



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR
CURSO DE GRADUAÇÃO EM BIOTECNOLOGIA

DÉBORA ALBUQUERQUE DE LIMA PACHECO

**DESIMUNIZAÇÃO DE L-ASPARAGINASE DE *ESCHERICHIA COLI* ATRAVÉS DE
ALGORITMO GENÉTICO**

FORTALEZA

2019

DÉBORA ALBUQUERQUE DE LIMA PACHECO

DESIMUNIZAÇÃO DE L-ASPARAGINASE DE *ESCHERICHIA COLI* ATRAVÉS DE
ALGORITMO GENÉTICO

Trabalho de Conclusão de Curso apresentado
ao Curso de Graduação em Biotecnologia do
Centro de Ciências da Universidade Federal do
Ceará, como requisito parcial à obtenção do
grau de bacharel em Biotecnologia.

Orientador: Dr. Raphael Trevizani

Coorientadora: Dra. Márjory Lima Holanda
Araújo

FORTALEZA

2019

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

A299d Albuquerque de Lima Pacheco, Débora.
Desimunização de L-asparaginase de Escherichia coli através de Algoritmo Genético / Débora
Albuquerque de Lima Pacheco. – 2019.
86 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Ciências,
Curso de Biotecnologia, Fortaleza, 2019.

Orientação: Prof. Dr. Raphael Trevizani.

Coorientação: Profa. Dra. Márjory Lima Holanda Araújo.

1. Imunogenicidade. 2. Engenharia de proteínas. 3. Leucemia linfoblástica aguda. I. Título.

CDD 661

DÉBORA ALBUQUERQUE DE LIMA PACHECO

DESIMUNIZAÇÃO DE L-ASPARAGINASE DE *ESCHERICHIA COLI* ATRAVÉS DE
ALGORITMO GENÉTICO

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Biotecnologia do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Biotecnologia.

Aprovada em:

BANCA EXAMINADORA

Dr. Raphael Trevizani (Orientador)
Fundação Oswaldo Cruz (Fiocruz)

Dra. Márjory Lima Holanda Araújo (Coorientadora)
Universidade Federal do Ceará (UFC)

Dr. Nicholas Costa Barroso Lima
Universidade Federal do Ceará (UFC)

Aos pacientes, familiares e profissionais de
saúde que lutam contra a leucemia todos os dias.

AGRADECIMENTOS

A **Deus**. Não por senso comum, não pelo senso poético, ou por minhas experiências pessoais. Mas por seguir o argumento até onde ele levar. Os fatos me levam a Ele, e Ele me trouxe até aqui. A gratidão é consequência de todo o caminho percorrido.

Ao meu pai, **José Antonio Pacheco**, por me inspirar em todos os âmbitos, investir de todas as formas e tentar ajudar das maneiras mais peculiares. Nunca vou esquecer seu cuidado e atenção para que esse trabalho fosse feito da melhor forma possível.

À minha mãe, **Rosilene Albuquerque de Lima Pacheco**, minha melhor amiga e a mulher que mais me inspira. Obrigada por me ajudar, me acolher e me incentivar todos os dias. Você é um dos meus maiores combustíveis.

Ao meu irmão, **Rafael Pacheco**, e à cunhada, **Marília Brilhante**. Obrigada pelas buscas de artigos, revisões e pelo apoio constante. Vocês são peças fundamentais na minha vida, e foram muito importantes durante o desenvolvimento desse trabalho.

Aos meus avôs **Sebastião Cota Pacheco** (*in memoriam*) e **Raimundo Albuquerque de Lima**; às avós **Terezinha de Jesus Pacheco** (*in memoriam*) e **Laís Helena de Lima**; às tias **Maria Abadia**, **Rosângela**, **Rosemary** e **Robersi**; aos primos **Marina**, **Gabriel**, **Sarah** e **Isabela**; aos tios agregados **Guido** e **Emerson**. Obrigada pelo amor, cuidado, carinho e incentivo constantes.

Ao Orientador Técnico-Científico, Dr. **Raphael Trevizani**, por seu comprometimento e competência, por me desafiar e me preparar para o futuro. Agradeço as instruções, correções e todo o conhecimento transmitido.

À Orientadora Pedagógica, Profa. Dra. **Márjory Holanda**, por sua gentileza, preocupação e acolhimento. Muito obrigada por sua orientação, disponibilidade e pelas extensas revisões.

À **Coordenação e Secretaria do Curso de Biotecnologia**, pelo senso de responsabilidade e compromisso.

Aos colegas de estágio: **Tayná Fiúza**, pela cumplicidade e por ter aberto o caminho para que esse trabalho fosse executado; e **Igor Duarte** pela parceria e ajuda no meu início em bioinformática.

Aos amigos **Amanda**, **Eduardo**, **Guilherme**, **Janise**, **Juliany**, **Luca**, **Ludmilla** e **Priscila**. Enfrentamos juntos a maior parte desses longos anos de graduação. A amizade, companheirismo, parceria e o apoio de vocês tornavam as maiores dificuldades mais fáceis de se

enfrentar. Obrigada pelos dias e noites de estudo em grupo, por terem sempre uma palavra de ânimo na ponta da língua e por não medirem esforços para ajudar uns aos outros. Com vocês, aprendi algo muito mais importante do que o que qualquer Universidade pode ensinar: o valor de uma amizade sincera.

Ao **Boris** e à **Lucy**, obrigada por me fazerem tão bem. Basta olhar para vocês e meu coração sorri.

À **Universidade Federal do Ceará** e à **University of Technology: Sydney** pelo aprendizado, pelos professores brilhantes e pelos recursos disponibilizados; e à **Fiocruz**, por me proporcionar avançar no aprendizado e aplicação dos conhecimentos através do Programa de Estágio.

"Success is not final, failure is not fatal: it is the courage to continue that counts."

(Winston Churchill)

RESUMO

A L-asparaginase é uma proteína terapêutica usada no tratamento de Leucemia Linfoblástica Aguda cuja utilização pode causar efeitos colaterais severos ou resultar na neutralização da proteína devido à ativação da resposta imune. Componentes da imunidade adaptativa, como o MHC II (Complexo Principal de Histocompatibilidade II), são responsáveis por apresentar peptídeos exógenos processados por células apresentadoras de antígenos (APCs). Ferramentas de imunoinformática podem ser usadas para prever a afinidade de ligação de um peptídeo com a fenda do MHC II, que resultam na imunogenicidade do peptídeo. Além disso, métodos de desimunização por mutação de epítomos podem ser empregados na tentativa de melhorar a eficácia dessas proteínas e reduzir sua imunogenicidade. Nesse trabalho o método de predição de afinidade peptídeo-MHC II, TEPITOPE, foi aplicado para calcular os escores de imunogenicidade de epítomos da L-asparaginase de *Escherichia coli*, e essas informações foram usadas em um Algoritmo Genético (AG) de desimunização de proteínas. O Algoritmo utilizou um conjunto de sequências da proteína e realizou operações de recombinação e mutação a cada ciclo nas regiões imunogênicas da proteína, considerando apenas mutações neutras (cujo $|\Delta\Delta G|$ entre proteína selvagem e mutante era menor que 0,5 kcal/mol) e substituindo sequências de maior imunogenicidade por aquelas de menor. Ao final de 30 aplicações, o algoritmo foi capaz de gerar 953 variantes estáveis com imunogenicidade menor que a sequência original. Destas, foram analisadas as sequências de três, quatro e cinco mutações com a menor imunogenicidade encontrada. A partir dessa análise foi possível verificar as mutações que contribuíram para a redução da imunogenicidade e as que foram necessárias para que a proteína se mantivesse estável. O AG pode ser aplicado para outras sequências proteicas, com potencial de parametrização para variar a dinâmica do algoritmo ou para implantar novas funções que aumentem sua eficiência.

Palavras-chave: Imunogenicidade. Engenharia de proteínas. Leucemia linfoblástica aguda.

ABSTRACT

L-asparaginase is a therapeutic protein used in Acute Lymphoblastic Leukemia treatment, whose administration may cause severe side effects or result in protein neutralization due to the activation of the immune response. Adaptive immunity components, such as MHC II (Major Histocompatibility Complex class II), are responsible for presenting exogenous peptides processed by antigen presenting cells (APCs). Immunoinformatics tools can be used to predict the binding affinity between a peptide and MHC-II pocket, which result in peptide immunogenicity. In addition, deimmunization methods by epitope mutation may be employed in an attempt to improve therapeutic proteins efficacy and reduce its immunogenicity. In this work, a peptide-MHC II affinity prediction method, TEPITOPE, was applied to calculate the immunogenicity scores of *Escherichia coli* L-asparaginase epitopes, and this information was used in a Genetic Algorithm (GA) of protein deimmunization. The Algorithm used a set of protein sequences and performed crossover and mutations at each cycle in the immunogenic regions of the protein, considering only neutral mutations (when $\Delta\Delta G$ between wild and mutant protein was less than 0,5 kcal/mol), substituting sequences of higher immunogenicity by those of lower. At the end of 30 applications, the algorithm was able to generate 953 stable variants with less immunogenicity than the original sequence. Of these, the sequences of three, four and five mutations with the lowest immunogenicity were analyzed. From this analysis it was possible to verify the mutations that contributed to the reduction of the immunogenicity and those that were necessary for the protein to remain stable. The GA can be applied to other protein sequences, with parameterization potential to vary the dynamics of the algorithm or to implement new functions that increase its efficiency.

