-294,35
5	-292,21	-272,45	-270,36	-293,88
6	-291,92	-280,54	-245,51	-251,45
3	-290,89	-275,66	-289,52	-268,83
2	-283,22	-280,45	-279,27	-261,41

**Fonte:** Elaborado pelo Autor.

Quanto aos modelos 1 e 4, as interações com o receptor ocorreram, principalmente, em relação aos resíduos dos motivos LRR de TLR3 (Figura 13). O modelo 1 formou, na presente simulação, 14 ligações de hidrogênio e 9 pontes salinas, dentre outras 134 interações não ligadas. As ligações de hidrogênio apresentaram distância média de  $2,8 \text{ \AA}$ . Por sua vez, o modelo 4 formou 6 ligações de hidrogênio e 3 pontes salinas, dentre outras 287 interações não ligadas. A descrição detalhada de interações está contida nos Apêndices F e G.

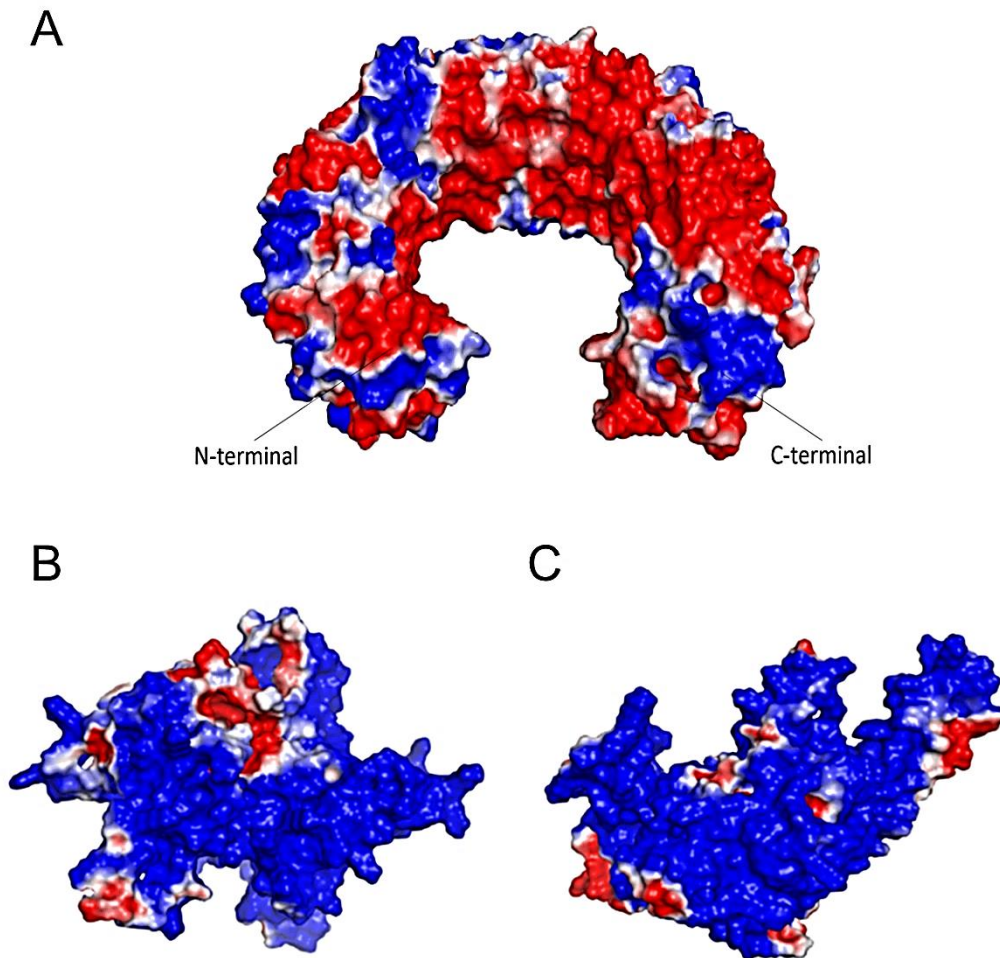
**Figura 13-** Interações intermoleculares obtidas por docagem molecular entre TLR3 e modelos da proteína multipeptípo. A) Modelo 1. B) Modelo 4.



**Fonte:** Elaborado pelo Autor.

Adicionalmente, o mapa de potencial eletrostático dos modelos 1 (Figura 14B) e 4 (Figura 14C) apresentou a maioria de resíduos em regiões com potencial positivo (azul) e uma minoria em regiões negativas (vermelho) e neutras (branco). Essas características estruturais são compatíveis com o reconhecimento pelo TLR3, que apresenta predominância de cargas negativas em seu domínio de ligação extracelular (Figura 14A).

**Figura 14** - Mapas de potencial eletrostático. A) TLR3. B) Modelo 1. C) Modelo 2.



**Fonte:** Elaborado pelo Autor.

De modo geral, os resultados obtidos foram consistentes com estudos computacionais anteriores que mostraram a interação de TLR3 com proteínas multi-epítipo projetadas para NiV (OJHA et al., 2019) e MERS-CoV (SRIVASTAVA et al., 2018).

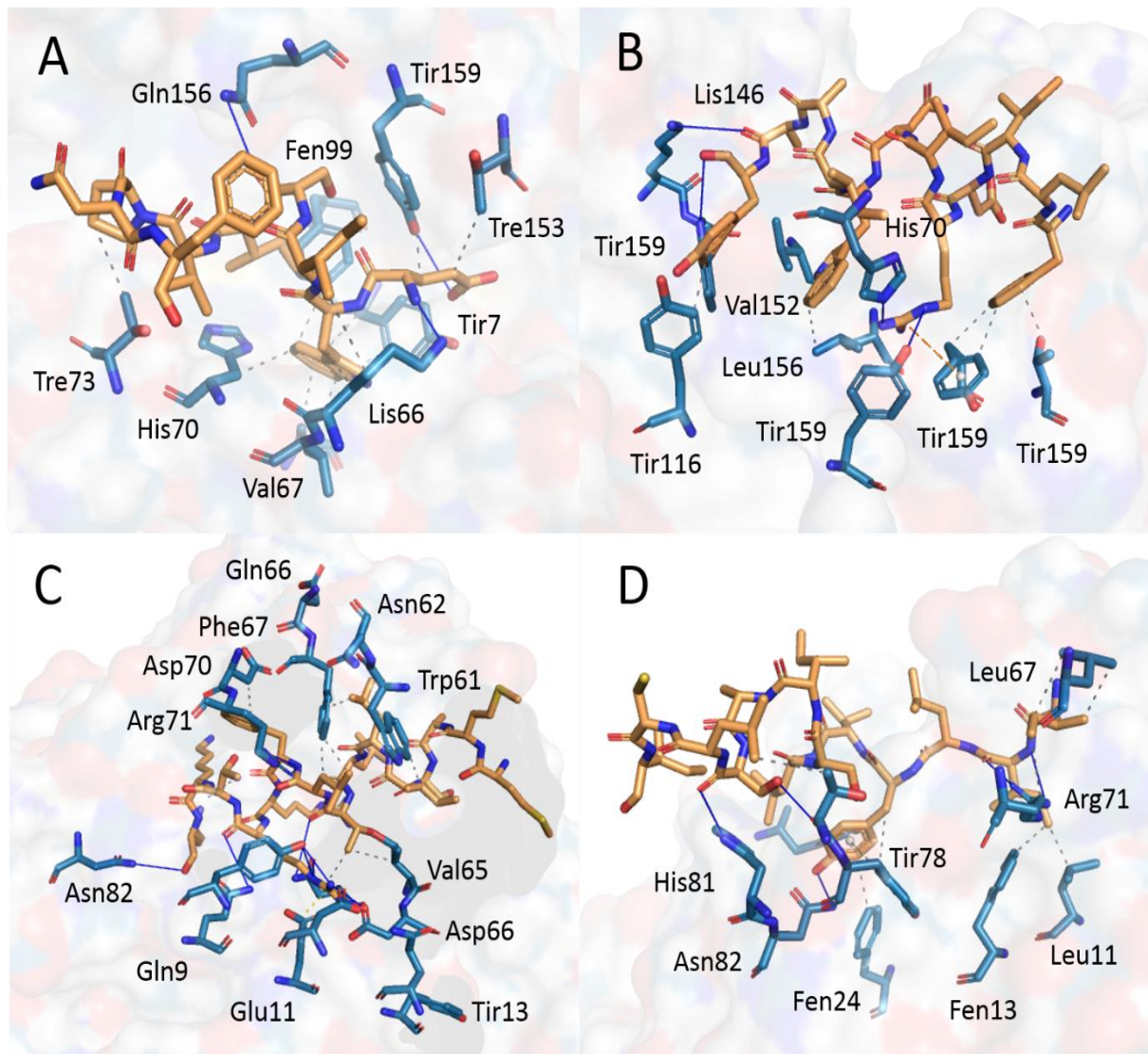
### 5.7.1 Interação entre peptídeos epitópicos e MHC classes I e II

As análises de docagem molecular dos epítipos preditos com seus respectivos receptores MHC-I e MHC-II revelaram interações moleculares favoráveis e afinidade energética aos receptores; vários resíduos dos sítios de ligação formaram ligações de hidrogênio e contatos hidrofóbicos em relação às proteínas multi-epítipo.

No complexo formado entre o epítipo EWISIVPNF obtido a partir da proteína F e o supertipo MHC-I A24 (PDB ID: 3VXR), os resíduos Lys45, Gln156 e Tyr159 formaram

ligações de hidrogênio e contatos hidrofóbicos com vários resíduos no sítio de ligação, de forma semelhante com os epítomos de HIV-1 cristalográficos (Figura 15A). Da mesma forma, o epítomo FLIDRINWI obtido a partir da proteína G formou ligações de hidrogênio com os resíduos His70, Tyr99, Lys146, Trp147 e Val152 localizados no sítio de ligação do supertipo A2 MHC-I (PDB ID: 3VXR), bem como contatos hidrofóbicos com os resíduos Tyr159 e Thr163 de forma semelhante, o que ocorre com o receptor cristalográfico MART-1 (Figura 15B). Em relação ao MHC-II DRB5\*01:01 (PDB ID: 1FV1) e DRB1\*07:01 (PDB ID: 3PDO), os epítomos MMASILLTLFRRTKK e IILYVLSIASLCIGL também formaram ligações de hidrogênio semelhantes àquelas formadas por ligantes cristalográficos; o primeiro formou interações com Tyr13, Val65 e Phe67; o último formou ligações de hidrogênio com os resíduos Phe13, Leu6 e Tyr78 (Figuras 15C e 15D). A descrição completa das energias de ligação obtidas pelo docking molecular entre os peptídeos previstos como epítomos e os receptores MHC-I e -II são mostradas no Apêndice H.

**Figura 15** - Interações intermoleculares obtidas por docagem molecular entre moléculas do MHC e epítotos. preditos. A) MHC-I supertipo A24 complexado com epítoto EWISIVPNF. B) MHC-I supertipo A2 complexado com epítoto FLITRINWI. C) MHC-II molécula DRB5\*01:01 complexada com epítoto MMASILLTLFRRTKK. D) MHC-II molécula DRB\*07:01 em complexo com epítoto IILYVLSIASLCIGL.

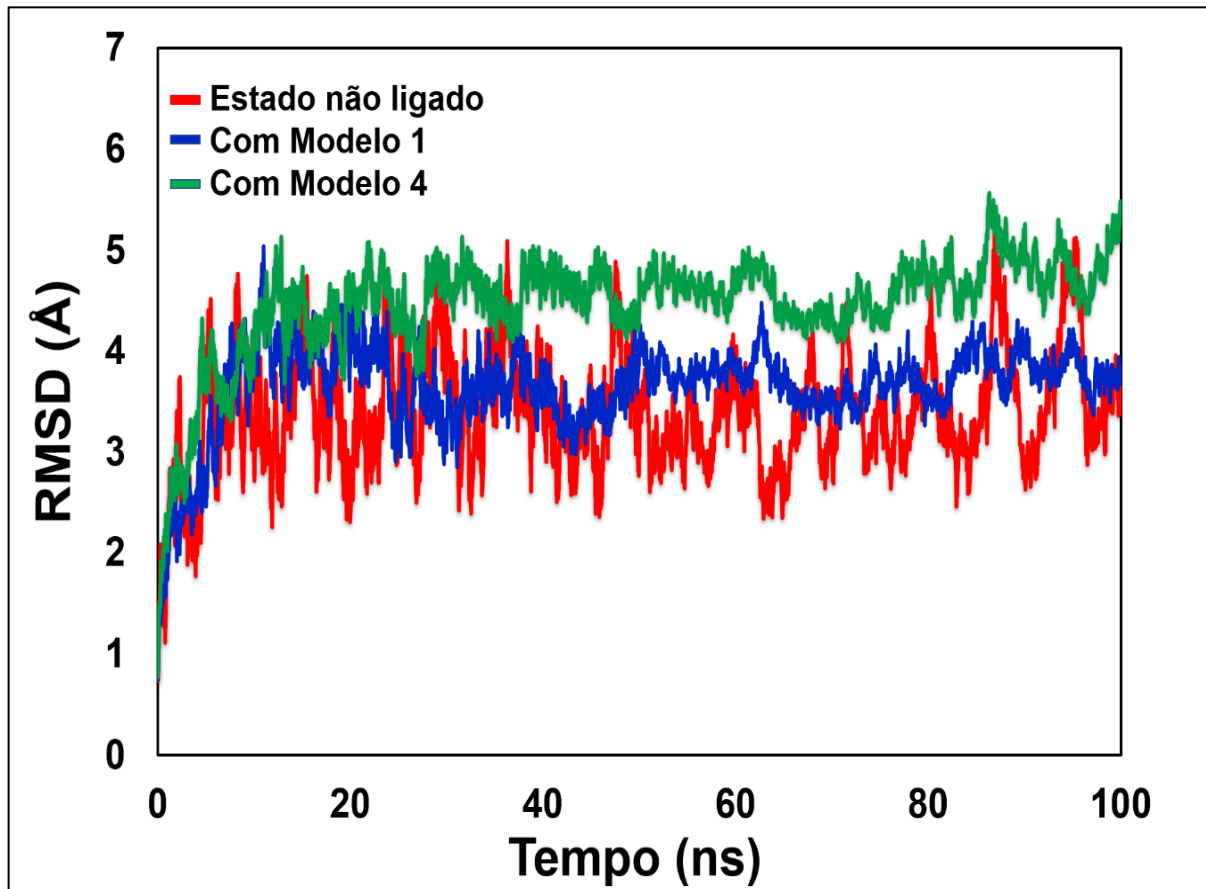


Fonte: Elaborado pelo Autor.

## 5.8 Dinâmica Molecular

Para analisar a estabilidade e as interações intermoleculares dos complexos formados entre os modelos 1 e 4 e o receptor *toll-like 3*, foram realizadas simulações de DM com um tempo total de 100ns. A análise dos valores RMSD da estrutura TLR3 no estado não ligado ao longo da trajetória de DM demonstra uma estrutura estável com ligeiras mudanças conformacionais com oscilações em torno de 2,98 Å a 5,52 Å; e um RMSD médio de 3,37Å. Os modelos 1 e 4 complexados com a estrutura TLR3 mostraram um RMSD médio de 3,47 e 5,16Å, respectivamente (Figura 16).

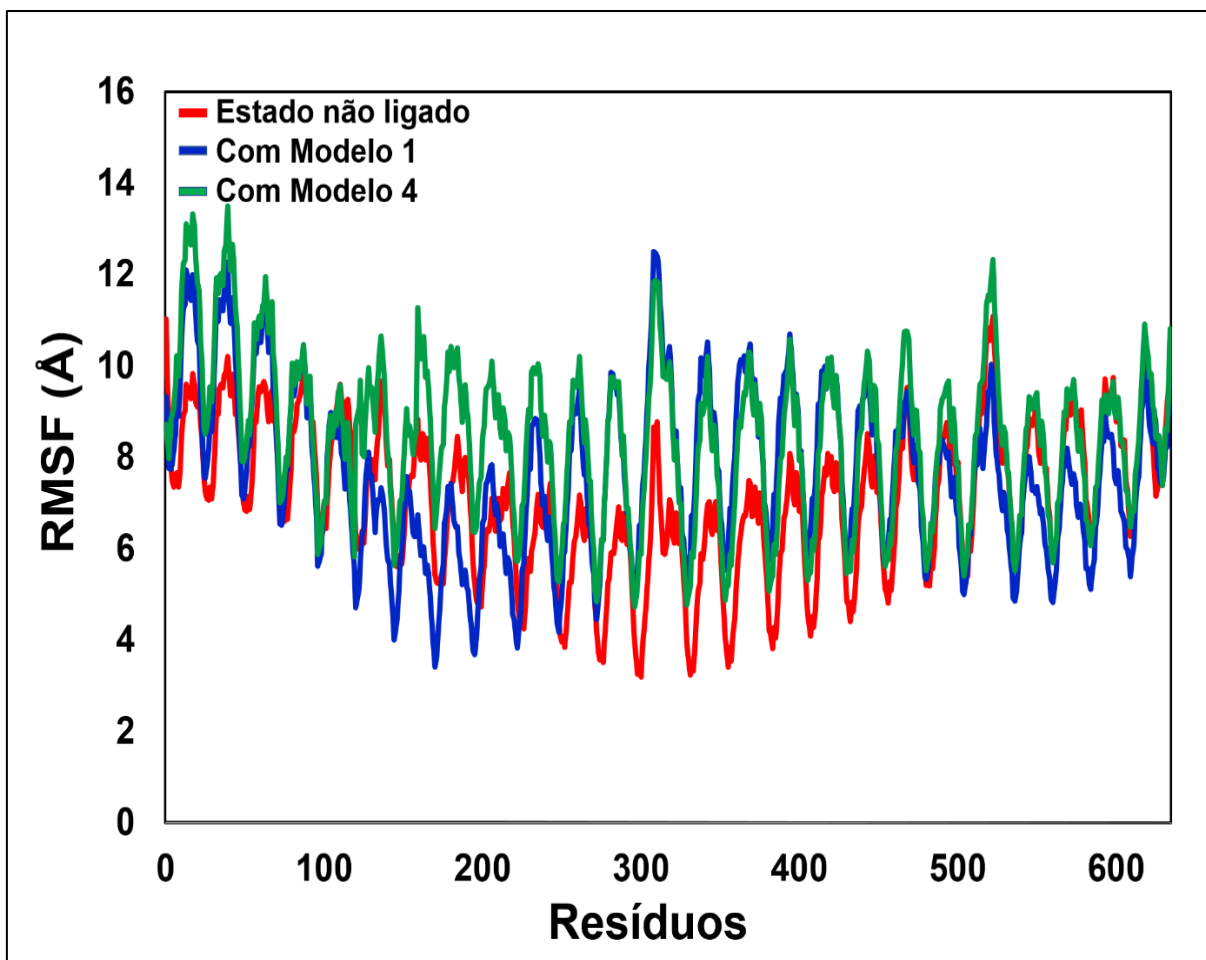
**Figura 16** - Gráfico de RMSD do receptor TLR3 no estado não ligado (livre de antígeno) e no estado ligado (complexado com os modelos multipéptido) obtidos ao longo de 100 ns de simulação de DM.



Fonte: Elaborado pelo Autor.

O gráfico de RMSF do TLR3 complexado com as proteínas multi-epítipo mostrou que os resíduos do sítio de ligação exibiram pouca variação em sua mobilidade quando comparados com o estado ilimitado (livre de antígeno), o que indica que os epítipos têm uma afinidade de ligação estável e estabilidade conformacional. A análise dos valores de RMSF das proteínas de múltiplos epítipos sobre a simulação de DM indicou que os resíduos do modelo 1 mostraram menos mobilidade (Figura 17).

**Figura 17-** Gráfico de RMSF do receptor TLR3 no estado não ligado (livre de antígeno) e no estado ligado (complexado com os modelos de multiepítotos) obtidos ao longo de 100 ns de simulação de DM.



Fonte: Elaborado pelo Autor.

Dessa forma, os complexos entre o TLR3 e os modelos projetados 1 e 4 mostraram-se estáveis ao longo da trajetória das simulações de DM, o que é um indicativo adicional da afinidade em relação ao receptor imunológico. Os valores de energia livre de ligação obtidos pelos métodos MM-PBSA e MM-GBSA reforçam que os antígenos desenhados apresentam grande estabilidade com o receptor TLR3 (Tabela 11).

**Tabela 11** - A energia de ligação livre do receptor TLR3 complexado com as proteínas multi-epítoto projetadas.

Complexos	$\Delta G_{\text{PBSA}}$ (kcal.mol <sup>-1</sup> )	$\Delta G_{\text{GBSA}}$ (kcal.mol <sup>-1</sup> )
Complexo 1	- 57,30 ± 1,07	- 118,62 ± 1,71
Complexo 4	- 65,22 ± 1,58	- 99,07 ± 2,50

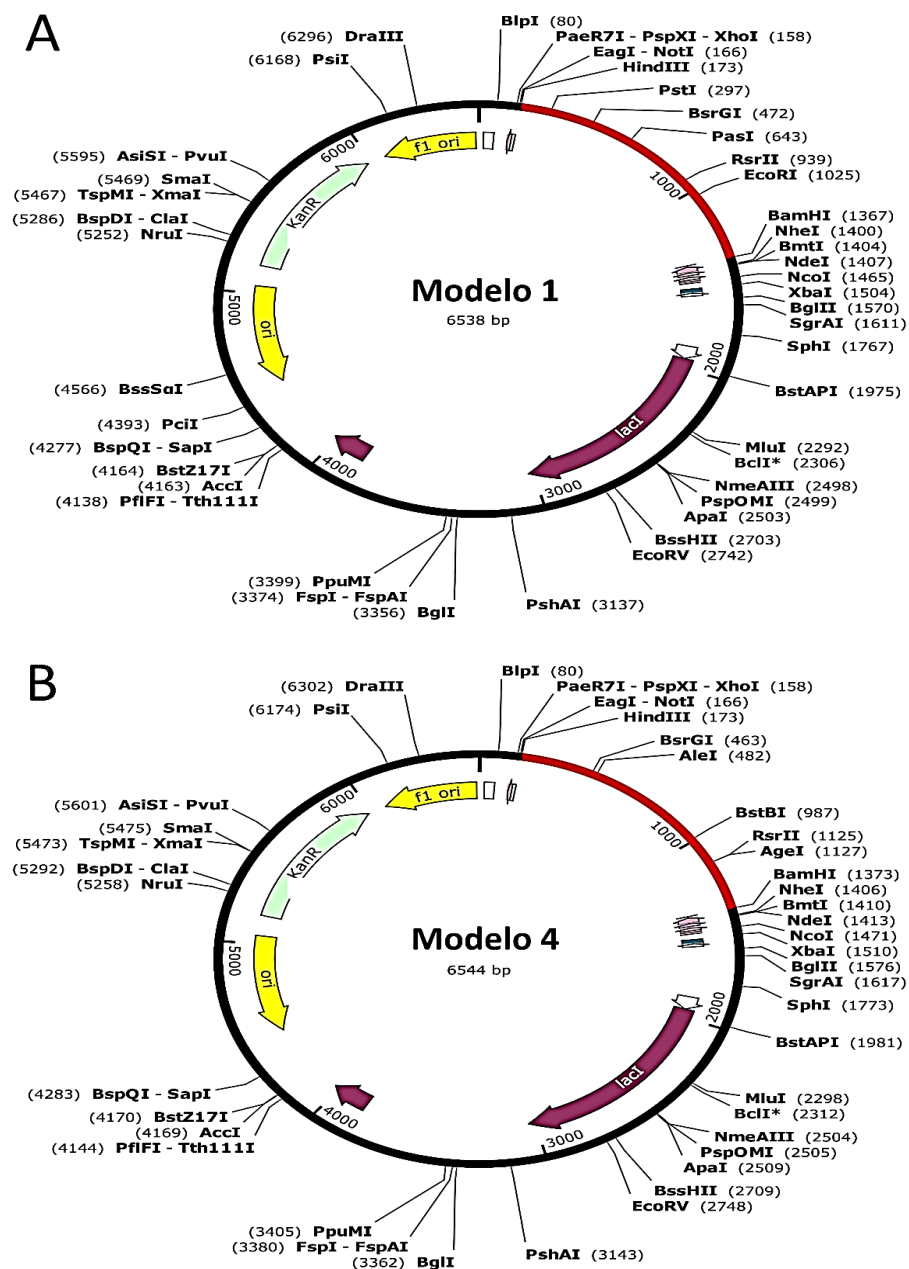
Fonte: Elaborado pelo Autor.



## 5.9 Clonagem *in silico*

As seqüências de cDNAs otimizadas dos modelos selecionados 1 e 4 foram geradas pelo servidor JCAT usando *E. coli* como sistema de expressão de referência (Apêndice I). Os fragmentos de DNA da construção, digeridos por BamHI/HindIII, foram inseridos nos locais de restrição BamHI/HindIII do vetor pET28a + digerido/desfosforilado. Os construtos foram marcados com 6X marcadores de histidina na extremidade 3' do cDNA (Figura 18).

**Figura 18** - Esquema de expressão *in silico*. A porção vermelha representa a inserção de proteínas de múltiplos epítomos no vetor de expressão pET28a (+) (preto). A) Modelo 1. B) Modelo 4.



Fonte: Elaborado pelo Autor.