Keywords: Immunogenicity. Protein engineering. Acute lymphoblastic leukemia.

LISTA DE FIGURAS

Figura 1 – Diferenciação de células sanguíneas a partir de uma célula-tronco hematopoiética	18
Figura 2 – Representação da hidrólise de L-asparagina por meio da ação da enzima L-asparaginase	20
Figura 3 – Estrutura cristalográfica de L-asparaginase de <i>Escherichia coli</i> no estado homotetramérico	22
Figura 4 – Classificação de epítomos baseada em sua estrutura tridimensional	24
Figura 5 – Representação da estrutura de MHC de classes I e II em um complexo de ligação epítomo-MHC-linfócito T	26
Figura 6 – Representação da estrutura da fenda de ligação do Complexo Principal de Histocompatibilidade (MHC) II	28
Figura 7 – Representação do cálculo dos escores de imunogenicidade de epítomos de L-asparaginase de <i>Escherichia coli</i> através do TEPITOPE	43
Figura 8 – Representação dos indivíduos e da população inicial no Algoritmo Genético	45
Figura 9 – Representação do operador de mutação utilizado pelo Algoritmo Genético para geração de novas sequências desimunizadas de L-asparaginase de <i>Escherichia coli</i>	46
Figura 10 – Representação do operador de recombinação utilizado pelo Algoritmo Genético para geração de novas sequências desimunizadas de L-asparaginase de <i>Escherichia coli</i>	47
Figura 11 – Representação do método de seleção de indivíduos parentais para recombinação utilizado pelo Algoritmo Genético para geração de novas sequências desimunizadas de L-asparaginase de <i>Escherichia coli</i>	48
Figura 12 – Função de restrição que verifica a criação de novos epítomos gerados pela alteração em uma região dos epítomos de L-asparaginase de <i>Escherichia coli</i>	49
Figura 13 – Fluxograma de Operação do Algoritmo Genético para obtenção de sequências desimunizadas de L-asparaginase de <i>Escherichia coli</i>	51
Figura 14 – Contribuições para a afinidade de ligação de epítomos de L-asparaginase de <i>E. coli</i> associadas à cada posição da fenda de MHC II	56
Figura 15 – Indivíduos gerados pelo Algoritmo Genético a partir de mutações nos epítomos M^{111} , V^{112} e V^{217} de L-asparaginase de <i>E. coli</i>	65

Figura 16 – Fronteira de Pareto dispondo todos os indivíduos produzidos pelo Algoritmo Genético	67
Figura 17 – Decaimento da similaridade de sequência em comparação à redução da imunogenicidade de mutantes obtidos pelo AG a partir de mutações nos epítomos de L-asparinase de <i>E. coli</i>	68

LISTA DE TABELAS

Tabela 1 – Formulações de L-asparaginase tipo II utilizadas no tratamento de LLA . . .	21
Tabela 2 – Fatores que contribuem para a imunogenicidade de proteínas terapêuticas . . .	23
Tabela 3 – Preditores de epítomos de linfócitos T CD4+ disponíveis online para uso público	32
Tabela 4 – Métodos de predição de epítomos de MHC classe II em larga escala	34
Tabela 5 – Epítomos de L-asparaginase de <i>Escherichia coli</i> determinados experimentalmente que foram alvos de mutagênese	42
Tabela 6 – Alelos de HLA utilizados na predição da imunogenicidade de epítomos de L-asparaginase	42
Tabela 7 – Resumo dos parâmetros utilizados no Algoritmo Genético para obtenção de sequências desimunizadas de L-asparaginase de <i>Escherichia coli</i>	51
Tabela 8 – Matriz de Pontuação BLOSUM62 utilizada para avaliar a diferença entre as sequências de L-asparaginase de <i>Escherichia coli</i> selvagem e as sequências mutantes propostas pelo Algoritmo Genético	53
Tabela 9 – Escores de imunogenicidade calculados para os epítomos de L-asparaginase de <i>Escherichia coli</i>	55
Tabela 10 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*1501	59
Tabela 11 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB5*0101	60
Tabela 12 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*1302	61
Tabela 13 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*0401	62
Tabela 14 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*0701	63
Tabela 15 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*0101	64
Tabela 16 – $\Delta\Delta G$ das mutações propostas nos resíduos correspondentes às posições 111, 112 e 113 da sequência de L-asparaginase de <i>E. coli</i>	71

LISTA DE QUADROS

Quadro 1 – Alelos de MHC para os quais o TEPITOPE disponibiliza matrizes de afinidade de ligação	35
--	----

LISTA DE ABREVIATURAS E SIGLAS

AG	Algoritmo Genético
ANN	Rede Neural Artificial
APC	Células Apresentadoras de Antígenos
AUC	Área sob a Curva ROC
BCR	Receptores de Linfócitos B
FDA	<i>Food and Drug Administration</i>
HLA	<i>Human Leukocyte Antigen</i>
LLA	Leucemia Linfoblástica Aguda
MHC	Complexo Principal de Histocompatibilidade
MM	Matrizes de Motivos
PDB	<i>Protein Data Bank</i>
PEG	Polietilenoglicol
QAM	Matrizes Quantitativas de Afinidade
QSAR	Modelo Quantitativo de Relação Estrutura-Atividade
RMSD	<i>Root-Mean-Square Deviation</i>
SM	Motivos de Sequência
SVM	Máquinas de Vetores de Suporte
TCR	Receptores de Linfócitos T

LISTA DE SÍMBOLOS

Å	Angstrom
G	Energia Livre de <i>Gibbs</i>
Δ	Varição

SUMÁRIO

1	INTRODUÇÃO	18
1.1	Leucemia Linfoblástica Aguda	18
1.2	L-asparaginase no tratamento de Leucemia Linfoblástica Aguda	19
<i>1.2.1</i>	<i>Histórico</i>	19
<i>1.2.2</i>	<i>Aplicação Terapêutica</i>	20
1.3	Imunogenicidade	22
<i>1.3.1</i>	<i>Desenvolvimento da resposta imunológica contra proteínas terapêuticas</i>	23
<i>1.3.2</i>	<i>Efeitos adversos da administração de L-asparaginase</i>	29
1.4	Mapeamento de Epítomos Lineares	30
<i>1.4.1</i>	<i>TEPITOPE</i>	35
1.5	Desimunização	36
<i>1.5.1</i>	<i>Algoritmo Genético</i>	38
2	OBJETIVOS	40
2.1	Objetivo Geral	40
2.2	Objetivos Específicos	40
3	METODOLOGIA	41
3.1	Determinação da Imunogenicidade de Epítomos Lineares	41
<i>3.1.1</i>	<i>Identificação de epítomos de L-asparaginase de Escherichia coli e seleção de alelos de HLA-II</i>	41
<i>3.1.2</i>	<i>Cálculo da imunogenicidade dos epítomos</i>	42
3.2	Desimunização da L-asparaginase de Escherichia coli	44
<i>3.2.1</i>	<i>Indivíduo e População</i>	44
<i>3.2.2</i>	<i>Operador de Mutação</i>	45
<i>3.2.3</i>	<i>Operador de Recombinação</i>	46
<i>3.2.3.1</i>	<i>Seleção Parental</i>	47
<i>3.2.4</i>	<i>Funções de Restrição</i>	47
<i>3.2.4.1</i>	<i>Estabilidade Estrutural</i>	47
<i>3.2.4.2</i>	<i>Criação de Novos Epítomos</i>	48
<i>3.2.5</i>	<i>Função de Aptidão</i>	49
<i>3.2.6</i>	<i>Substituição Parental</i>	50

3.2.7	<i>Fluxograma do Algoritmo Genético</i>	50
3.3	Análise e Seleção dos Mutantes Obtidos	52
3.3.1	<i>Fronteira de Pareto</i>	52
3.3.2	<i>Avaliação da Similaridade de Sequência</i>	52
3.3.3	<i>Avaliação da Variação Estrutural</i>	53
4	RESULTADOS E DISCUSSÃO	55
4.1	Determinação da Imunogenicidade de Epítomos Lineares	55
4.2	Desimunização da L-asparaginase de <i>Escherichia coli</i>	63
4.3	Análise e Seleção dos Mutantes Obtidos	66
4.3.1	<i>Triplo-Mutante</i>	69
4.3.2	<i>Quádruplo-Mutantes</i>	71
4.3.3	<i>Quíntuplo-Mutante</i>	72
5	CONCLUSÃO	74
	REFERÊNCIAS	75
	APÊNDICES	86
	APÊNDICE A – Termos de Energia Considerados pelo Foldx	86
	APÊNDICE B – Matriz de Pontuação BLOSUM62	87