As sequências de cDNA otimizadas dos modelos 1 e 4 exibiram teor de GC% igual a 59,38% e 58,41%, respectivamente. Ambos os parâmetros mostraram taxas favoráveis para a expressão de proteínas em um hospedeiro bacteriano. Estudos mostraram que o teor de GC% deve ser de 30% a 70% (GIBSON, 2012). A análise do índice de adaptação de códons (CAI) mostrou que os modelos contêm valores de 0,85 e 0,91 com raros códons em tandem de 0%, respectivamente. As pontuações obtidas indicaram que a expressão da proteína no hospedeiro não deve ser prejudicada pela frequência de uso de códons de sequências. Além disso, a frequência obtida de códons tandem raros indicou que as sequências de cDNA projetadas serão potencialmente expressas no hospedeiro bacteriano quando comparadas com os valores de referência (<30%). Códons em tandem raros podem impedir a expressão de cDNA ou interromper o mecanismo de tradução. Portanto, espera-se que cDNAs otimizados sejam altamente expressos em linhas celulares de hospedeiros bacterianos.

## 6 CONSIDERAÇÕES FINAIS

O vírus Nipah é um dos vírus zoonóticos emergentes mais perniciosos, cuja expansão da infecção viral através de surtos representa um sério perigo à saúde pública mundial. A taxa de letalidade que pode chegar a 91% e sequelas neurológicas de longo prazo, mesmo após a sobrevivência à doença, impõem a necessidade do desenvolvimento de uma vacina eficaz contra esse vírus. Aqui, foram projetados modelos de proteínas multi-epítipo baseados no proteoma viral, usando diferentes ferramentas e servidores disponíveis.

Inicialmente, foi possível verificar a dissimilaridade de sequências das proteínas virais em relação a proteínas humanas e a potencial antigenicidade das proteínas virais, especialmente V, W e P, o que é também atestado por evidências experimentais disponíveis. Ainda, o mapeamento dos epítipos mais imunogênicos mostrou que esses estão, em sua maioria, com localização nas regiões externas dos modelos proteicos tridimensionais.

Neste estudo, foi proposto o desenho de dois modelos de proteínas multi-epítipo contra NiV que consistem em epítipos virais atóxicos e não alergênicos reconhecidos por linfócitos T citotóxicos, linfócitos T auxiliares e linfócitos B. Os epítipos selecionados foram validados por métodos *in silico* quanto à interação molecular com receptores do complexo principal de histocompatibilidade classes I e II. A cobertura populacional dos modelos projetados superou 98% da população mundial. Verificou-se que ambos continham epítipos IFN- $\gamma$  e epítipos descontínuos de células B. Através da ferramenta C-QUARK foi possível prever e validar a estrutura tridimensional das proteínas desenhadas de forma satisfatória.

Além disso, os modelos multi-epítipo 1 e 4 também mostraram interação estável com o receptor TLR3, analisada através de docagem molecular, simulações de dinâmica molecular e cálculos de energia livre de ligação. Com base no projeto e na validação *in silico* dos construtos, observa-se que a aplicação dos mesmos como componentes vacinais potencialmente induza uma resposta imune humoral e mediada por células. Foram obtidas evidências de que os cDNAs baseados em códons dos respectivos modelos multi-epítipo têm uma expressão potencialmente elevada em *E. coli*; portanto, poderiam ser clonados e expressos em nível de laboratório para ensaios *in vivo* em camundongos que expressam HLA humanizado para estudos posteriores.

Este estudo apresenta proteínas candidatas a componentes de vacina que, portanto, podem ser uma solução altamente eficaz para a infecção do vírus Nipah. Tal natureza profilática poderá ser atestada através de estudos computacionais adicionais e, posteriormente, testes pré-clínicos.

## REFERÊNCIAS

- AALBERSE, Rob C.; STADLER, Beda M. In silico predictability of allergenicity: from amino acid sequence via 3-D structure to allergenicity. **Molecular nutrition & food research**, v. 50, n. 7, p. 625-627, 2006.
- ABBAS, Abul K.; LICHTMAN, Andrew H.; PILLAI, Shiv. **Cellular and molecular immunology E-book**. Elsevier Health Sciences, 2012.
- AGUILAR, H. C. et al. Paramyxovirus Glycoproteins and the Membrane Fusion Process. **Current Clinical Microbiology Reports**, v. 3, n. 3, p. 142–154, 5 set. 2016.
- AGUILAR, J. C.; RODRIGUEZ, E. G. Vaccine adjuvants revisited. **Vaccine**, v. 25, n. 19, p. 3752-3762, 2007.
- ALI, Mudassar et al. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. **Scientific reports**, v. 7, n. 1, p. 1-13, 2017.
- ALI, M. T.; MORSHED, M. M.; HASSAN, F. A Computational Approach for Designing a Universal Epitope-Based Peptide Vaccine Against Nipah Virus. **Interdisciplinary Sciences: Computational Life Sciences**, 2015.
- ARANKALLE, V. A. et al. Genomic characterization of nipah virus, West Bengal, India. **Emerging Infectious Diseases**, 2011.
- ARORA, G.; MISRA, R.; SAJID, A. Model Systems for Pulmonary Infectious Diseases: Paradigms of Anthrax and Tuberculosis. **Current Topics in Medicinal Chemistry**, 2017.
- ARTIMO, P. et al. ExPASy: SIB bioinformatics resource portal. **Nucleic Acids Research**, 2012.
- ARUNKUMAR, G. et al. Outbreak Investigation of Nipah Virus Disease in Kerala, India, 2018. **The Journal of Infectious Diseases**, 2018.
- BACKERT, Linus; KOHLBACHER, Oliver. Immunoinformatics and epitope prediction in the age of genomic medicine. **Genome medicine**, v. 7, n. 1, p. 119, 2015.
- BASLER, C. F. Nipah and hendra virus interactions with the innate immune system. **Current Topics in Microbiology and Immunology**, 2012.
- BATEMAN, A. et al. UniProt: The universal protein knowledgebase. **Nucleic Acids Research**, 2017.
- BERMAN, H. M. et al. The Protein Data Bank archive as an open data resource. **Journal of Computer-Aided Molecular Design**, v. 28, n. 10, p. 1009–1014, 2014.
- BHARAJ, P. et al. The Matrix Protein of Nipah Virus Targets the E3-Ubiquitin Ligase TRIM6 to Inhibit the IKK $\epsilon$  Kinase-Mediated Type-I IFN Antiviral Response. **PLoS Pathogens**, 2016.
- BIASINI, M. et al. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. **Nucleic Acids Research**, 2014.

- BIRYUKOV, Sergei; STOUTE, José A. The Complement System. In: **Complement Activation in Malaria Immunity and Pathogenesis**. Springer, Cham, 2018. p. 1-29.
- BLUM, J. S.; WEARSCH, P. A.; CRESSWELL, P. Pathways of Antigen Processing. **Annual Review of Immunology**, 2013.
- BOSSART, K. N. et al. A neutralizing human monoclonal antibody protects against lethal disease in a new ferret model of acute Nipah virus infection. **PLoS Pathogens**, 2009.
- BOSSART, K. N. et al. A Hendra virus G glycoprotein subunit vaccine protects African green monkeys from Nipah virus challenge. **Science Translational Medicine**, 2012.
- BRENNICK, Cory A. et al. Neoepitopes as cancer immunotherapy targets: key challenges and opportunities. **Immunotherapy**, v. 9, n. 4, p. 361-371, 2017.
- BRERETON, R. G. Steepest Ascent, Steepest Descent, and Gradient Methods. In: **Comprehensive Chemometrics**, 2009.
- BRODER, Christopher C. et al. A treatment for and vaccine against the deadly Hendra and Nipah viruses. **Antiviral research**, v. 100, n. 1, p. 8-13, 2013.
- BRODER, Christopher C.; WONG, Kum Thong. Henipaviruses. In: **Neurotropic Viral Infections**. Springer, Cham, 2016. p. 45-83.
- BRUHN, J. F. et al. Crystal Structure of the Nipah Virus Phosphoprotein Tetramerization Domain. **Journal of Virology**, 2014.
- BUI, H. H. et al. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. **BMC Bioinformatics**, 2007.
- BYRUM, S. D.; TAVERNA, S. D.; TACKETT, A. J. Post-translational modifications. In: **Epigenetics, Environment, and Genes**. Apple Academic Press, 2013. p. 193–203.
- CALIS, J. J. A. et al. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. **PLoS Computational Biology**, 2013.
- CARBO, A. et al. Systems modeling of the role of interleukin-21 in the maintenance of effector CD4+T cell responses during chronic Helicobacter pylori infection. **mBio**, 2014.
- CARUGO, O.; DJINOVIĆ-CARUGO, K. A proteomic Ramachandran plot (PRplot). **Amino Acids**, v. 44, n. 2, p. 781–790, 2013.
- CHADHA, M. S. et al. Nipah virus-associated encephalitis outbreak, Siliguri, India. **Emerging infectious diseases**, 2006.
- CHAKRABORTY, Hirak Jyoti et al. Protein Structure Prediction. In: **Applying Big Data Analytics in Bioinformatics and Medicine**. IGI Global, 2018. p. 48-79.
- CHANG, L. Y. et al. Human neuronal cell protein responses to Nipah virus infection. **Virology Journal**, 2007.
- CHAUHAN, V. et al. Designing a multi-epitope based vaccine to combat Kaposi Sarcoma

utilizing immunoinformatics approach. **Scientific Reports**, 2019.

CHEN, Chuming; HUANG, Hongzhan; WU, Cathy H. Protein bioinformatics databases and resources. In: **Protein Bioinformatics**. Humana Press, New York, NY, 2017. p. 3-39.

CHEN, X.; ZARO, J. L.; SHEN, W.-C. Fusion protein linkers: property, design and functionality. **Advanced drug delivery reviews**, 2013.

CHEN, Xiaoying; ZARO, Jennica L.; SHEN, Wei-Chiang. Fusion protein linkers: property, design and functionality. **Advanced drug delivery reviews**, v. 65, n. 10, p. 1357-1369, 2013.

CHENG, J. et al. SCRATCH: A protein structure and structural feature prediction server. **Nucleic Acids Research**, 2005.

CHIN, J. X.; CHUNG, B. K. S.; LEE, D. Y. Codon Optimization OnLine (COOL): A web-based multi-objective optimization platform for synthetic gene design. **Bioinformatics**, v. 30, n. 15, p. 2210–2212, 2014.

CHING, P. K. G. et al. Outbreak of henipavirus infection, Philippines, 2014. **Emerging Infectious Diseases**, 2015.

CHONG, H. T. et al. Treatment of acute Nipah encephalitis with ribavirin. **Annals of Neurology**, 2001.

CHOU, Kuo-Chen. Structural bioinformatics and its impact to biomedical science and drug discovery. In: **Frontiers in medicinal chemistry**. Bentham Science Publishers, 2006. p. 455-502.

CHUA, K. B. et al. Fatal encephalitis due to Nipah virus among pig-farmers in Malásia. **Lancet**, 1999.

CHUA, K. B. Nipah virus: A recently emergent deadly paramyxovirus. **Science**, 2000.

CHUA, K. B. et al. The presence of nipah virus in respiratory secretions and urine of patients during an outbreak of nipah virus encephalitis in Malásia. **Journal of Infection**, 2001.

CHUNG, S. et al. Impact of linker strain and flexibility in the design of a fragment-based inhibitor. **Nature Chemical Biology**, v. 5, n. 6, p. 407–413, 2009.

CLAYTON, B. A. et al. Transmission routes for Nipah virus from Malásia and Bangladesh. **Emerging Infectious Diseases**, 2012.

CLAYTON, Bronwyn Anne. Nipah virus: transmission of a zoonotic paramyxovirus. **Current opinion in virology**, v. 22, p. 97-104, 2017.

CLAYTON, B. A.; WANG, L. F.; MARSH, G. A. Henipaviruses: An Updated Review Focusing on the Pteropid Reservoir and Features of Transmission. **Zoonoses and Public Health**, 2013.

CRAGG, G. M.; GROTHAUS, P. G.; NEWMAN, D. J. New horizons for old drugs and drug leads. **Journal of Natural Products**, v. 77, n. 3, p. 703–723, 2014.

CRESSWELL, Peter et al. Mechanisms of MHC class I-restricted antigen processing and cross-presentation. **Immunological reviews**, v. 207, n. 1, p. 145-157, 2005.

CROSER, E. L.; MARSH, G. A. The changing face of the henipaviruses. **Veterinary Microbiology**, 2013.

DA COSTA, K. S. et al. Structural analysis of viral infectivity factor of HIV type 1 and its interaction with A3G, EloC and EloB. **PLoS ONE**, v. 9, n. 2, 2014.

DAPKŪNAS, J.; OLECHNOVIČ, K.; VENCLOVAS, Č. Structural modeling of protein complexes: Current capabilities and challenges. **Proteins: Structure, Function and Bioinformatics**, 2019.

DASGUPTA, A. et al. Cowpox Virus Evades CTL Recognition and Inhibits the Intracellular Transport of MHC Class I Molecules. **The Journal of Immunology**, 2014.

DAWES, B. E. et al. Favipiravir (T-705) protects against Nipah virus infection in the hamster model. **Scientific Reports**, 2018.

DEAKIN, J. E. et al. Evolution and comparative analysis of the MHC Class III inflammatory region. **BMC Genomics**, 2006.

DE MOLFETTA, F. A. et al. Docagem and molecular dynamics simulation of quinone compounds with trypanocidal activity. **Journal of Molecular Modeling**, v. 15, n. 10, p. 1175–1184, 2009.

DE VIVO, Marco et al. Role of molecular dynamics and related methods in drug discovery. **Journal of medicinal chemistry**, v. 59, n. 9, p. 4035-4061, 2016.

DEBUYSSCHER, B. L. et al. Single-dose live-attenuated Nipah virus vaccines confer complete protection by eliciting antibodies directed against surface glycoproteins. **Vaccine**, 2014.

DEFANG, G. N. et al. Induction of neutralizing antibodies to Hendra and Nipah glycoproteins using a Venezuelan equine encephalitis virus in vivo expression system. **Vaccine**, 2010.

DELANY, I.; RAPPUOLI, R.; SEIB, K. L. Vaccines, reverse vaccinology, and bacterial pathogenesis. **Cold Spring Harbor Perspectives in Medicine**, v. 3, n. 5, 2013.

DHANDA, S. K. et al. Predicting HLA CD4 immunogenicity in human populations. **Frontiers in Immunology**, 2018.

DIEDERICH, Sandra; MAISNER, ANDREA. Molecular characteristics of the Nipah virus glycoproteins. **Annals of the New York Academy of Sciences**, v. 1102, n. 1, p. 39-50, 2007.

EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic acids research**, 2004.

EIDHAMMER, I. et al. Quantification of Post-Translational Modifications. In: **Computational and Statistical Methods for Protein Quantification by Mass Spectrometry**. John Wiley & Sons, Inc, 2013. p. 244–253.

EL-MANZALAWY, Y.; HONAVAR, V. B Cell Epitope Prediction. In: **Encyclopedia of**

**Systems Biology**. Springer-Verlag New York, 2013. p. 63–67.

ENGLUND, J. A.; CHU, H. Y. Vaccines Against Respiratory Syncytial Virus: The Time Has Come. **The Journal of infectious diseases**, v. 215, n. 1, p. 4–7, 2017.

ESCAFFRE, O. et al. Henipavirus Pathogenesis in Human Respiratory Epithelial Cells. **Journal of Virology**, 2013.

FERRON, François et al. A practical overview of protein disorder prediction methods. **Proteins: Structure, Function, and Bioinformatics**, v. 65, n. 1, p. 1-14, 2006.

FISER, A.; SALI, A. Modeller: generation and refinement of homology-based protein structure models. **Methods in enzymology**, v. 374, p. 461–91, jan. 2003.

FRANKE, E. D. et al. Pan DR binding sequence provides T-cell help for induction of protective antibodies against Plasmodium yoelii sporozoites. **Vaccine**, v. 17, n. 9–10, p. 1201–1205, 1999.

FURUSE, Y.; SUZUKI, A.; OSHITANI, H. Origin of measles virus: Divergence from rinderpest virus between the 11th and 12th centuries. **Virology Journal**, 2010.

GASTEIGER, E. et al. ExPASy: The proteomics server for in-depth protein knowledge and analysis. **Nucleic Acids Research**, 2003.

GASTEIGER, Elisabeth et al. Protein identification and analysis tools on the ExPASy server. In: **The proteomics protocols handbook**. Humana press, 2005. p. 571-607.

GEISBERT, T. W. et al. Therapeutic treatment of Nipah virus infection in nonhuman primates with a neutralizing human monoclonal antibody. **Science Translational Medicine**, 2014.

GENHEDEN, S.; RYDE, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. **Expert Opinion on Drug Discovery**, v. 10, n. 5, p. 449–461, 2015.

GIBSON, D. G. Gene Synthesis Methods. **Genscript**, 2012.

GIULIANI, M. M. et al. A universal vaccine for serogroup B meningococcus. **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 29, p. 10834–10839, 2006.

GOH, K. J. et al. Clinical Features of Nipah Virus Encephalitis among Pig Farmers in Malásia. **New England Journal of Medicine**, 2002.

GOLDBERG, T. L. Forest Fragmentation as Cause of Bacterial Transmission among Nonhuman Primates, Humans, and Livestock, Uganda. **Emerging Infectious Diseases**, 2008.

GOLDSMITH, C. S. et al. Elucidation of Nipah virus morphogenesis and replication using ultrastructural and molecular approaches. **Virus Research**, 2003.

GREENBAUM, J. et al. Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes. **Immunogenetics**, v. 63, n. 6, p. 325–335, 9 jun. 2011.

GROTE, A. et al. JCat: A novel tool to adapt codon usage of a target gene to its potential



expression host. **Nucleic Acids Research**, 2005.

GUILLAUME, V. et al. Nipah Virus: Vaccination and Passive Protection Studies in a Hamster Model. **Journal of Virology**, 2004.

GUPTA, S. et al. In Silico Approach for Predicting Toxicity of Peptides and Proteins. **PLoS ONE**, 2013.

GURLEY, E. S. et al. Person-to-person transmission of Nipah virus in a Bangladeshi community. **Emerging Infectious Diseases**, 2007.

HALPIN, K. et al. Pteropid bats are confirmed as the reservoir hosts of henipaviruses: A comprehensive experimental study of virus transmission. **American Journal of Tropical Medicine and Hygiene**, 2011.

HARCOURT, B. H. et al. Molecular characterization of Nipah virus, a newly emergent paramyxovirus. **Virology**, 2000.

HARCOURT, B. H. et al. Genetic characterization of Nipah virus, Bangladesh, 2004. **Emerging Infectious Diseases**, 2005.

HASAN, M. A. et al. A comprehensive immunoinformatics and target site study revealed the corner-stone toward Chikungunya virus treatment. **Molecular Immunology**, 2015.

HAYMAN, Colin M.; HERMANS, Ian F.; PAINTER, Gavin F. Increased efficacy of NKT cell-adjuvanted peptide vaccines through chemical conjugation. In: **Coupling and Decoupling of Diverse Molecular Units in Glycosciences**. Springer, Cham, 2018. p. 309-335.

HE, R. et al. Efficient control of chronic LCMV infection by a CD4 T cell epitope-based heterologous prime-boost vaccination in a murine model. **Cellular and Molecular Immunology**, 2018.

HEINECKE, Alexander et al. Molecular dynamics simulation. In: **Supercomputing for Molecular Dynamics Simulations**. Springer, Cham, 2015. p. 11-29.

HEO, L.; PARK, H.; SEOK, C. GalaxyRefine: Protein structure refinement driven by side-chain repacking. **Nucleic acids research**, 2013.

HOFACKER, I. L. Vienna RNA secondary structure server. **Nucleic Acids Research**, 2003.

HOFFMANN, M. et al. Tetherin Inhibits Nipah Virus but Not Ebola Virus Replication in Fruit Bat Cells. **Journal of Virology**, 2018.

HOSSAIN, M. J. et al. Clinical Presentation of Nipah Virus Infection in Bangladesh. **Clinical Infectious Diseases**, 2008.

HOWELL, M.; SHEPHERD, M. **The immune system** *Anaesthesia and Intensive Care Medicine*, 2018.

HSU, V. P. et al. Nipah virus encephalitis reemergence, Bangladesh. **Emerging Infectious Diseases**, 2004.

HUANG, M. et al. Determination of a phosphorylation site in Nipah virus nucleoprotein and its involvement in virus transcription. **Journal of General Virology**, 2011.

ICDDR,B. Nipah Encephalitis Outbreak Over Wide Area of Western Bangladesh, 2004. **ICDDR,B Health and Science Bulletin**, 2004.

IKRAM, A. et al. Exploring NS3/4A, NS5A and NS5B proteins to design conserved subunit multi-epitope vaccine against HCV utilizing immunoinformatics approaches. **Scientific Reports**, 2018.

ISLAM, M. S. et al. Nipah virus transmission from bats to humans associated with drinking traditional liquor made from date palm sap, Bangladesh, 2011-2014. **Emerging Infectious Diseases**, 2016.

IVANCIUC, Ovidiu; SCHEIN, Catherine H.; BRAUN, Werner. SDAP: database and computational tools for allergenic proteins. **Nucleic acids research**, v. 31, n. 1, p. 359-362, 2003.

JARDETZKY, Theodore S.; LAMB, Robert A. Activation of paramyxovirus membrane fusion and virus entry. **Current opinion in virology**, v. 5, p. 24-33, 2014.

JENSEN, M. R. et al. STRUCTURAL DESCRIPTION OF THE NIPAH VIRUS PHOSPHOPROTEIN AND ITS INTERACTION WITH STAT1. **Biophysical Journal**, 2020.

JESPERSEN, M. C. et al. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. **Nucleic Acids Research**, 2017a.

JESPERSEN, M. C. et al. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. **Nucleic Acids Research**, v. 45, n. W1, p. W24–W29, 2017b.

JOHNSON, M. et al. NCBI BLAST: a better web interface. **Nucleic acids research**, 2008.

JONES, K. E. et al. Global trends in emerging infectious diseases. **Nature**, 2008.

KAHL, G. Codon optimization. In: **The Dictionary of Genomics, Transcriptomics and Proteomics**. Wiley-VCH, 2015. p. 108–119.

KAISER, W. J. et al. Toll-like receptor 3-mediated necrosis via TRIF, RIP3, and MLKL. **Journal of Biological Chemistry**, 2013.

KANDUC, D. Peptide cross-reactivity: The original sin of vaccines. **Frontiers in Bioscience - Scholar**, 2012.

KANZLER, Holger et al. Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. **Nature medicine**, v. 13, n. 5, p. 552-559, 2007..

KARESH, William B. et al. Ecology of zoonoses: natural and unnatural histories. **The Lancet**, v. 380, n. 9857, p. 1936-1945, 2012.

KARUPIAH, G.; CHAUDHRI, G. IMMUNOLOGY, INFECTION, AND IMMUNITY. **Immunology and Cell Biology**, 2004.

KAUR, Bani Preet; SECORD, Elizabeth. Innate Immunity. **Pediatric clinics of North America**, v. 66, n. 5, p. 905-911, 2019.

KINDT, T.; GOLDSBY, R.; OSBORNE, B. Kuby Immunology 6th ed, ed. **WHFa Company**, 2006.

KLEYWEGT, Gerard J.; VELANKAR, Sameer; PATWARDHAN, Ardan. Structural biology data archiving—where we are and what lies ahead. **FEBS letters**, v. 592, n. 12, p. 2153-2167, 2018.

KONG, D. et al. Newcastle disease virus-vectored Nipah encephalitis vaccines induce B and T cell responses in mice and long-lasting neutralizing antibodies in pigs. **Virology**, 2012.

KONSTANTINOOU, George N. T-Cell Epitope Prediction. In: **Food Allergens**. Humana Press, New York, NY, 2017. p. 211-222.

KR??UTLER, V.; VAN GUNSTEREN, W. F.; H??NENBERGER, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. **Journal of Computational Chemistry**, v. 22, n. 5, p. 501–508, 2001.

KSIAZEK, T. G. et al. Hendra and Nipah Viruses. In: **Manual of Clinical Microbiology, 10th Edition**. ASM Press, 2015. p. 283-295.

KULKARNI, D. D. et al. Nipah virus infection: Current scenario. **Indian Journal of Virology**, 2013.

KUMAR, H.; KAWAI, T.; AKIRA, S. Pathogen recognition by the innate immune system. **International Reviews of Immunology**, 2011.

KUMAR, V. et al. Molecular modeling and in-silico engineering of Cardamom mosaic virus coat protein for the presentation of immunogenic epitopes of *Leptospira* LipL32. **Journal of Biomolecular Structure and Dynamics**, 2016.

KUO, Tiffany et al. The challenges and opportunities for the development of a T-cell epitope-based herpes simplex vaccine. **Vaccine**, v. 32, n. 50, p. 6733-6745, 2014.

LASKOWSKI, R. A. et al. PDBsum: Structural summaries of PDB entries. **Protein Science**, 2018.

LATA, K. S. et al. Exploring Leptospiral proteomes to identify potential candidates for vaccine design against Leptospirosis using an immunoinformatics approach. **Scientific Reports**, 2018.

LATA, Sneha; BHASIN, Manoj; RAGHAVA, Gajendra PS. Application of machine learning techniques in predicting MHC binders. In: **Immunoinformatics**. Humana Press, 2007. p. 201-215.

LEE, J.; FREDDOLINO, P. L.; ZHANG, Y. Ab initio protein structure prediction. In: **From Protein Structure to Function with Bioinformatics: Second Edition**. Springer, 2017. p. 3–35.

LENNERZ, V. et al. Immunologic response to the survivin-derived multi-epitope vaccine EMD640744 in patients with advanced solid tumors. **Cancer Immunology, Immunotherapy**,

2014.

LI, Ziqiang et al. The generation of antibody diversity through somatic hypermutation and class switch recombination. **Genes & development**, v. 18, n. 1, p. 1-11, 2004.

LI, Z. et al. Beilong virus, a novel paramyxovirus with the largest genome of non-segmented negative-stranded RNA viruses. **Virology**, 2006.