1 INTRODUÇÃO

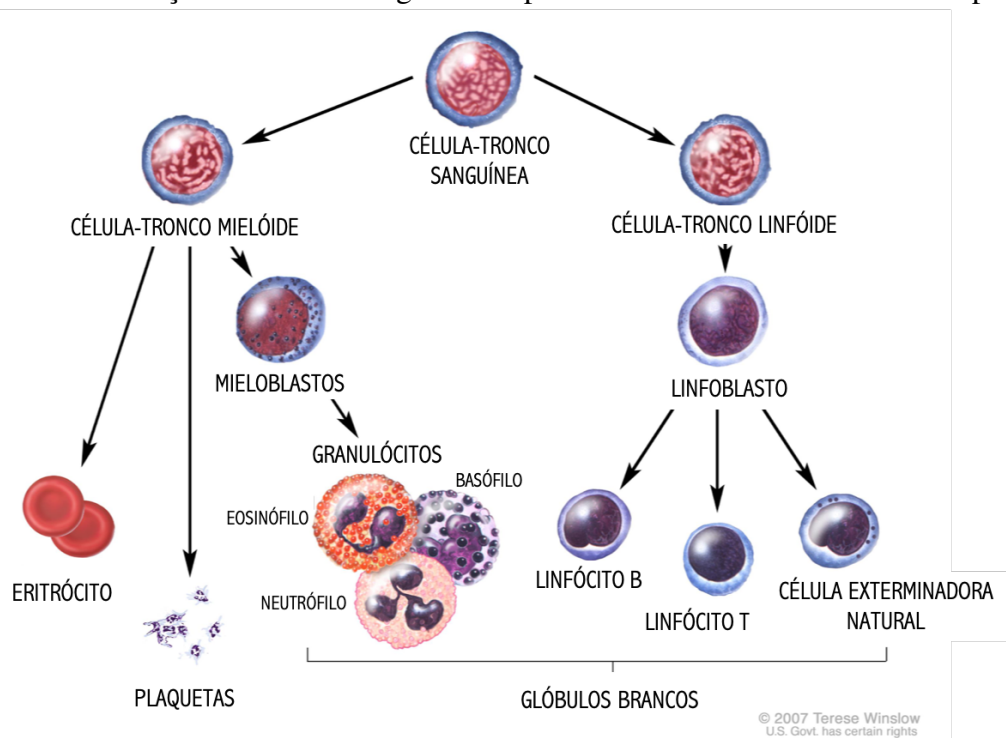
1.1 Leucemia Linfoblástica Aguda

A Leucemia Linfoblástica Aguda (LLA) é um câncer sanguíneo no qual há uma produção desordenada de células brancas imaturas ou linfócitos na medula óssea. É o tipo de câncer mais comum em crianças, representando 29% dos casos diagnosticados em indivíduos de até 15 anos (NATIONAL CANCER INSTITUTE, 2018). Aproximadamente 3 em cada 100.000 indivíduos manifestam a doença, comumente antes dos 6 anos de idade, com maior incidência entre 2 e 4 anos de idade (FERNANDES *et al.*, 2016).

Em um indivíduo saudável, a medula óssea produz células-tronco hematopoiéticas que se tornam células sanguíneas maduras ao longo do tempo, podendo se diferenciar primariamente em duas linhagens principais: mielóides ou linfóides.

As células-tronco **mielóides** se desenvolvem em três tipos de células sanguíneas maduras: hemácias, que transportam oxigênio e outras substâncias para os tecidos do corpo; plaquetas, responsáveis pela formação de coágulos; e os mieloblastos, que posteriormente se diferenciam em neutrófilos, eosinófilos e basófilos (SILVERTHORN, 2017), como ilustra a Figura 1.

Figura 1 – Diferenciação de células sanguíneas a partir de uma célula-tronco hematopoiética



Fonte: Adaptado de NATIONAL CANCER INSTITUTE (2018).

Células-tronco **linfóides**, por sua vez, também se diferenciam em três tipos celulares: linfócitos B, que combatem infecções; linfócitos T, que criam anticorpos para o combate de doenças; e células chamadas de exterminadoras naturais (NK) (Figura 1), que são importantes no combate a infecções virais devido ao seu rápido mecanismo de resposta em relação às células B e T (SILVERTHORN, 2017).

Em um indivíduo portador de LLA, grandes quantidades de células-tronco sanguíneas se diferenciam em linfoblastos, linfócitos B ou T, mas que não são totalmente desenvolvidos e, portanto, não são aptos ao combate de infecções (EGLER *et al.*, 2016). Além disso, devido ao aumento no número de células leucêmicas, as concentrações de glóbulos brancos saudáveis, hemácias e plaquetas diminuem, levando a sintomas como anemia, diátese hemorrágica e infecções (NATIONAL CANCER INSTITUTE, 2018).

1.2 L-asparaginase no tratamento de Leucemia Linfoblástica Aguda

1.2.1 Histórico

Em um trabalho pioneiro, Clementi (1922) detalhou as propriedades enzimáticas de uma solução hidrofílica encontrada no tecido de diversos animais de várias classes de vertebrados capaz de hidrolisar asparagina e a denominou L-asparaginase (número EC 3.5.1.1). Ao enumerar os animais cujos tecidos continham L-asparaginase, o autor observou que apenas cobaias apresentavam L-asparaginase também no plasma sanguíneo. Anos depois, foi reportado que linfomas de ratos e camundongos regrediam quando tratados com soro de cobaias, enquanto linfomas tratados com soro de cavalo, porco e coelho cresciam e matavam os hospedeiros (KIDD, 1953). Quando Neuman e McCoy demonstraram que asparagina exógena era necessária para o crescimento tumoral (NEUMAN; MCCOY, 1956), o soro de cobaias tornou-se um agente potencial terapêutico contra linfomas transplantados graças à presença de L-asparaginase (BROOME, 1961; BROOME, 1963).

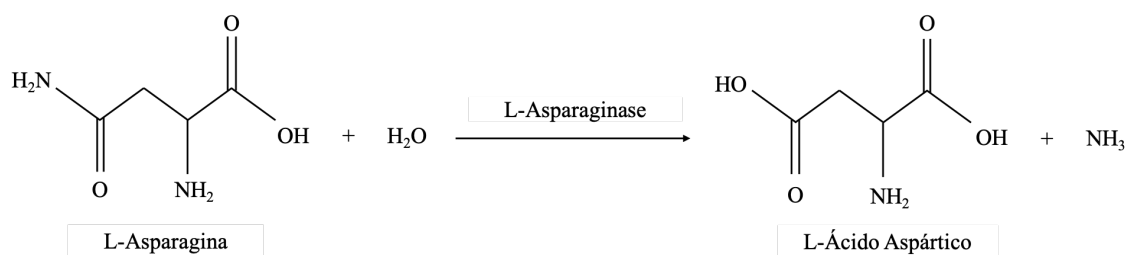
Apesar desse potencial terapêutico, a L-asparaginase de cobaias foi desconsiderada para uso clínico porque sua produção não é adequadamente abundante, levando pesquisadores a se voltarem para fontes bacterianas da enzima (OFFMAN *et al.*, 2011). Finalmente, uma L-asparaginase purificada de *Escherichia coli* (MASHBURN; WRISTON, 1964; CAMPBELL; MASHBURN, 1969) demonstrou potencial antineoplásico comparável ao de cobaias (PRA-GER; BACHYNSKY, 1968) com a vantagem de ser viável para a produção em larga escala

(CAMPBELL *et al.*, 1967; HO *et al.*, 1970). Em 1967, Ohnuma identificou duas isozimas de L-asparaginase em *E. coli*, tipos I e II (OHNUMA *et al.*, 1967). A L-asparaginase tipo I inclui enzimas citosólicas expressadas constitutivamente e com baixa afinidade pela L-asparagina, resultando em aplicações não-terapêuticas. A L-asparaginase tipo II localiza-se no espaço periplasmático da bactéria e tem alta-afinidade por asparagina. Conjuntamente, esses esforços culminaram na aprovação da primeira L-asparaginase tipo II de *E. coli* pelo *Food and Drug Administration* (FDA) em 1978 para uso terapêutico nos Estados Unidos.

1.2.2 Aplicação Terapêutica

A asparagina é um aminoácido não-essencial para humanos, sintetizada por meio da asparagino-sintetase, uma enzima que catalisa a conversão de aspartato a asparagina, através de glutamina e nitrogênio (RICHARDS; SCHUSTER, 1998). Entretanto, células tumorais não são capazes de sintetizar asparagina endogenamente em quantidade suficiente para sobrevivência e multiplicação, o que as torna dependente de um estoque externo. O uso da L-asparaginase por pacientes de LLA catalisa a hidrólise de asparagina em ácido aspártico e amônia (Figura 2), eliminando a asparagina circulante necessária para a sobrevivência de células tumorais, culminando em sua apoptose (STAMS *et al.*, 2005; NARTA *et al.*, 2007). Células saudáveis, capazes de produzir asparagina endogenamente, permanecem inafetadas.

Figura 2 – Representação da hidrólise de L-asparagina por meio da ação da enzima L-asparaginase



Fonte: Elaborado pela autora.

Atualmente, três formulações de L-asparaginase tipo II bacteriana são usadas: uma formulação nativa de *E. coli*, uma formulação de *E. coli* covalentemente ligada a Polietilenoglicol (PEG) e uma de *Erwinia chrysanthemi*, aprovada pelo FDA em 2011 (Tabela 1). Essas L-asparaginases foram selecionadas por sua potência terapêutica e baixa toxicidade relativa quando comparadas com demais fontes bacterianas (DUVAL *et al.*, 2002).