LIN, X. et al. Chimerically fused antigen rich of overlapped epitopes from latent membrane protein 2 (LMP2) of Epstein-Barr virus as a potential vaccine and diagnostic agent. **Cellular and Molecular Immunology**, 2016.

LIONTA, E. et al. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. **Current Topics in Medicinal Chemistry**, 2014.

LITMAN, Gary W.; CANNON, John P.; DISHAW, Larry J. Reconstructing immune phylogeny: new perspectives. **Nature Reviews Immunology**, v. 5, n. 11, p. 866-879, 2005.

LLOYD-SMITH, James O. et al. Epidemic dynamics at the human-animal interface. **science**, v. 326, n. 5958, p. 1362-1367, 2009.

LO, M. K. et al. Characterization of the antiviral and inflammatory responses against Nipah virus in endothelial cells and neurons. **Virology**, 2010.

LO, M. K. et al. Characterization of nipah virus from outbreaks in Bangladesh, 2008-2010. **Emerging Infectious Diseases**, 2012.

LU, C. et al. A novel multi-epitope vaccine from MMSA-1 and DKK1 for multiple myeloma immunotherapy. **British Journal of Haematology**, 2017a.

LU, I. N. et al. Identification of a CD4 T-cell epitope in the hemagglutinin stalk domain of pandemic H1N1 influenza virus and its antigen-driven TCR usage signature in BALB/c mice. **Cellular and Molecular Immunology**, 2017b.

LUBY, S. P. et al. Recurrent zoonotic transmission of Nipah virus into humans, Bangladesh, 2001-2007. **Emerging Infectious Diseases**, 2009.

LUBY, Stephen P. The pandemic potential of Nipah virus. **Antiviral research**, v. 100, n. 1, p. 38-43, 2013.

LUNDGREN, A. et al. Helicobacter pylori-specific CD4<sup>+</sup> CD25<sup>high</sup> regulatory T cells suppress memory T-cell responses to H. pylori in infected individuals. **Infection and Immunity**, 2003.

LYNN, D. J. et al. InnateDB: Facilitating systems-level analyses of the mammalian innate immune response. **Molecular Systems Biology**, 2008.

LZAGUIRRE, J. A. et al. Langevin stabilization of molecular dynamics. **Journal of Chemical Physics**, 2001.

MAGRATH, I. Immunobiology. In: **The Lymphoid Neoplasms 3ed.** [s.l: s.n.].

- MANAVALAN, B. et al. iBCE-EL: A New Ensemble Learning Framework for Improved Linear B-Cell Epitope Prediction. **Frontiers in immunology**, v. 9, p. 1695, 2018.
- MARK, P.; NILSSON, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. **Journal of Physical Chemistry A**, v. 105, n. 43, p. 9954–9960, 2001.
- MATHIEU, C. et al. Nipah Virus Uses Leukocytes for Efficient Dissemination within a Host. **Journal of Virology**, 2011.
- MATHIEU, C. et al. Nonstructural Nipah Virus C Protein Regulates both the Early Host Proinflammatory Response and Viral Virulence. **Journal of Virology**, 2012.
- MAYO, M. A. A summary of taxonomic changes recently approved by ICTV. **Archives of Virology**, 2002.
- MCARTHUR, Donna Behler. Emerging Infectious Diseases. **Nursing Clinics**, v. 54, n. 2, p. 297-311, 2019.
- MCEACHERN, J. A. et al. A recombinant subunit vaccine formulation protects against lethal Nipah virus challenge in cats. **Vaccine**, 2008.
- MCLELLAN, J. S. et al. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. **Science**, v. 342, n. 6158, p. 592–598, 2013.
- MHOD NOR, M. N.; GAN, C. H.; ONG, B. L. Nipah virus infection of pigs in peninsular Malásia. **Revue Scientifique et Technique de l'OIE**, 2016.
- MIRE, C. E. et al. Single injection recombinant vesicular stomatitis virus vaccines protect ferrets against lethal Nipah virus disease. **Virology Journal**, 2013.
- MIRE, C. E. et al. Pathogenic Differences between Nipah Virus Bangladesh and Malásia Strains in Primates: Implications for Antibody Therapy. **Scientific Reports**, 2016.
- MOAL, Iain H. et al. Scoring functions for protein–protein interactions. **Current opinion in structural biology**, v. 23, n. 6, p. 862-867, 2013.
- MOISE, Leonard; MOSS, Steven F.; DE GROOT, Anne S. Moving Helicobacter pylori vaccine development forward with bioinformatics and immunomics. **Expert Review of Vaccines**, v. 11, n. 9, p. 1031-1033, 2012.
- MORSE, Stephen S. et al. Prediction and prevention of the next pandemic zoonosis. **The Lancet**, v. 380, n. 9857, p. 1956-1965, 2012.
- MORTIMER, Stefanie A.; KIDWELL, Mary Anne; DOUDNA, Jennifer A. Insights into RNA structure and function from genome-wide studies. **Nature Reviews Genetics**, v. 15, n. 7, p. 469-479, 2014.
- MUNGALL, B. A. et al. Feline Model of Acute Nipah Virus Infection and Protection with a Soluble Glycoprotein-Based Subunit Vaccine. **Journal of Virology**, 2006.
- NEEFJES, Jacques et al. Towards a systems understanding of MHC class I and MHC class II antigen presentation. **Nature Reviews Immunology**, v. 11, n. 12, p. 823-836, 2011.

- NEGRETE, O. A. et al. EphrinB2 is the entry receptor for Nipah virus, an emergent deadly paramyxovirus. **Nature**, 2005.
- NUCCITELLI, A. et al. Structure-based approach to rationally design a chimeric protein for an effective vaccine against Group B Streptococcus infections. **Proceedings of the National Academy of Sciences of the United States of America**, v. 108, n. 25, p. 10278–10283, 2011.
- OJHA, R. et al. Strategic Development of a Next-Generation Multi-Epitope Vaccine to Prevent Nipah Virus Zoonotic Infection. **ACS Omega**, 2019.
- OYARZÚN, P.; KOBE, B. Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production. **Human Vaccines and Immunotherapeutics**, 2016.
- PALLISTER, J. A. et al. Vaccination of ferrets with a recombinant G glycoprotein subunit vaccine provides protection against Nipah virus disease for over 12 months. **Virology Journal**, 2013.
- PANDEY, R. K.; PRAJAPATI, V. K. Exploring sand fly salivary proteins to design multiepitope subunit vaccine to fight against visceral leishmaniasis. **Journal of Cellular Biochemistry**, v. 120, n. 2, p. 1141–1155, 2019.
- PARASHAR, U. D. et al. Case-Control Study of Risk Factors for Human Infection with a New Zoonotic Paramyxovirus, Nipah Virus, during a 1998–1999 Outbreak of Severe Encephalitis in Malásia. **The Journal of Infectious Diseases**, 2002.
- PARK, M.-S. et al. Newcastle Disease Virus (NDV)-Based Assay Demonstrates Interferon-Antagonist Activity for the NDV V Protein and the Nipah Virus V, W, and C Proteins. **Journal of Virology**, v. 77, n. 2, p. 1501–1511, 15 jan. 2003.
- PATON, N. I. et al. Outbreak of Nipah-virus infection among abattoir workers in Singapore. **Lancet**, 1999.
- PATRONOV, A.; DOYTCHINOVA, I. T-cell epitope vaccine design by immunoinformatics. **Open Biology**, 2013.
- PAUL, S. et al. TepiTool: A pipeline for computational prediction of T cell epitope candidates. **Current Protocols in Immunology**, v. 2016, p. 18.19.1-18.19.24, 2016.
- PEITSCH, M. C.; SCHWEDE, T. Protein Homology Modelling. In: **eLS**. John Wiley & Sons, 2011. p. 1-10.
- PENG, Jian; XU, Jinbo. A multiple-template approach to protein threading. **Proteins: Structure, Function, and Bioinformatics**, v. 79, n. 6, p. 1930-1939, 2011.
- PICKERING, B. S. et al. Protection against henipaviruses in swine requires both, cell-mediated and humoral immune response. **Vaccine**, 2016.
- PIROVANO, Walter; HERINGA, Jaap. Protein secondary structure prediction. In: **Data Mining Techniques for the Life Sciences**. Humana Press, 2010. p. 327-348.
- PIZZA, M. et al. Identification of vaccine candidates against serogroup B meningococcus by

whole-genome sequencing. **Science**, 2000.

PLOQUIN, A. et al. Protection against henipavirus infection by use of recombinant adeno-associated virus-vector vaccines. **Journal of Infectious Diseases**, 2013.

PLOTKIN, Susan L.; PLOTKIN, Stanley A. A short history of vaccination. **Vaccines**, v. 5, p. 1-16, 2004.

PLOWRIGHT, R. K. et al. Prioritizing surveillance of nipah virus in India. **PLoS Neglected Tropical Diseases**, 2019.

PONOMARENKO, J. et al. ElliPro: A new structure-based tool for the prediction of antibody epitopes. **BMC Bioinformatics**, 2008.

PRESCOTT, Joseph et al. The immune response to Nipah virus infection. **Archives of virology**, v. 157, n. 9, p. 1635-1641, 2012.

PULLIAM, J. R. C. et al. Agricultural intensification, priming for persistence and the emergence of Nipah virus: A lethal bat-borne zoonosis. **Journal of the Royal Society Interface**, 2012.

RAHMAN, S. A. et al. Characterization of Nipah virus from naturally infected Pteropus vampyrus bats, Malásia. **Emerging Infectious Diseases**, 2010.

RAJALINGAM, R. et al. Major histocompatibility complex. In: **Transplant Immunology**. John Wiley & Sons, 2016. p. 85-112,

RAMPHUL, K. et al. The Killer Virus Called Nipah: A Review. **Cureus**, 2018.

RAPPUOLI, Rino. Bridging the knowledge gaps in vaccine design. **Nature biotechnology**, v. 25, n. 12, p. 1361-1366, 2007.

RAVEENDRAN, A. V. et al. Nipah virus infection. **Journal of Association of Physicians of India**, 2018.

REN, S. et al. Design and evaluation of a multi-epitope assembly peptide vaccine against Acinetobacter baumannii infection in mice. **Swiss Medical Weekly**, v. 149, n. 23–24, 2019.

REUSKEN, C. et al. WHO Research and Development Blueprint 2018 Annual review of diseases prioritized under the Research and Development Blueprint Informal consultation. **Geneva, Switzerland**, 2018.

REYNES, Jean-Marc et al. Nipah virus in Lyle's flying foxes, Cambodia. **Emerging infectious diseases**, v. 11, n. 7, p. 1042, 2005.

ROCKX, B. et al. A Novel Model of Lethal Hendra Virus Infection in African Green Monkeys and the Effectiveness of Ribavirin Treatment. **Journal of Virology**, 2010.

RODRIGUES, R. L. et al. Prediction of MAYV peptide antigens for immunodiagnostic tests by immunoinformatics and molecular dynamics simulations. **Scientific Reports**, 2019.

RODRÍGUEZ-ORTEGA, M. J. et al. Characterization and identification of vaccine candidate

proteins through analysis of the group A Streptococcus surface proteome. **Nature Biotechnology**, v. 24, n. 2, p. 191–197, 2006.

RODRIGUEZ, J. J.; PARISIEN, J.-P.; HORVATH, C. M. Nipah Virus V Protein Evades Alpha and Gamma Interferons by Preventing STAT1 and STAT2 Activation and Nuclear Accumulation. **Journal of Virology**, 2002.

ROJAS, Manuel et al. Molecular mimicry and autoimmunity. **Journal of Autoimmunity**, v. 95, p. 100-123, 2018.

RUBINSTEIN, N. D.; MAYROSE, I.; PUPKO, T. A machine-learning approach for predicting B-cell epitopes. **Molecular Immunology**, v. 46, n. 5, p. 840–847, 2009.

RUPPRECHT, Charles E.; TURMELLE, Amy; KUZMIN, Ivan V. A perspective on lyssavirus emergence and perpetuation. **Current opinion in virology**, v. 1, n. 6, p. 662-670, 2011.

SAADI, M.; KARKHAH, A.; NOURI, H. R. Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. **Infection, Genetics and Evolution**, 2017.

SAHA, C. K. et al. In silico identification and characterization of common epitope-based peptide vaccine for Nipah and Hendra viruses. **Asian Pacific Journal of Tropical Medicine**, 2017.

SAHA, S.; RAGHAVA, G. P. S. AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. **Nucleic Acids Research**, 2006.

SAINI, S. et al. Computational prediction of miRNAs in Nipah virus genome reveals possible interaction with human genes involved in encephalitis. **Molecular Biology Research Communications**, 2018.

SAKIB, M. S. et al. Prediction of Epitope-Based Peptides for the Utility of Vaccine Development from Fusion and Glycoprotein of Nipah Virus Using in Silico Approach. **Advances in Bioinformatics**, 2014.

SALMASO, V. et al. Combining self- and cross-docking as benchmark tools: the performance of DockBench in the D3R Grand Challenge 2. **Journal of Computer-Aided Molecular Design**, v. 32, n. 1, p. 251–264, 24 jan. 2018.

SALOMON-FERRER, R.; CASE, D. A.; WALKER, R. C. An overview of the Amber biomolecular simulation package. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, 2013.

SATTERFIELD, B. A. et al. The immunomodulating v and W proteins of Nipah virus determine disease course. **Nature Communications**, 2015.

SATTERFIELD, B. A.; DAWES, B. E.; MILLIGAN, G. N. Status of vaccine research and development of vaccines for Nipah virus. **Vaccine**, 2016a.

SATTERFIELD, B. A.; DAWES, B. E.; MILLIGAN, G. N. Status of vaccine research and development of vaccines for Nipah virus. **Vaccine**, v. 34, n. 26, p. 2971–2975, jun. 2016b.



SAYERS, E. W. et al. Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**, 2019.

SAZZAD, H. M. S. et al. Nipah virus infection outbreak with nosocomial and corpse-to-human transmission, Bangladesh. **Emerging Infectious Diseases**, 2013.

SCARSELLI, M. et al. Rational design of a meningococcal antigen inducing broad protective immunity. **Science Translational Medicine**, v. 3, n. 91, 2011.

SCHOMACKER, H.; COLLINS, P. L.; SCHMIDT, A. C. In silico identification of a putative new paramyxovirus related to the Henipavirus genus. **Virology**, 2004.

SCHOUNTZ, Tony. Immunology of bats and their viruses: challenges and opportunities. **Viruses**, v. 6, n. 12, p. 4880-4901, 2014.

SCHOUNTZ, T. et al. Immunological control of viral infections in bats and the emergence of viruses highly pathogenic to humans. **Frontiers in Immunology**, 2017.

SETO, J. et al. Novel Nipah Virus Immune-Antagonism Strategy Revealed by Experimental and Computational Study. **Journal of Virology**, 2010.

SHAN, Y. et al. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. **Journal of Chemical Physics**, 2005.

SHAW, M. L. et al. Nuclear Localization of the Nipah Virus W Protein Allows for Inhibition of both Virus- and Toll-Like Receptor 3-Triggered Signaling Pathways. **Journal of Virology**, 2005.

SHAW, Megan L. Henipaviruses employ a multifaceted approach to evade the antiviral interferon response. **Viruses**, v. 1, n. 3, p. 1190-1203, 2009.

SIDNEY, J. et al. HLA class I supertypes: A revised and updated classification. **BMC Immunology**, v. 9, 2008.

SINGH, H.; ANSARI, H. R.; RAGHAVA, G. P. S. Improved method for linear B-cell epitope prediction using antigen's primary sequence. **PloS one**, 2013.

SINGH, R. K. et al. Nipah virus: epidemiology, pathology, immunobiology and advances in diagnosis, vaccine designing and control strategies - a comprehensive review. **The Veterinary quarterly**, 2019.

ŚLEDŹ, Paweł; CAFLISCH, Amedeo. Protein structure-based drug design: from docagem to molecular dynamics. **Current opinion in structural biology**, v. 48, p. 93-102, 2018.

SRINIVASAN, P. et al. Epitope-based immunoinformatics and molecular docagem studies of nucleocapsid protein and ovarian tumor domain of Crimean-Congo hemorrhagic fever virus. **Frontiers in Genetics**, 2011.

SRIVASTAVA, S. et al. Structural basis of development of multi-epitope vaccine against middle east respiratory syndrome using in silico approach. **Infection and Drug Resistance**, 2018.

- STACHOWIAK, B.; WEINGARTL, H. M. Nipah virus infects specific subsets of porcine peripheral blood mononuclear cells. **PLoS ONE**, 2012.
- STEPHENS, R. S. Chlamydial Genomics and Vaccine Antigen Discovery. **The Journal of Infectious Diseases**, v. 181, n. s3, p. S521–S523, 2000.
- STERN, A. et al. The Evolutionary Pathway to Virulence of an RNA Virus. **Cell**, 2017.
- STONE, J. A. et al. Multiple Strategies Reveal a Bidentate Interaction between the Nipah Virus Attachment and Fusion Glycoproteins. **Journal of Virology**, 2016.
- SUGAI, A. et al. Nipah and Hendra Virus Nucleoproteins Inhibit Nuclear Accumulation of Signal Transducer and Activator of Transcription 1 (STAT1) and STAT2 by Interfering with Their Complex Formation. **Journal of Virology**, 2017.
- TAI, W. et al. Rational Design of Zika Virus Subunit Vaccine with Enhanced Efficacy. **Journal of Virology**, v. 93, n. 17, 2019.
- TAN, C. T. et al. Relapsed and late-onset Nipah encephalitis. **Annals of Neurology**, 2002.
- TENZER, S. et al. Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. **Cellular and Molecular Life Sciences**, 2005.
- THIBAUT, Patricia A. et al. Zoonotic potential of emerging paramyxoviruses: Knowns and unknowns. In: **Advances in virus research**. Academic Press, 2017. p. 1-55.
- THOMAS, Sunil; DILBAROVA, Rima; RAPPUOLI, Rino. Future challenges for vaccinologists. In: **Vaccine Design**. Humana Press, New York, NY, 2016. p. 41-55.
- THORPE, C. et al. Discovery of a vaccine antigen that protects mice from Chlamydia pneumoniae infection. **Vaccine**, v. 25, n. 12, p. 2252–2260, 2007.
- TIAN, W. et al. CASTp 3.0: computed atlas of surface topography of proteins. **Nucleic acids research**, 2018.
- TOM, J. K. et al. **Applications of Immunomodulatory Immune Synergies to Adjuvant Discovery and Vaccine Development Trends in Biotechnology**, 2019.
- TOMAR, N.; DE, R. K. Immunoinformatics: A brief review. **Methods in Molecular Biology**, 2014.
- TOMAR, Shailly; KAUR, Ramanjit; SINGH, Vedita Anand. Structure–Function Relationship of Negative-Stranded Viral RNA Polymerases: Prospectives for Antiviral Therapy. In: **Viral Polymerases**. Academic Press, 2019. p. 43-67.
- USMAN MIRZA, M. et al. Towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins. **Scientific Reports**, 2016.
- VAISHNAV, Nishita et al. Overview of computational vaccinology: vaccine development through information technology. **Journal of applied genetics**, v. 56, n. 3, p. 381-391, 2015.

- VANGURI, V. K. The adaptive immune system. In: Pathobiology of Human Disease. Mc Manus LM, 2014. p. 1-4.
- VERA-VELASCO, N. M. et al. Proteomic composition of Nipah virus-like particles. **Journal of Proteomics**, 2018.
- VITA, R. et al. The immune epitope database (IEDB) 3.0. **Nucleic Acids Research**, 2015.
- WALPITA, P. et al. A VLP-based vaccine provides complete protection against Nipah virus challenge following multiple-dose or single-dose vaccination schedules in a hamster model. **npj Vaccines**, 2017.
- WANG, L. F. et al. The exceptionally large genome of Hendra virus: support for creation of a new genus within the family Paramyxoviridae. **Journal of virology**, 2000.
- WANG, Lin-Fa et al. Molecular biology of Hendra and Nipah viruses. **Microbes and infection**, v. 3, n. 4, p. 279-287, 2001.
- WATKINSON, Ruth E.; LEE, Benhur. Nipah virus matrix protein: expert hacker of cellular machines. **FEBS letters**, v. 590, n. 15, p. 2494-2511, 2016.
- WEINGARTL, H. et al. Invasion of the Central Nervous System in a Porcine Host by Nipah Virus. **Journal of Virology**, 2005.
- WEINGARTL, H. M. et al. Recombinant Nipah Virus Vaccines Protect Pigs against Challenge. **Journal of Virology**, 2006.
- WIEDERSTEIN, M.; SIPPL, M. J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research**, v. 35, n. SUPPL.2, 2007.
- WIETHOELTER, Anke K. et al. Global trends in infectious diseases at the wildlife–livestock interface. **Proceedings of the National Academy of Sciences**, v. 112, n. 31, p. 9662-9667, 2015.
- WILSON, Sarah S.; WIENS, Mayim E.; SMITH, Jason G. Antiviral mechanisms of human defensins. **Journal of molecular biology**, v. 425, n. 24, p. 4965-4980, 2013.
- WOLF, M. C. et al. A broad-spectrum antiviral targeting entry of enveloped viruses. **Proceedings of the National Academy of Sciences of the United States of America**, 2010.
- WOLFE, Nathan D.; DUNAVAN, Claire Panosian; DIAMOND, Jared. Origins of major human infectious diseases. **Nature**, v. 447, n. 7142, p. 279-283, 2007.
- YABUKARSKI, F. et al. Structure of Nipah virus unassembled nucleoprotein in complex with its viral chaperone. **Nature Structural and Molecular Biology**, 2014.
- YAN, Y. et al. HDock: A web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. **Nucleic Acids Research**, v. 45, n. W1, p. W365–W373, 2017.
- YAO, B. et al. Conformational B-Cell Epitope Prediction on Antigen Protein Structures: A Review of Current Algorithms and Comparison with Common Binding Site Prediction

Methods. **PLoS ONE**, 2013.

YEE, P. T. I. et al. Development of live attenuated Enterovirus 71 vaccine strains that confer protection against lethal challenge in mice. **Scientific Reports**, v. 9, n. 1, 2019.

YONEDA, M. et al. Recombinant Measles Virus Vaccine Expressing the Nipah Virus Glycoprotein Protects against Lethal Nipah Virus Challenge. **PLoS ONE**, 2013.

YU, M. et al. The attachment protein of Hendra virus has high structural similarity but limited primary sequence homology compared with viruses in the genus paramyxovirus. **Virology**, 1998.

ZHANG, Y.; SKOLNICK, J. Scoring function for automated assessment of protein structure template quality. **Proteins: Structure, Function, and Bioinformatics**, v. 68, n. 4, p. 1020–1020, 2007.

ZHENG, W. et al. Deep-learning contact-map guided protein structure prediction in CASP13. **Proteins: Structure, Function and Bioinformatics**, v. 87, n. 12, p. 1149–1164, 2019.

ZHOU, P. et al. Type III IFN receptor expression and functional characterisation in the pteropid bat, *Pteropus alecto*. **PLoS ONE**, 2011a.

ZHOU, P. et al. Type III IFNs in Pteropid Bats: Differential Expression Patterns Provide Evidence for Distinct Roles in Antiviral Immunity. **The Journal of Immunology**, 2011b.

ZHOU, P. et al. Contraction of the type I IFN locus and unusual constitutive expression of IFN- $\alpha$  in bats. **Proceedings of the National Academy of Sciences**, 2016.

ZHU, Z. et al. Potent Neutralization of Hendra and Nipah Viruses by Human Monoclonal Antibodies. **Journal of Virology**, 2005.

ZHU, Z. et al. Exceptionally Potent Cross-Reactive Neutralization of Nipah and Hendra Viruses by a Human Monoclonal Antibody. **The Journal of Infectious Diseases**, 2008.

ZUKER, Michael. Mfold web server for nucleic acid folding and hybridization prediction. **Nucleic acids research**, v. 31, n. 13, p. 3406-3415, 2003.

## APÊNDICES

### Apêndice A

**Tabela A1** - Proteomas do vírus Nipah obtidos do banco de dados Uniprot.

Número	Proteoma (ID)	Sequência (pb)	País de origem
1	UP000170143	18,246	Malásia
2	UP000002330	18,246	Malásia
3	UP000128950	18,252	Índia
4	UP000007527	18,246	Malásia
5	UP000008676	18,246	Malásia
6	UP000130871	18,246	Malásia
7	UP000246210	18,167	Bangladesh
8	UP000246966	18,212	Malásia
9	UP000120177	18,252	Bangladesh
10	UP000100567	18,246	Malásia
11	UP000103103	18,252	Bangladesh
12	UP000110983	18,246	Malásia
13	UP000246915	18,214	Malásia
14	UP000136952	18,252	Bangladesh
15	UP000180669	18,373	Malásia

### Apêndice B

**Tabela A2** - Variantes e respectivos supertipos do complexo principal de histocompatibilidade utilizados como parâmetro de predição de epítomos.

MHC (Classe)	Alelo/Molécula	Supertipo
Classe I	A*01:01	A1
	A*02:01	A2
	A*03:01	A3
	A*24:02	A24
	A*26:02	A26
	B*07:02	B7
	B*08:01	B8
	B*27:05	B27
	B*39:01	B39
	B*40:01	B44
	B*58:01	B58
	B*15:01	B62
	DRB1*01:01	DR
	DRB1*03:01	DRB3
	DRB1*04:01	DR4
DRB1*04:05	DR4	
DRB1*07:01	DR	

Classe II	DRB1*08:02	DR4
	DRB1*09:01	DR
	DRB1*11:01	DR
	DRB1*12:01	DR
	DRB1*13:02	DRB3
	DRB1*15:01	DR
	DRB3*01:01	DRB3
	DRB3*02:02	DRB3
	DRB4*01:01	DRB3
	DRB5*01:01	DR
	DPA1*01:01/DPB1*04:01	DP2
	DPA1*01:03/DPB1*02:01	DP2
	DPA1*02:01/DPB1*01:01	DP
	DPA1*02:01/DPB1*05:01	DP
	DPA1*03:01/DPB1*04:02	DP
	DQA1*01:01/DQB1*05:01	DQ
	DQA1*01:02/DQB1*06:02	DQ7
	DQA1*03:01/DQB1*03:02	DQ7
	DQA1*04:01/DQB1*04:02	DQ
	DQA1*05:01/DQB1*02:01	DQ
DQA1*05:01/DQB1*03:01	DQ7	

### Apêndice C

**Tabela A3:** Moléculas do complexo principal de histocompatibilidade com estrutura terciária predita por modelagem de homologia.