A forma nativa de *E. coli* é a mais usada globalmente e aplicada como primeira linha

Tabela 1 – Formulações de L-asparaginase tipo II utilizadas no tratamento de LLA

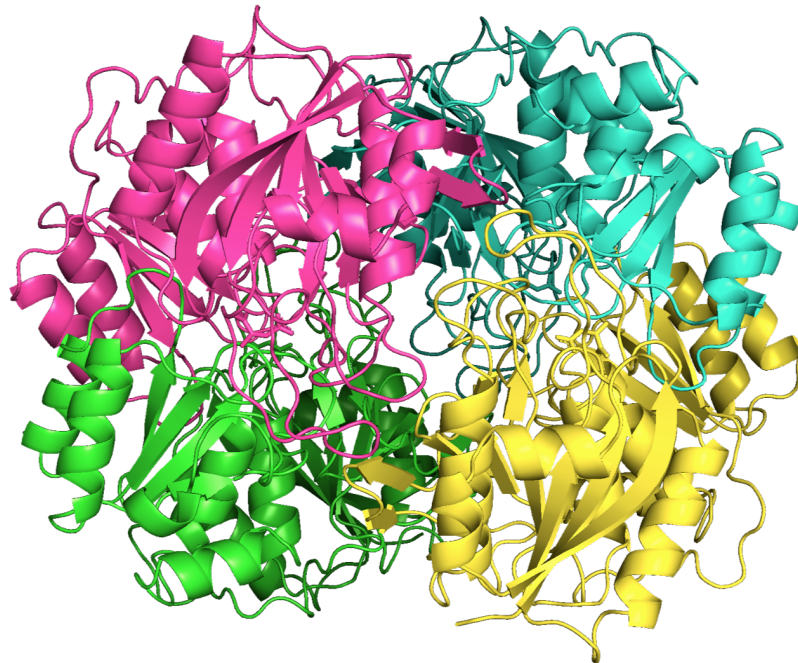
Formulação	Dose	Frequência	Meia-Vida
L-asparaginase nativa de <i>Escherichia coli</i>	6000 UI/m ²	3 vezes na semana	26 a 30 horas
PEG-asparaginase de <i>Escherichia coli</i>	2000–2500 UI/m ²	A cada 2 ou 4 semanas	5 a 7 dias
L-asparaginase nativa de <i>Erwinia chrysanthemi</i>	6000 UI/m ²	3 vezes na semana	16 horas

Fonte: Elaborado pela autora.

de tratamento na Europa (MULLER; BOOS, 1998; DUVAL *et al.*, 2002). A PEG-asparaginase de *E. coli*, que é usada como terapia de primeira linha nos EUA, tem meia-vida cinco vezes maior do que sua forma nativa e por esse motivo uma dosagem menor é necessária (ETTINGER *et al.*, 1994). A formulação de L-asparaginase de *Erwinia* é usada como uma opção a pacientes que apresentam efeitos colaterais nos tratamentos anteriores, pois apresenta menor imunogenicidade e menos complicações de coagulação sanguínea; contudo, sua principal desvantagem é a meia-vida de apenas 16 horas (DUVAL *et al.*, 2002; RIZZARI *et al.*, 2013) (Tabela 1). O uso da L-asparaginase no tratamento de Leucemia Linfoblástica Aguda permitiu elevar a taxa de sobrevivência dos pacientes acima de 90% (SCHMIEGELOW *et al.*, 2010; PUI *et al.*, 2015).

A L-asparaginase pode ser encontrada em diferentes estados funcionais, como monômeros, octômeros e dodecâmeros, mas a estrutura quaternária em que a enzima é mais eficaz é um homotetrâmero (PIETERS *et al.*, 2008). Este tetrâmero (Figura 3) representa a maior proporção contida nas formulações comerciais e é melhor descrito como sendo um dímero de dímeros: dois monômeros se ligam de maneira próxima e formam um par denominado dímero íntimo, e dois dímeros íntimos são combinados formando o tetrâmero catalítico final (GESTO *et al.*, 2013). O sítio de ligação é parcialmente formado quando o monômero se enovela em sua estrutura terciária, e é completado quando dois monômeros se juntam no dímero íntimo. Como resultado, a cavidade catalítica repousa parcialmente na interface dimérica e, por isso, modificações realizadas na interface do dímero íntimo podem alterar a atividade da enzima (OFFMAN *et al.*, 2011).

Figura 3 – Estrutura cristalográfica de L-asparaginase de *Escherichia coli* no estado homotetramérico



Fonte: *Protein Data Bank* (PDB): 3ECA. Elaborado pela autora.

1.3 Imunogenicidade

Proteínas terapêuticas têm a vantagem de serem mais específicas e menos tóxicas quando comparadas a outras moléculas terapêuticas (WAN, 2016). Entretanto, quando administrada a pacientes, essas proteínas apresentam o potencial de incitar respostas imunes que podem impactar negativamente a atividade da proteína, afetando também a segurança e eficácia do tratamento (WAN, 2016; AHMAD *et al.*, 2016). A capacidade de uma substância ou seus componentes de induzir uma resposta imunológica específica quando exposta a elementos do sistema imune denomina-se imunogenicidade (FLOWER, 2007).

Proteínas com potencial terapêutico são produzidas em linhagens celulares que são derivadas de fontes variadas, incluindo mamíferos, insetos, bactérias, plantas e fungos. Pequenas diferenças na sequência da proteína ou modificações pós-traducionais como glicosilação, oxidação, desaminação, acilação e alquilação podem contribuir para a imunogenicidade de uma proteína (DEEHAN *et al.*, 2015). Além disso, durante o processo de produção, que envolve múltiplas etapas (regulação da expressão gênica, purificação, concentração, formulação e estabilização a longo prazo) existe o potencial de introdução de modificações bioquímicas e biofísicas que podem influenciar o perfil imunogênico do produto biológico (SCHELLEKENS, 2002).

A rota de administração também é um fator contribuinte para a imunogenicidade

de proteínas terapêuticas. Proteínas que são administradas por via intraperitoneal, intravenosa, oral ou através de aerossol são menos imunogênicas do que as que são administradas por via sub-cutânea ou intradermal, que se assemelham à imunização ativa (RIZZARI *et al.*, 2013). O risco de imunogenicidade também pode aumentar com a frequência do tratamento (BRYSON *et al.*, 2010).

Além disso, a imunogenicidade também pode ser causada devido à presença de impurezas e possíveis contaminantes nas amostras ou devido a alterações estruturais que podem ocorrer durante o processo de produção (SCHELLEKENS, 2002). O armazenamento e manipulação imprópria também podem desnaturar as proteínas e torná-las mais imunogênicas (FLOWER, 2007).

Portanto, a atividade imunogênica de proteínas terapêuticas pode ser atribuída a muitos fatores que estão relacionados não somente às características da proteína, mas incluem também fatores relacionados à produção, a própria administração durante terapia e características do paciente (Tabela 2).

Tabela 2 – Fatores que contribuem para a imunogenicidade de proteínas terapêuticas

Características de Produção	Características da Proteína	Condução da Terapia	Características do Paciente
Célula hospedeira	Modificações pós-traducionais	Dosagem	Predisposição genética
Aspectos físico-químicos da produção	Diferenças na sequência entre a proteína terapêutica e endógena	Frequência de administração	Exposição prévia a terapêutico similar
Processo de purificação	Epítomos imunodominantes	Rota de administração	Estado imune

Fonte: Elaborado pela autora.

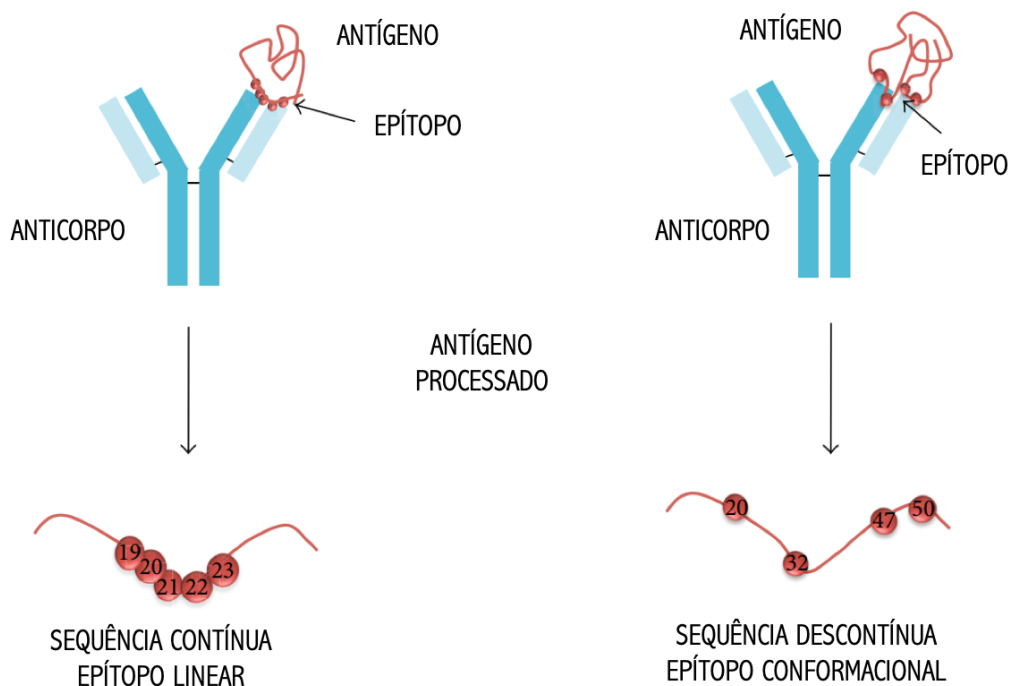
1.3.1 Desenvolvimento da resposta imunológica contra proteínas terapêuticas

A resposta imunológica contra proteínas terapêuticas pode se desenvolver de duas maneiras: através da imunidade inata ou através da imunidade adaptativa/adquirida. A imunidade inata corresponde a uma resposta imediata contra patógenos, porém inespecífica. Em repetidas exposições, os estímulos gerados pela resposta imune inata serão os mesmos. Por outro lado, a imunidade adaptativa gera memória imunológica, age de maneira específica, porém mais

demorada devido à necessidade de recrutamento de componentes da imunidade inata para ser ativada, e a cada exposição, há um aumento da magnitude e capacidade de defesa contra um patógeno específico (ABBAS *et al.*, 2015). As substâncias que entram no organismo e induzem a resposta imune, sejam patógenos ou corpos estranhos, são denominadas de antígenos, e as porções determinantes dessas substâncias são chamadas de epítopos (SILVERTHORN, 2017).

Epítopos podem ser classificados em dois tipos principais baseados em sua estrutura: lineares ou conformacionais (Figura 4). Epítopos lineares se encontram dispostos de forma contínua na estrutura primária da proteína, ao passo que epítopos conformacionais são formados por segmentos de aminoácidos que estão em forma descontínua na sequência e são reconhecidos apenas após o enovelamento da proteína, em sua estrutura terciária ou quaternária (ABBAS *et al.*, 2015). Linfócitos T reconhecem apenas epítopos lineares, enquanto que linfócitos B reconhecem cerca de 10% de epítopos lineares e 90% de epítopos conformacionais (FLOWER, 2007).

Figura 4 – Classificação de epítopos baseada em sua estrutura tridimensional



Fonte: Adaptado de Sanchez-Trincado *et al.* (2017).

A imunidade inata ocorre em todos os animais, é inespecífica e atua sobre qualquer agente agressor. É composta por barreiras físicas (epitélios), barreiras químicas (substâncias antimicrobianas produzidas pelas células epiteliais), células que realizam o processo da fagocitose (macrófagos e neutrófilos), células dendríticas e células exterminadoras naturais (NK), citocinas (proteínas produzidas por diversos tipos celulares que regulam as atividades das células da

imunidade inata) e demais proteínas do sistema complemento (ABBAS *et al.*, 2015).

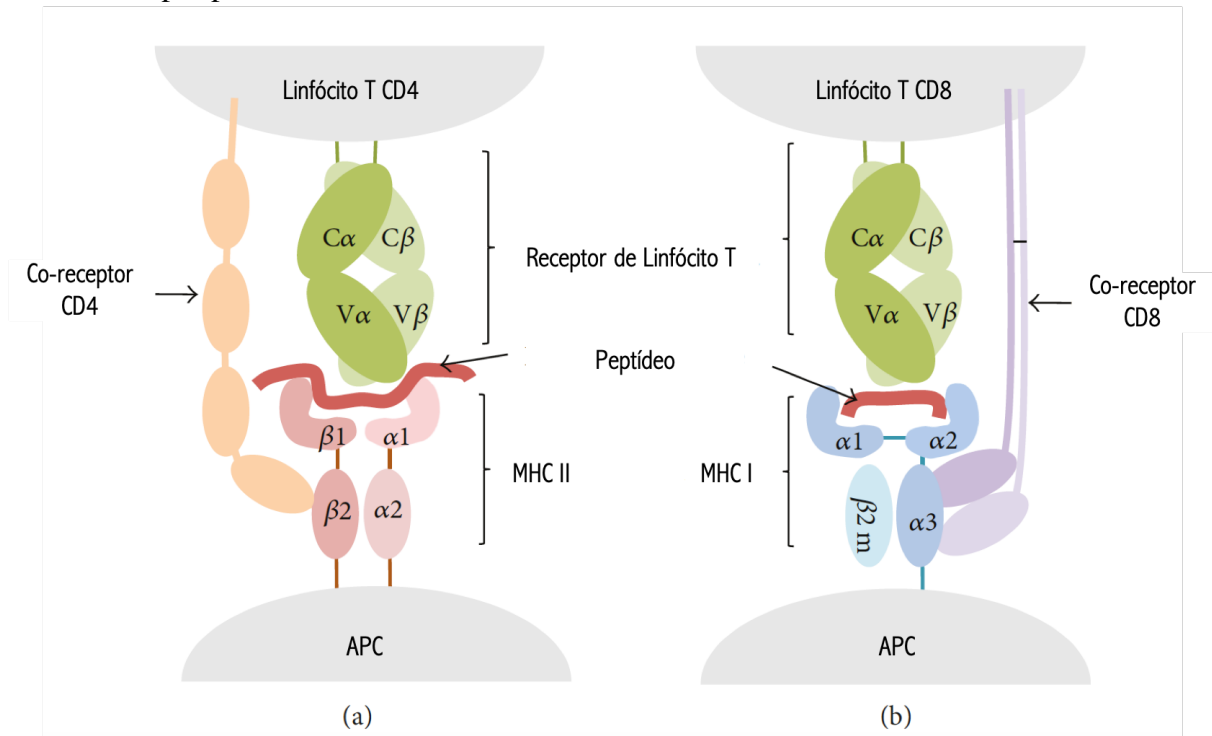
A resposta imune inata fornece sinais que atuam em conjunto com o antígeno para estimular a proliferação e diferenciação de linfócitos B e T específicos contra o agente agressor. Dessa forma, à medida que a resposta imune está fornecendo a defesa inicial contra patógenos, ela também põe em movimento a resposta imune adaptativa. A ativação dos linfócitos necessita de dois sinais distintos, sendo primeiramente o antígeno, e em seguida, as moléculas que são produzidas durante as respostas imunes inatas (ABBAS *et al.*, 2015).

Linfócitos B reconhecem antígenos expostos ao solvente através de Receptores de Linfócitos B (BCR), que consistem em imunoglobulinas ligadas à membrana. Após ativação, os linfócitos B se diferenciam e secretam formas solúveis de imunoglobulinas, conhecidas como anticorpos, que irão mediar a imunidade humoral adaptativa. Anticorpos produzidos por linfócitos B podem apresentar diferentes funções, as quais são desencadeadas pela ligação a seus respectivos antígenos, que incluem a neutralização de toxinas e patógenos e sua identificação para destruição (SILVERTHORN, 2017). Um epítipo de célula B é a porção do antígeno que se liga ao BCR ou anticorpo. Epítopos reconhecidos por células B podem estar em qualquer região exposta ao solvente no antígeno e podem apresentar natureza química diversa (SANCHEZ-TRINCADO *et al.*, 2017).

Por outro lado, linfócitos T reconhecem apenas antígenos expostos por receptores específicos presentes em sua superfície, chamados de Receptores de Linfócitos T (TCR). O TCR é um heterodímero composto de uma cadeia α e uma β de tamanho aproximadamente igual. Ambas as cadeias têm uma região transmembrana que compreende os aminoácidos hidrofóbicos pelos quais a molécula é ancorada na membrana celular, além de uma região constante (C) e uma região variável (V), semelhante às cadeias de imunoglobulina (Figura 5). A região variável de ambas as cadeias determinam a especificidade para o antígeno. Os linfócitos T têm uma especificidade restrita para antígenos; reconhecem apenas peptídeos derivados das proteínas estranhas que estão ligadas às proteínas do hospedeiro, denominadas moléculas do Complexo Principal de Histocompatibilidade (MHC) (ABBAS *et al.*, 2015).

Os genes que codificam para MHCs em humanos são chamados de *Human Leukocyte Antigen* (HLA) (sistema antígeno leucocitário humano). Os MHCs exercem a função de se ligarem aos antígenos peptídicos e apresentá-los para o reconhecimento por linfócitos T específicos, e como resultado, essas células T reconhecem e respondem aos antígenos associados à superfície celular (Figura 5).

Figura 5 – Representação da estrutura de MHC de classes I e II em um complexo de ligação epítipo-MHC-linfócito T



Os epítomos de células T são peptídeos derivados de antígenos e são reconhecidos pelo Receptor de Linfócitos T (TCR) quando ligado a moléculas de MHC exibidas na superfície celular de APCs. (a) Linfócitos T CD4 expressam o co-receptor CD4, que se liga ao MHC II, e reconhece peptídeos apresentados por moléculas de MHC II. (b) Linfócitos T CD8 expressam o co-receptor CD8, que se liga ao MHC I, e reconhece peptídeos apresentados por moléculas de MHC I. $C\alpha$ e $C\beta$ correspondem a regiões constantes de cadeias α e β de TCR, ao passo que $V\alpha$ e $V\beta$ correspondem a regiões variáveis das cadeias α e β de TCR.