HLA molécula	Classe	PDB ID	Resolução	Cobertura	Identidade
DPA1*02:01/DPB1*05:01	Class II	3WEX	2.4Å	71%	99.73%
DPA1*01/DPB1*04:01	Class II	4P5K	2.6Å	79%	95.08%
DQA1*05:01/DQB1*02:01	Class II	4D8P	3.0Å	79%	94.00%
DPA1*02:01/DPB1*01:01	Class II	3WEX	2.4Å	71%	98.08%
DPA1*03:01/DPB1*04:02	Class II	4P5K	2.6Å	79%	95.08%
DQA1*01:02/DQB1*06:02	Class II	6BLQ	1.8Å	70%	73.57%
DQA1*04:01/DQB1*04:02	Class II	4D8P	3.0Å	79%	90.50%
DRB1*03:01	Class II	6CPN	2.0Å	75%	94.21%
A*26:01	Class I	6O9B	2.2Å	77%	92.14%

## Apêndice D

**Esquema A1** - Desenho de sequências de proteínas com múltiplos epítomos. A construção consiste em resíduos de aminoácidos de sequências adjuvantes ( $\beta$ -defensinas) nos terminais N e C. O ligante EAAAK foi adicionado para ligar os epítomos e adjuvantes. Os epítomos de CTL, HTL e linfócitos B lineares foram unidos com ligantes KK, ligantes GPGPG e ligantes AAY, respectivamente.

### Modelo 1

GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKPEAAAKDSKILSAFNTVIALLGPGPGNSEWI  
SIVPNFILVRGPGPGIILYVLSIASLCIGLGPGPGSAVYNNEFYVLCVAVGPGPGGVYNDAFLIDRINWIGP  
GPGFISFIIVEKKRNTYSGPGPGSKILSAFNTVIALLGPGPGAMDEGYFAYAAAYVLFANCISVAAAYLMM  
TRLAVKAAAYEWISIVPNFAAYETLLRTLGYAAAYIVRVYFPILAAAYAQITAGVALAAAYTVNPLVVNWAA  
YKIGLVKGGKKTAGVALYKKIQELLPVSKKISIVPNFIKKVYHCSAVKKYFPAVGFLKKPSKVIKSEAAAK  
GIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKK

### Modelo 2

GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKPEAAAKAMDEGYFAYAAAYVLFANCISVA  
AYLMMTRLAVKAAAYEWISIVPNFAAYETLLRTLGYAAAYIVRVYFPILAAAYAQITAGVALAAAYTVNPLV  
VNWAAAYDSKILSAFNTVIALLGPGPGNSEWISIVPNFILVRGPGPGIILYVLSIASLCIGLGPGPGSAVYNN  
EFYVLCVAVGPGPGGVYNDAFLIDRINWIGPGPGFISFIIVEKKRNTYSGPGPGSKILSAFNTVIALLGPGPG  
PKIGLVKGGKKTAGVALYKKIQELLPVSKKISIVPNFIKKVYHCSAVKKYFPAVGFLKKPSKVIKSEAAA  
KGIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKK

### Modelo 3

GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKPEAAAKKIGLVKGGKKTAGVALYKKIQELL  
PVSKKISIVPNFIKKVYHCSAVKKYFPAVGFLKKPSKVIKSKKDSKILSAFNTVIALLGPGPGNSEWISIVP  
NFILVRGPGPGIILYVLSIASLCIGLGPGPGSAVYNNEFYVLCVAVGPGPGGVYNDAFLIDRINWIGPGPG  
FISFIIVEKKRNTYSGPGPGSKILSAFNTVIALLGPGPGAMDEGYFAYAAAYVLFANCISVAAAYLMMTRL  
AVKAAAYEWISIVPNFAAYETLLRTLGYAAAYIVRVYFPILAAAYAQITAGVALAAAYTVNPLVVNWEAAAK  
GIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKK

### Modelo 4

GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKPEAAAKADKVLEYAPIMKAHAGPGPGAVE  
IISDIGNYVEETGPGPGMMASILLTLFRRTKKGPGPGKRKKIRTIAAYPLGVGPGPGFEFEDEFAGSSSEVIV  
GPGPGNSEWISIVPNFILVRGPGPGSKILSAFNTVIALLGPGPGSHMWRELALAAAYWTQWLQTLYAAAY  
AQITAGVALAAAYAMDEGYFAYAAAYKSRGIPIKKAAYMYLICYGFAAYFLIDRINWIAAAYRASAAAT  
LAAAYIQELLPVSKKYHDHGGEKKLYILRQSKQGDKKELVNDGLKKVLEESIQTKKEGVYNKKTAGVA  
LYEAAAKGIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKK

### Modelo 5

GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKPEAAAKSHMWRELALAAAYWTQWLQTLY  
AAAYAQITAGVALAAAYAMDEGYFAYAAAYKSRGIPIKKAAYMYLICYGFAAYFLIDRINWIAAAYRASAA  
TATLAAAYADKVLEYAPIMKAHAGPGPGAVEIISDIGNYVEETGPGPGMMASILLTLFRRTKKGPGPGKR  
KKIRTIAAYPLGVGPGPGFEFEDEFAGSSSEVIVGPGPGNSEWISIVPNFILVRGPGPGSKILSAFNTVIALLG  
GPGPGIQELLPVSKKYHDHGGEKKLYILRQSKQGDKKELVNDGLKKVLEESIQTKKEGVYNKKTAGV  
ALYEAAAKGIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKK

### Modelo 6

GIGDPVTCLKSGAICHPVFCPRRYKQIGTCGLPGTKCCKKPEAAAKIQELLPVSKKYHDHGGEKKLYILR  
QSKQGDKKELVNDGLKKVLEESIQTKKEGVYNKKTAGVALYKKADKVLEYAPIMKAHAGPGPGAV  
EISDIGNYVEETGPGPGMMASILLTLFRRTKKGPGPGKRKKIRTIAAYPLGVGPGPGFEFEDEFAGSSSEVI





## Apêndice F

**Tabela A5** - Docagem molecular entre epítomos reconhecidos por CTL e moléculas de classe I do complexo principal de histocompatibilidade.

<b>Epítopo</b>	<b>Docagem Score</b>	<b>Receptor/Supertipo</b>	<b>Protein Data Bank ID</b>
WTQWLQTTY	-272.10	A*01:01 (A1)	4UQ3
	-245,3	A*26:01 (A26)	6ENY (Homology modelling)
VLFANCISV	-192.22	A*02:01 (A2)	3QEQ
EWISIVPNF	-210.33	A*24:02 (A24)	3VXR
	-213,24	A*26:01 (A26)	6ENY (Homology modelling)
ETLLRTLGY	-207.68	A*01:01 (A1)	4UQ3
	-235,72	A*26:01 (A26)	6ENY (Homology modelling)
IVRVYFPIL	-207.83	B*08:01 (B8)	1MI5
	-241,1	B*39:01 (B39)	4O2C
	-228,23	B*15:01 (B62)	5TXS
AQITAGVAL	-244,14	B*40:01 (B44)	5IEK
LMMTRLAVK	-209,4	A*03:01 (A3)	6ENY
TVNPLVVNW	-212,1	B*58:01 (B58)	5VWH
AMDEGYFAY	-258.55	A*01:01 (A1)	4UQ3
FLIDRINWI	-246.32	A*02:01 (A2)	3QEQ
SHMWRELAL	-209.94	B*08:01 (B8)	1MI5
	-251,07	B*39:01 (B39)	4O2C
YMYLICYGF	-258.76	A*24:02 (A24)	3VXR
	-237,28	B*15:01 (B62)	5TXS
RASAATATL	-245,38	B*58:01 (B58)	5VWH
	-201,06	B*07:02 (B7)	6AVG
KSRGIPIKK	-224,09	A*03:01 (A3)	6ENY

## Apêndice G

**Tabela A7:** Docagem molecular entre o receptor *toll-like 3* e o modelo de proteína imunogênica 1.

<b>TLR-3 (Átomo)</b>	<b>TLR-3 (Resíduo)</b>	<b>Modelo (Átomo)</b>	<b>Modelo (Resíduo)</b>	<b>Distância</b>	<b>Tipo de interação</b>
OE2	GLU33	NZ	LYS342	2.79	Ligações de hidrogênio
NE2	HIS39	O	LYS36	2.80	Ligações de hidrogênio
ND2	ASN57	OD1	ASN67	2.99	Ligações de hidrogênio
NE2	GLN76	OE1	GLN382	2.67	Ligações de hidrogênio
OE1	GLN174	N	GLY186	3.07	Ligações de hidrogênio
NE2	GLN174	O	GLY186	2.85	Ligações de hidrogênio
OE1	GLU175	NZ	LYS10	2.68	Ligações de hidrogênio
NE	ARG251	OD1	ASP189	2.92	Ligações de hidrogênio
NH2	ARG251	OD2	ASP189	3.06	Ligações de hidrogênio
OH	TYR302	OD1	ASN130	2.98	Ligações de hidrogênio
OE1	GLU363	NZ	LYS25	3.06	Ligações de hidrogênio

OD1	ASN380	NE1	TRP140	2.83	Ligações de hidrogênio
NZ	LYS382	OD2	ASP136	2.77	Ligações de hidrogênio
OE2	GLU460	NE	ARG22	3.18	Ligações de hidrogênio
OE2	GLU33	NZ	LYS342	2.79	Interações hidrofóbicas
NE2	HIS39	O	LYS36	2.80	Interações hidrofóbicas
NE2	GLN174	O	GLY186	2.85	Interações hidrofóbicas
OE1	GLU175	NZ	LYS10	2.68	Interações hidrofóbicas
OE2	GLU175	NZ	LYS10	2.89	Interações hidrofóbicas
OD1	ASN380	NE1	TRP140	2.83	Interações hidrofóbicas
OE1	GLU33	NZ	LYS342	2.79	Pontes salinas
OD2	ASP36	NZ	LYS40	3.61	Pontes salinas
OE2	GLU175	NZ	LYS10	2.68	Pontes salinas
NE	ARG251	OD2	ASP189	2.92	Pontes salinas
OE1	GLU363	NZ	LYS25	3.06	Pontes salinas
NZ	LYS382	OD1	ASP136	2.77	Pontes salinas
OE1	GLU434	NH1	ARG137	2.73	Pontes salinas
OD2	ASP437	NH1	ARG22	2.82	Pontes salinas
OE1	GLU460	NH1	ARG22	2.59	Pontes salinas

### Apêndice H

**Tabela A8:** Docagem molecular entre o receptor *toll-like* 3 e o modelo de proteína imunogênica 4.

<b>TLR-3 (Átomo)</b>	<b>TLR-3 (Resíduo)</b>	<b>Modelo (Átomo)</b>	<b>Modelo (Resíduo)</b>	<b>Distância</b>	<b>Tipo de interação</b>
NZ	LYS 12	O	LYS353	2.33	Ligações de hidrogênio
OE1	GLU146	NZ	LYS36	2.59	Ligações de hidrogênio
NZ	LYS301	OE1	GLU69	2.27	Ligações de hidrogênio
NH1	ARG460	OG	SER273	3.08	Ligações de hidrogênio
O	LYS502	OH	TYR222	2.11	Ligações de hidrogênio
O	LEU586	OH	TYR207	2.72	Ligações de hidrogênio
ND1	HIS 10	CA	GLY354	2.44	Interações hidrofóbicas
CE	LYS 12	O	LYS353	2.44	Interações hidrofóbicas
NZ	LYS 12	O	LYS353	2.33	Interações hidrofóbicas
OE1	GLU146	NZ	LYS36	2.59	Interações hidrofóbicas
OE2	GLU146	NZ	LYS36	2.69	Interações hidrofóbicas
CE	LYS172	NZ	LYS36	2.56	Interações hidrofóbicas
NZ	LYS172	CE	LYS36	2.26	Interações hidrofóbicas
NZ	LYS172	NZ	LYS36	1.53	Interações hidrofóbicas
OE1	GLU146	NZ	LYS36	2.59	Pontes salinas
NZ	LYS301	OE1	GLU69	2.27	Pontes salinas
OD2	ASP563	NH2	ARG263	3.02	Pontes salinas

## Apêndice I

**Tabela A9:** Docagem molecular entre epítomos reconhecidos por HTL e moléculas de classe II do complexo principal de histocompatibilidade.