Fonte: Adaptado de Sanchez-Trincado *et al.* (2017).

Epítomos de células T podem ser apresentados por moléculas de MHC de classe I ou II. MHCs de classe I são moléculas presentes em todas as células nucleadas do corpo humano, enquanto que apenas Células Apresentadoras de Antígenos (APC) possuem também o MHC da classe II. Ambas as classes de moléculas do MHC consistem em uma fenda de ligação ao peptídeo extracelular, uma região tipo imunoglobulina não polimórfica, uma região transmembrana e uma região citoplasmática (ABBAS *et al.*, 2015). MHCs são reconhecidos por dois subconjuntos de linfócitos T: CD8 e CD4. Células T CD8 se tornam citotóxicas e interagem com epítomos expostos por MHC I, ao passo que células T CD4 podem se tornar auxiliares (amplificam a resposta imune) ou reguladoras (emitem sinais de inibição à resposta imune) e reconhecem epítomos ligados ao MHC II. Portanto, há epítomos de CD8 e de CD4, que se diferenciam por sua ligação ao MHC de classe I e de classe II, respectivamente (Figura 5).

As moléculas de MHC I possuem uma cadeia α longa e uma cadeia β curta chamada

β 2-microglobulina (Figura 5). A cadeia α contém uma região de ligação ao peptídeo altamente polimórfica formada a partir dos domínios α 1 e α 2 (ABBAS *et al.*, 2015). O co-receptor de CD8 se liga ao domínio α 3, altamente conservado, ao qual a cadeia β 2-microglobulina associa-se (Figura 5), auxiliando a manter a conformação apropriada da molécula (ABBAS *et al.*, 2015).

As moléculas de MHC classe II, no entanto, são compostas de duas cadeias polipeptídicas α e β de tamanho aproximadamente igual (Figura 5). A região de ligação a peptídeo é altamente polimórfica, formada pelos domínios α 1 e β 1 (ABBAS *et al.*, 2015). Os domínios α 2 e β 2, altamente conservados, formam um bolsão ao qual o co-receptor de CD4 se liga (ABBAS *et al.*, 2015).

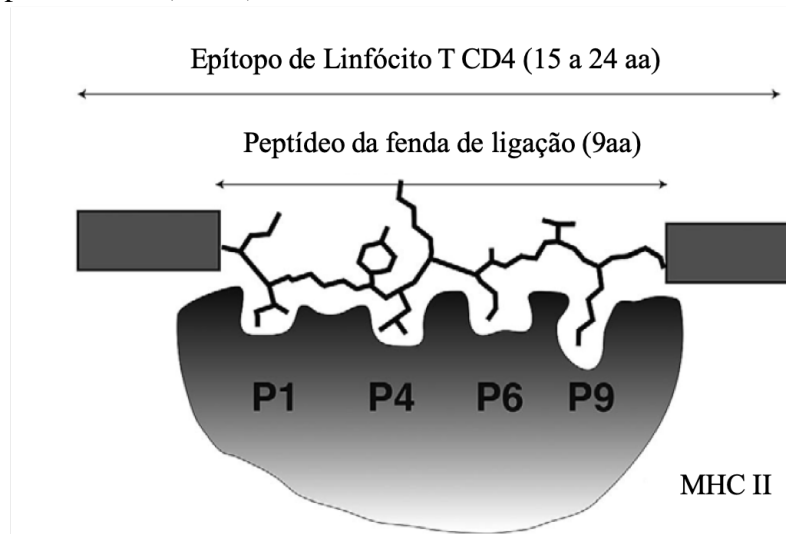
Cada molécula do MHC de classe I ou II tem uma única fenda de ligação que se liga a um peptídeo de cada vez, mas cada molécula de MHC pode se ligar a diversos peptídeos diferentes (ABBAS *et al.*, 2015). Os peptídeos que se ligam a moléculas de MHC possuem características estruturais semelhantes que promovem esta interação. Uma dessas características é o tamanho do peptídeo. A fenda de ligação de MHC classe I é mais fechada, permitindo que apenas epítomos mais curtos de aproximadamente 8 a 12 aminoácidos se liguem (BACKERT; KOHLBACHER, 2015), enquanto que a fenda de ligação de MHC classe II é aberta, podendo acomodar peptídeos mais longos de aproximadamente 15 a 24 aminoácidos, com alguns dos resíduos localizados fora da fenda de ligação (BACKERT; KOHLBACHER, 2015).

O que determina se um peptídeo irá se ligar à fenda de MHC ou não são os aminoácidos que estiverem localizados dentro da fenda de ligação, que se limitam a um conjunto de nove resíduos (SANCHEZ-TRINCADO *et al.*, 2017). Esses nove aminoácidos definem a afinidade e estabilidade à fenda de ligação de MHC e, por essa razão, métodos de predição de epítomos lineares consideram apenas peptídeos de nove aminoácidos de comprimento (nonâmeros). Os demais aminoácidos que compõem os epítomos são conhecidos como flanqueadores, responsáveis por ajudar na estabilidade do complexo através de ligações de hidrogênio (JAWA *et al.*, 2013).

Além disso, outros resíduos de aminoácidos de um peptídeo podem conter cadeias laterais que se encaixam em bolsões específicos e se ligam aos aminoácidos complementares na molécula do MHC por meio de interações eletrostáticas, ligações de hidrogênio ou interações de Van der Waals (ABBAS *et al.*, 2015). Estes resíduos peptídicos são chamados de resíduos de ancoragem, porque contribuem com a maior parte das interações favoráveis para a ligação (isto é, ancoram o peptídeo na fenda da molécula do MHC)(SILVERTHORN, 2017)

As posições de ancoragem na fenda de ligação do MHC II são determinantes para

Figura 6 – Representação da estrutura da fenda de ligação do Complexo Principal de Histocompatibilidade (MHC) II



As posições P1, P4, P6 e P9 que correspondem às cavidades na fenda de ligação das moléculas de MHC II são determinantes para a ligação do epítipo à fenda de MHC II.

Fonte: Adaptado de Malherbe (2009).

a ligação a um peptídeo, pois representam os maiores valores de contribuição para o peptídeo ser considerado epítipo. Âncoras estão presentes em todos os motivos de ligação a HLA, e são limitadas a um grupo de aminoácidos que compartilha características físico-químicas como polaridade, carga ou hidrofobicidade (HAMMER *et al.*, 1994). A interação entre o epítipo de célula T CD4 e a molécula de MHC II, por exemplo, é mediada principalmente pela contribuição de resíduos-âncora nas posições P1, P4, P6 e P9, que correspondem às cavidades na fenda de ligação das moléculas de MHC II (Figura 6).

A contribuição de linfócitos T CD4 é de grande importância em estratégias de desimunização, pois essas células são as responsáveis pela indução da produção de anticorpos anti-terapêuticos e pela produção de sinais inibitórios em resposta a epítopos (SHARMA; HOLT, 2014). Portanto, a identificação de epítopos de células T CD4 é a de maior interesse numa abordagem que objetive a redução da imunogenicidade (RIZZARI *et al.*, 2014).

Linfócitos T CD4 possuem pelo menos três *loci* receptores de antígeno: DR, DQ, e DP. As fendas de ligação das moléculas codificadas por esses genes são capazes de interagir com motivos de ligação específicos. O conhecimento acerca das fendas codificadas por esses genes polimórficos e suas frequências é de extrema importância para realizar a predição dos peptídeos que serão apresentados pelos MHCs aos linfócitos T (BACKERT; KOHLBACHER, 2015).

Muitos complexos MHC-DR compartilham o mesmo padrão de ligação à molécula

de MHC, o que indica que sua promiscuidade (inespecificidade) é maior do que a alelos de MHC classe I. Portanto, a identificação de epítomos de MHC classe II pode ser um forte indicador da resposta imunológica de linfócitos T, pois eles representam regiões da proteína em que as sequências possuem alta afinidade com múltiplos alelos de MHC e onde múltiplos padrões estão localizados (SANCHEZ-TRINCADO *et al.*, 2017). Existem poucos dados na literatura acerca da relação dos alelos MHC-DP e MHC-DQ com a imunogenicidade, o que pode ser devido ao fato de eles serem menos definidos do que os alelos MHC-DR, ou ao fato de que esses alelos apresentam uma baixa contribuição para a imunogenicidade (SHARMA; HOLT, 2014).

1.3.2 Efeitos adversos da administração de L-asparaginase

Apesar do avanço significativo no tratamento de Leucemia com L-asparaginase, a terapia promove uma série de efeitos adversos prejudiciais incluindo respostas imunológicas, variando de reações alérgicas ao choque anafilático fatal, distúrbios de coagulação, pancreatite, hiperglicemia, hepatotoxicidade e inibição da síntese de proteínas (SHRIVASTAVA *et al.*, 2015). Além disso, pode ocorrer também uma diminuição da atividade antitumoral de L-asparaginase, que está diretamente associada à produção de anticorpos contra a enzima pelo paciente, o que leva à redução de sua concentração na corrente sanguínea prejudicando a eficácia do tratamento (LOPES *et al.*, 2017). Esses sintomas desencadeados pela administração de L-asparaginase podem se enquadrar em duas categorias principais: os que pertencem à hipersensibilidade à proteína exógena e os efeitos adversos relacionados à inibição da síntese proteica (STRICKLER *et al.*, 2018).

A manifestação clínica mais comum da hipersensibilidade é a urticária, sendo também relatadas indurações, edema, inchaço, frio, febre, sensibilidade e erupções cutâneas (NARTA *et al.*, 2007). Existem casos em que a urticária pode ser tratada ou prevenida através de anti-histamínicos. No entanto, reações mais severas geralmente exigem a descontinuação da formulação administrada. Tais reações têm de 13 a 30% de frequência de ocorrência causada por L-asparaginase de *E. coli* (AMYLON *et al.*, 1999). Fatores de risco para hipersensibilidade à L-asparaginase incluem doses acima 6000 UI/m² por dia via intravenosa, repetições frequentes do tratamento e utilização da enzima como agente único de tratamento, ao invés da utilização de uma quimioterapia combinada (TRUEWORTHY *et al.*, 1978).

Em caso de toxicidades decorrentes da síntese proteica inibida, os tecidos com elevadas taxas de síntese proteica (fígado, pâncreas e sistema de coagulação) são os mais

frequentemente afetados pela terapia com L-asparaginase (OHNUMA *et al.*, 1970). A maioria dos pacientes apresentam evidências de hepatotoxicidade. Os efeitos no sistema de coagulação, evidenciados por desequilíbrios na formação de fatores de coagulação, também são aspectos comuns associados à terapia com L-asparaginase. A pancreatite é uma complicação comumente documentada, ocorrendo em cerca de 10 a 16% dos pacientes com taxa de mortalidade entre 1,8 e 4,6% (OHNUMA *et al.*, 1970). Os pacientes também podem desenvolver sinais e sintomas de diabetes devido à danos causados nas ilhotas pancreáticas e diminuição subsequente da síntese de insulina. A hiperglicemia pode se tornar mais severa quando a enzima é administrada em combinação com prednisona. Porém, o risco pode ser reduzido se a administração de prednisona for realizada antes de L-asparaginase (SUTOW *et al.*, 1976). A neurotoxicidade, que pode ser um resultado da ausência de L-asparagina e L-glutamina no cérebro, é um sintoma ocorre em aproximadamente 25% dos pacientes adultos tratados com L-asparaginase, contudo é um sintoma raro em crianças (STRICKLER *et al.*, 2018).

A administração a pacientes geralmente resulta em uma rápida decadência nos níveis de L-asparaginase sanguínea devido a reação imunológica de inativação da enzima exógena, levando à necessidade de uma alta frequência de administração, o que pode favorecer a produção de anticorpos anti-terapêuticos (FERNANDES *et al.*, 2016). A produção de anticorpos contra L-asparaginase foi observada em cerca de 60% dos pacientes tratados com L-asparaginase de *E. coli* (SHRIVASTAVA *et al.*, 2015), o que resulta na redução da concentração ótima da proteína na corrente sanguínea, e conseqüentemente na diminuição da eficácia do tratamento. Este quadro representa um dos maiores desafios em relação ao uso de proteínas terapêuticas.

A identificação de sequências peptídicas em proteínas terapêuticas que são reconhecidas como epítomos por algum MHC é de grande importância para o entendimento da origem de doenças, monitoramento da resposta imune, desenvolvimento de ensaios diagnósticos e *design* de vacinas, além de ser necessária para a determinação do potencial imunogênico de terapêuticos, servindo como ponto de partida em inúmeras abordagens de desimunização (AHMAD *et al.*, 2016).

1.4 Mapeamento de Epítomos Lineares

O mapeamento de epítomos lineares se refere ao processo de identificação das menores sequências de aminoácidos de um antígeno específico que é reconhecido por receptores de células T, e ao mesmo tempo, têm potencial de incitar uma resposta imune citotóxica duradoura

(AHMAD *et al.*, 2016). A predição de epítomos é baseada nas interações entre aminoácidos da proteína e moléculas de MHC de um indivíduo. Quanto mais estável a ligação de um peptídeo à fenda de ligação de uma molécula de MHC, maior a probabilidade de que esse peptídeo incite uma resposta imune (WEAVER *et al.*, 2008).

A principal limitação comum a todos os métodos de análise de ligação de MHC de classe II (*in silico* e *in vitro*) são o grande número de peptídeos falso-positivos que são identificados como ligantes de MHC classe II, mas que não estimulam resposta de células T nem *in vitro* quanto *in vivo* (BAKER *et al.*, 2010). Isso ocorre em grande parte devido à dificuldade de se considerar e prever todos os fatores que influenciam a formação de epítomos. Etapas como o processamento da proteína, o reconhecimento de epítomos por TCRs e a tolerância de células T a determinados peptídeos são fatores complexos e extremamente variados para cada proteína e cada organismo (AHMAD *et al.*, 2016).

Para a identificação de epítomos em larga escala foram desenvolvidos algoritmos que preveem a capacidade de um peptídeo se ligar à molécula de MHC e quantificam a afinidade desta ligação. Em adição à vantagem de se rastrear rapidamente um grande número de sequências, estes métodos são consideravelmente menos dispendiosos do que ensaios *in vitro*, que consomem mais tempo e demandam mais recursos (GROOT; SCOTT, 2007; BAKER *et al.*, 2010). As ferramentas *in silico* atuais são precisas quanto à predição de peptídeos de ligação ao MHC II contra um grande número de alelos de MHC (PARKER *et al.*, 2010; BRENNAN *et al.*, 2010; GUPTA *et al.*, 2011; DROUIN *et al.*, 2013).

A Tabela 3 indica preditores de epítomos lineares de ligação ao MHC II que se encontram disponíveis online para uso público. Essas ferramentas permitem a identificação de epítomos de inúmeras sequências de proteínas com base na afinidade por diversos alelos de MHC e podem ser categorizadas como as que se baseiam em critérios (i) *bioquímicos*, utilizando metodologias como motivos de ligação, *docking* e relação estrutura-atividade; e as que se baseiam em critérios (ii) *estatísticos*, se utilizando algoritmos de aprendizagem de máquina, como máquina de vetores de suporte e rede neural artificial. Isso significa que uma função de predição é construída com base em um conjunto de dados para treinamento.

Métodos *bioquímicos* se valem da existência de aminoácidos em determinadas posições dos epítomos, conhecidos como resíduos-âncora (THAKALLAPALLY *et al.*, 2000). As primeiras predições da ligação peptídeo-MHC foram realizadas usando este fundamento através dos chamados Motivos de Sequência (SM), que refletem as preferências das moléculas de MHC

Tabela 3 – Preditores de epítomos de linfócitos T CD4+ disponíveis online para uso público

Ferramenta	URL	Método	Referência
EpiTOP	http://www.pharmfac.net/EpiTOP	QSAR	(DIMITROV <i>et al.</i> , 2010)
IEDB-MHCII	http://tools.immuneepitope.org/mhcii/	Consenso	(ZHANG <i>et al.</i> , 2008)
IL4pred	http://webs.iiitd.edu.in/raghava/il4pred/index.php	SVM	(DHANDA <i>et al.</i> , 2013)
MHCPred	http://www.ddg-pharmfac.net/mhcpred/MHCPred/	QSAR	(GUAN <i>et al.</i> , 2003)
MHC2PRED	http://www.imtech.res.in/raghava/mhc2pred/index.html	SVM	(BHASIN; RAGHAVA, 2004)
Motif Scan	https://www.hiv.lanl.gov/content/immunology/motif_scan/motif_scan	SM	(THAKALLAPALLY <i>et al.</i> , 2000)
MULTIPRED2	http://cvc.dfci.harvard.edu/multipred2/index.php	ANN	(ZHANG <i>et al.</i> , 2011)
NetMHCII	http://www.cbs.dtu.dk/services/NetMHCII/	ANN	(NIELSEN <i>et al.</i> , 2007a)
NetMHCIIpan	http://www.cbs.dtu.dk/services/NetMHCIIpan/	ANN	(NIELSEN <i>et al.</i> , 2008)
PREDIVAC	http://predivac.biosci.uq.edu.au/	MM	(OYARZUN <i>et al.</i> , 2013)
Propred	http://www.imtech.res.in/raghava/propred/	QAM	(SINGH; RAGHAVA, 2001)
Rankpep	http://imed.med.ucm.es/Tools/rankpep.html	MM	(RECHE <i>et al.</i> , 2004)
SYFPEITHI	http://www.syfpeithi.de/	MM	(RAMMENSEE <i>et al.</i> , 1999)
SVMHC	http://abi.inf.uni-tuebingen.de/Services/SVMHC/	SVM	(DONES; KOHLBACHER, 2006)
SVRMHC	http://us accurascience.com/SVRMHCdb/	SVM	(WAN <i>et al.</i> , 2006)
TEPITOPE	http://datamining-iip.fudan.edu.cn/service/TEPITOPEpan	QAM	(STURNIOLO <i>et al.</i> , 1999)
Vaxign	http://www.violinet.org/vaxign/	MM	(XIANG; HE, 2009)

ANN: Rede Neural Artificial; DS-QM: Matriz Quantitativa baseada em Escores de Docking; Consenso: Combinação de Métodos; MM: Matriz de Motivos; SM: Motivo de Sequência; SVM: Máquinas de Vetores de Suporte; QAM: Matriz de Afinidade Quantitativa; QSAR: Modelo Quantitativo de Relação Estrutura-Atividade.