Epítopo	Receptor/Supertipo	Protein Data Bank ID	Docagem Score
MMASILLTLFRRTKK	HLA-DRB5*01:01	1FV1	-230,32
	HLA-DPA1*02:01/DPB1*05:01	3wex (Homology Modelling)	-258,76
	HLA-DPA1*03:01/DPB1*04:02	4p5k (Homology Modelling)	-292,07
	HLA-DPA1*01:03/DPB1*02:01	4P4K	-293,4
	HLA-DPA1*01/DPB1*04:01	4P5K (Homology Modelling)	-301,17
NSEWISIVPNFILVR	HLA-DRB1*15:01	5V4M	-245,55
	HLA-DRB1*13:02	6ATF	-277,15
	HLA-DRB1*04:05	6PX6	-265,24
	HLA-DQA1*05:01/DQB1*02:01	4d8p (Homology Modelling)	-258,23
	HLA-DRB1*04:01	5JLZ	-251,12
	HLA-DRB1*08:02	6CPL	-239,11
	HLA-DRB1*12:01	6ATF	-260,09
IILYVLSIASLCIGL	HLA-DRB1*07:01	3PDO	-301,35
	HLA-DRB1*04:01	5JLZ	-298,12
	HLA-DRB1*15:01	5V4M	-288,41
	HLA-DRB1*11:01	6CPL	-287,61
	HLA-DRB5*01:01	1FV1	-290,72
	HLA-DRB3*02:02	3C5J	-281,34
	HLA-DRB4*01:01	1AQD	-289,15
FISFIIIVEKKRNTYS	HLA-DRB1*11:01	6CPL	-266,07
	HLA-DRB1*12:01	6ATF	-275,47
NIEIGFCLITKRSVI	HLA-DRB1*09:01	3PDO	-259,23
	HLA-DRB5*01:01	1FV1	-251,32
VPNFILVRNTLISNI	HLA-DRB1*04:01	5JLZ	-273,63
	HLA-DRB1*08:02	6CPL	-262,78
	HLA-DRB1*13:02	6ATF	-244,51
	HLA-DRB1*15:01	5V4M	-232,63
TAVLGNVIISLGKYL	HLA-DRB1*13:02	6ATF	-213,06
	HLA-DPA1*03:01/DPB1*04:02	4p5k (Homology Modelling)	-229,01
	HLA-DRB1*07:01	3PDO	-231,75
NSEWISIVPNFILVR	HLA-DRB1*15:01	5V4M	-254,18
	HLA-DRB1*13:02	6ATF	-238,19
	HLA-DRB1*04:05	6PX6	-212,04
	HLA-DQA1*05:01/DQB1*02:01	5KSU	-259,59
	HLA-DRB1*04:01	5JLZ	-244,22
	HLA-DRB1*08:02	6CPL	-262,11
	HLA-DRB1*12:01	6ATF	-270,89
DSKILSAFNTVIALL	HLA-DRB1*01:01	4GBX	-278,25
	HLA-DPA1*02:01/DPB1*01:01	3wex	-280,09
	HLA-DRB1*13:02	6ATF	-255,34
	HLA-DPA1*03:01/DPB1*04:02	4p5k (Homology Modelling)	-302,02

	HLA-DQA1*04:01/DQB1*04:02	4d8p (Homology Modelling)	-294,93
	HLA-DRB3*02:02	3C5J	-257,77
	HLA-DPA1*01/DPB1*04:01	4p5k (Homology Modelling)	-252,36
	HLA-DPA1*01:03/DPB1*02:01	4P4K	-271,43
	HLA-DPA1*02:01/DPB1*05:01	3wex (Homology Modelling)	-245,38
SAVYNNEFYVLCVAV	HLA-DPA1*01/DPB1*04:01	4p5k (Homology Modelling)	-286,1
	HLA-DRB3*01:01	1A6A	-222,95
	HLA-DPA1*01:03/DPB1*02:01	4P4K	-247,85
	HLA-DQA1*01:01/DQB1*05:01	4GRL	-239,9
GVYNDAFLIDRINWI	HLA-DRB1*13:02	6ATF	-253,28
	HLA-DPA1*03:01	4P4K_A	-291,23
	HLA-DPA1*03:01/DPB1*04:02	4p5k (Homology Modelling)	-300,67
	HLA-DRB1*04:05	6PX6	-307,14
SKILSAFNTVIALLG	HLA-DRB1*07:01	3PDO	-281,5
	HLA-DRB1*08:02	6CPL	-288,22
	HLA-DRB1*04:01	5JLZ	-293,46
	HLA-DQA1*01:02/DQB1*06:02	1UVQ	-295,55
	HLA-DRB1*04:05	6PX6	-298,34
SKILSAFNTVIALLG	HLA-DRB1*07:01	3PDO	-290,18
	HLA-DRB1*08:02	6CPL	-264,79
	HLA-DRB1*04:01	5JLZ	-273,48
	HLA-DQA1*01:02/DQB1*06:02	6blq (Homology Modelling)	299,78
	HLA-DRB1*08:02	6CPL	-265,55
ADKVLEYAPIMKAHA	HLA-DRB1*03:01	1A6A (Homology Modelling)	-281,45
	HLA-DRB1*07:01	3PDO	-275,63
	HLA-DRB1*04:05	6PX6	-251,91
	HLA-DPA1*01:03/DPB1*02:01	4P4K	-231,27
	HLA-DRB5*01:01	1FV1	-259,11
	HLA-DRB1*15:01	1FV1	-273,45
	HLA-DRB1*11:01	5V4M	-233,48
	HLA-DPA1*02:01/DPB1*05:01	3wex (Homology Modelling)	-292,14
	HLA-DPA1*01/DPB1*04:01	4p5k (Homology Modelling)	-285,22
	HLA-DRB3*01:01	1A6A	-304,05
HLA-DRB1*07:01	3PDO	-307,32	
KRKKIRTIAAYPLGV	HLA-DRB3*02:02	3C5J	-299,45
	HLA-DRB1*13:02	6ATF	-285,33
	HLA-DRB5*01:01	1FV1	-279,65
	HLA-DRB1*09:01	3PDO	-301,02
	HLA-DRB1*13:02	6ATF	-306,12
AVEIISDIGNYVEET	HLA-DRB1*03:01	1A6A (Homology Modelling)	-304,54
	HLA-DQA1*05:01/DQB1*02:01	4d8p (Homology Modelling)	-307,88
	HLA-DRB3*02:02	3C5J	-306,54
	HLA-DPA1*01:03/DPB1*02:01	4P4K	-307,81
	HLA-DQA1*03:01/DQB1*03:02	5KS9	-303,9
	HLA-DQA1*05:01/DQB1*02:01	4d8p (Homology Modelling)	-287,17
	HLA-DPA1*03:01/DPB1*04:02	4p5k (Homology Modelling)	-283,12

	HLA-DPA1*02:01/DPB1*01:01	3wex (Homology Modelling)	-286,31
	HLA-DQA1*04:01/DQB1*04:02	4d8p (Homology Modelling)	-291,36
EFEDEFAGSSSEVIV	HLA-DRB1*07:01	3PDO	-303,51
	HLA-DQA1*04:01/DQB1*04:02	4d8p (Homology Modelling)	-275,43

## Apêndice J

**Esquema A2** - Sequências de cDNAs otimizadas dos modelos selecionados 1 (A) e 4 (B) com códon de iniciação inserido e sítios de restrição BamHI/HindIII.

### A) Model 1

ATGGGATCCGGTATCGGTGACCCGGTTACCTGCCTGAAATCTGGTGCTATCTGCCACCCGGTTTTCTG  
 CCCGCGTCGTTACAAACAGATCGGTACCTGCGGTCTGCCGGGTACCAAATGCTGCAAAAAACCGGAA  
 GCTGCTGCTAAAGACTCTAAAATCCTGTCTGCTTTCAACACCGTTATCGCTCTGCTGGGTCCGGGTCC  
 GGGTAACTCTGAATGGATCTCTATCGTTCCGAACTTCATCCTGGTTCGTGGTCCGGGTCCGGGTATCA  
 TCCGTGACGTTCTGTCTATCGCTTCTCTGTGCATCGGTCTGGGTCCGGGTCCGGGTCTGCTGTTTACA  
 ACAACGAATTCTACTACGTTCTGTGCGCTGTTGGTCCGGGTCCGGGTGGTGTTTACAACGACGCTTTC  
 CTGATCGACCGTATCAACTGGATCGGTCCGGGTCCGGGTTCATCTCTTTCATCATCGTTGAAAAAAA  
 ACGTAAACCTACTCTGGTCCGGGTCCGGGTCTAAAATCCTGTCTGCTTTCAACACCGTTATCGCTC  
 TGCTGGGTGGTCCGGGTCCGGGTGCTATGGACGAAGGTTACTTCGCTTACGCTGCTTACGTTCTGTTT  
 GCTAACTGCATCTCTGTTGCTGCTTACCTGATGATGACCCGCTCTGGCTGTTAAAGCTGCTTACGAATG  
 GATCTCTATCGTTCGAACTTCGCTGCTTACGAAACCCTGCTGCGTACCCTGGGTACGCTGCTTACA  
 TCGTTCGTGTTTACTTCCCAGATCCTGGCTGCTTACGCTCAGATCACCGCTGGTGTGGTCTGGCTGCTT  
 ACACCGTTAACCCGCTGGTGTGTTAACTGGGCTGCTTACAAAATCGGTCTGGTTAAAGGTAAAAAAC  
 CGCTGGTGTGCTCTGTACAAAAAATCCAGGAACTGCTGCCGGTTTCTAAAAAATCTCTATCGTTC  
 CGAACTTCATCAAAAAAGTTTACCACTGCTCTGCTGTTAAAAAATACTTCCCAGCTGTTGGTTTCTG  
 AAAAAACCGTCTAAAGTTATCAAATCTGAAGCTGCTGCTAAAGGTATCATCAACACCCTGCAGAAAT  
 ACTACTGCCGTGTTTCGTGGTGGTTCGTTGCGCTGTTCTGTCTTGCCTGCCGAAAGAAGAACAGATCGGT  
 AAATGCTCTACCCGTGGTTCGTAATGCTGCCGTCGTAAAAAAAGCTT

### B) Model 4

ATGGGATCCGGTATCGGTGACCCGGTTACCTGCCTGAAATCTGGTGCTATCTGCCACCCGGTTTTCTG  
 CCCGCGTCGTTACAAACAGATCGGTACCTGCGGTCTGCCGGGTACCAAATGCTGCAAAAAACCGGAA  
 GCTGCTGCTAAAGCTGACAAAGTTCTGGAATACGCTCCGATCATGAAAGCTCACGCTGGTCCGGGTCC  
 CGGTGCTGTTGAAATCATCTCTGACATCGGTAACCTGTTGAAGAAACCGGTCCGGGTCCGGGTAT  
 GATGGCTTCTATCCTGCTGACCCTGTTCCGTCGTACCAAAAAAGGTCCGGGTCCGGGTAAACGTAAA  
 AAAATCCGTACCATCGCTGCTTACCCGCTGGGTGTTGGTCCGGGTCCGGGTGAATTCGAAGACGAAT  
 TCGCTGGTTCTTCTTCTGAAGTTATCGTTGGTCCGGGTCCGGGTAACTCTGAATGGATCTCTATCGTTC  
 CGAACTTCATCCTGGTTCGTGGTCCGGGTCCGGGTCTAAAATCCTGTCTGCTTTCAACACCCTGATC  
 GCTCTGCTGGGTGGTCCGGGTCCGGGTCTCACATGTGGCGTGAAGTGGCTCTGGCTGCTTACTGGAC  
 CCAGTGGCTGCAGACCCTGTACGCTGCTTACGCTCAGATCACCGCTGGTGTGGTCTGGCTGCTTACG  
 CTATGGACGAAGGTTACTTCGCTTACGCTGCTTACAAATCTCGTGGTATCCCAGTCAAAAAAGCTGCT  
 TACTACATGTACCTGATCTGCTACGGTTTCGCTGCTTACTTCCCTGATCGACCGTATCAACTGGATCGC  
 TGCTTACCGTGCTTCTGCTGCTACCGTACCCTGGCTGCTTACATCCAGGAACTGCTGCCGGTTTCTA  
 AAAAAATACCACGACCACGGTGGTGAATAAAAAAAGTGTACATCCTGCGTCAGTCTAAACAGGGTGACA  
 AAAAAAGAACTGGTTAACGACGGTCTGAAAAAAGTTCTGCTGGAAGAATCTATCCAGACCAAAAAAG  
 AAGGTGTTTACAACAAAAAACCCTGGTGTGGTCTGTACGAAGCTGCTGCTAAAGGTATCATCAA  
 CACCCTGCAGAAATACTACTGCCGTGTTTCGTGGTGGTTCGTTGCGCTGTTCTGTCTTGCCTGCCGAAAG  
 AAGAACAGATCGGTAATGCTCTACCCGTGGTTCGTAATGCTGCCGTCGTAAAAAAAGCTT