Fonte: Elaborado pela autora.

por aminoácidos em posições de ancoragem (D'AMARO *et al.*, 1995). Posteriormente, foram desenvolvidas Matrizes de Motivos (MM), que avaliam a contribuição de cada aminoácido

em cada posição da fenda de ligação ao MHC através de pontuações (NIELSEN *et al.*, 2004; RAMMENSEE *et al.*, 1999). A versão mais atualizada de Matrizes de Motivo consiste em perfis que são semelhantes àqueles usados para detectar homologia de sequências (RECHE *et al.*, 2004; XIANG; HE, 2009; OYARZUN *et al.*, 2013).

As Matrizes de Motivos não levam em consideração valores de afinidades de ligação nas pontuações da matriz e, portanto, as pontuações dos peptídeos resultantes não são adequadas para indicar a afinidade de ligação. Por isso, Matrizes Quantitativas de Afinidade (QAM) são treinadas em peptídeos com suas correspondentes afinidades de ligação, e objetivam prever a afinidade de ligação do complexo peptídeo-MHC (PARKER *et al.*, 1994). Várias abordagens foram desenvolvidas e aprimoradas para obter QAMs a partir de dados de afinidade de peptídeos e prever a ligação de peptídeos ao MHC (STURNIOLO *et al.*, 1999; SINGH; RAGHAVA, 2001).

Foi demonstrado que os resíduos não-âncora também contribuem para a capacidade de um peptídeo se ligar a uma dada molécula de MHC (BOUVIER; WILEY, 1994). Para calcular interferências que resíduos vizinhos de peptídeos possuem na afinidade de ligação foi introduzido o Modelo Quantitativo de Relação Estrutura-Atividade (QSAR), em que a afinidade dos peptídeos pelo MHC é calculada como a soma das contribuições dos aminoácidos em cada posição adicionada à contribuição de interações de cadeias laterais adjacentes (GUAN *et al.*, 2003).

Alternativamente, são utilizados métodos baseados em *aprendizagem estatística*, como Rede Neural Artificial (ANN) e Máquinas de Vetores de Suporte (SVM) (AHMED; MAEURER, 2009), treinados em conjuntos de dados que consistem em peptídeos com afinidade conhecida para moléculas de MHC.

Um estudo comparativo foi realizado para avaliação dos métodos de predição de epítomos lineares contra um vasto banco de dados (mais de 10.000 dados) de afinidades de ligação de MHC classe II-peptídeo obtido experimentalmente (WANG *et al.*, 2008). Para a avaliação dos métodos de predição foram escolhidas ferramentas capazes de realizar predições para MHC classe II que estivessem disponíveis para uso público e que permitissem submissão em larga escala, totalizando nove métodos que correspondiam a esses critérios (Tabela 4).

A performance geral das ferramentas selecionadas foi avaliada calculando-se a média da afinidade de ligação para todas as moléculas de MHC II disponíveis comparando com o banco de dados obtido experimentalmente que ainda não havia sido publicado. Os resultados

Tabela 4 – Métodos de predição de epítomos de MHC classe II em larga escala

Ferramenta	Método	Referência
ARB	Matriz de Ligação Relativa Média (ARB)	(BUI <i>et al.</i> , 2005)
PRORED	Matriz TEPITOPE	(SINGH; RAGHAVA, 2001)
SVMHC	SVM	(DONNES; KOHLBACHER, 2006)
SYFPEITHI	Matrizes de Pontuação de Posição Específica	(RAMMENSEE <i>et al.</i> , 1999)
RANKPEP	Matrizes de Pontuação de Posição Específica	(RECHE <i>et al.</i> , 2004)
SMM-align	Matriz Estabilizada	(NIELSEN <i>et al.</i> , 2007b)
SVRMHC	Regressão SVM	(WAN <i>et al.</i> , 2006)
MHC2PRED	SVM	(BHASIN; RAGHAVA, 2004)
MHCPRED	Regressão QSAR	(GUAN <i>et al.</i> , 2003)

Fonte: Adaptado de Wang *et al.* (2008).

foram representados através de curvas Área sob a Curva ROC (AUC), que indicam a relação entre a sensibilidade e a especificidade de um teste diagnóstico quantitativo, ao longo de um contínuo de valores de corte, para se classificar e contabilizar o número de predições positivas e negativas. Para cada ponto de corte são calculados valores de sensibilidade e especificidade, que são dispostos em um gráfico denominado AUC (CRISTIANO, 2017). SMM-align e PRORED apresentaram predições mais acuradas, com resultado de AUC de 0,73, seguido do método ARB, que obteve o terceiro melhor resultado, com média de AUC de 0,71 (WANG *et al.*, 2008).

Posteriormente foi analisado se os métodos selecionados seriam capazes de identificar com acurácia os nonâmeros que se ligam à fenda de ligação. Um total de 29 estruturas associadas a 14 moléculas diferentes de MHC II foram obtidas a partir do PDB para tal avaliação (WANG *et al.*, 2008). Foi feita a comparação de cada resíduo obtido por cada ferramenta com os resíduos originais extraídos das estruturas cristalográficas. PRORED obteve a melhor performance, tendo sido capaz de prever corretamente todos os resíduos presentes nas estruturas de PDB (WANG *et al.*, 2008). O PRORED utiliza as matrizes de afinidade do TEPITOPE (STURNIOLO *et al.*, 1999), que será descrito detalhadamente a seguir.

1.4.1 TEPITOPE

O TEPITOPE é um método de predição de epítomos em proteínas que utiliza matrizes de afinidade quantitativa. Essas matrizes de afinidade foram criadas a partir de dados experimentais obtidos da afinidade entre peptídeos e diferentes alelos de MHC II. Inicialmente foram realizadas medições da afinidade de ligação individual entre cada aminoácido e a fenda de ligação dos diferentes MHC II. Posteriormente, foram criados perfis de ligação para alelos não testados através de alinhamentos sequencial e estrutural (STURNIOLO *et al.*, 1999).

O TEPITOPE utiliza a soma dos escores específicos para cada aminoácido em uma dada posição em um nonâmero para prever sua probabilidade de se ligar a um MHC II e incitar uma resposta imune (PARKER *et al.*, 2010).

As fendas (ou cavidades) de ligação de MHC II contêm nove resíduos no núcleo de ligação, aos quais porções específicas do epítomo se ligam. Cadeias laterais diferentes podem apresentar efeitos diversos na modulação da afinidade de ligação do epítomo. A vantagem dessa metodologia é que um determinado perfil de ligação conhecido pode ser aplicado a outros alelos que possuem a mesma fenda sem qualquer modificação independentemente do polimorfismo da molécula de MHC (STURNIOLO *et al.*, 1999).

Os alelos de MHC para os quais existem matrizes de afinidade de ligação disponíveis encontram-se listados no Quadro 1.

Quadro 1 – Alelos de MHC para os quais o TEPITOPE disponibiliza matrizes de afinidade de ligação

DRB1*0101	DRB1*0102	DRB1*0301	DRB1*0305	DRB1*0306
DRB1*0307	DRB1*0308	DRB1*0309	DRB1*0311	DRB1*0401
DRB1*0402	DRB1*0404	DRB1*0405	DRB1*0408	DRB1*0410
DRB1*0423	DRB1*0426	DRB1*0701	DRB1*0703	DRB1*0801
DRB1*0802	DRB1*0804	DRB1*0806	DRB1*0813	DRB1*0817
DRB1*1101	DRB1*1102	DRB1*1104	DRB1*1106	DRB1*1107
DRB1*1114	DRB1*1120	DRB1*1121	DRB1*1128	DRB1*1301
DRB1*1302	DRB1*1304	DRB1*1305	DRB1*1307	DRB1*1311
DRB1*1321	DRB1*1322	DRB1*1323	DRB1*1327	DRB1*1328
DRB1*1501	DRB1*1502	DRB1*1506	DRB5*0101	DRB5*0105

Fonte: Elaborado pela autora.

A identificação de epítomos é o primeiro passo para a redução da imunogenicidade de proteínas terapêuticas. Para se alcançar maior eficácia e segurança no tratamento de LLA, estratégias de desimunização de proteínas podem ser aplicadas a partir do mapeamento de epítomos, tais como glicosilação, peguilação, humanização e mutagênese.

1.5 Desimunização

A imunogenicidade é uma das maiores limitações em relação ao uso de proteínas terapêuticas (Seção 1.3). A principal causa da imunogenicidade é a presença de epítomos de linfócitos T na sequência da proteína que podem ativar células T auxiliares, resultando na produção de anticorpos contra a proteína e na neutralização do efeito terapêutico (JONES *et al.*, 2009).

Os anticorpos produzidos contra L-asparaginase são a principal causa da resistência à esta proteína e resultam na redução da atividade enzimática e em efeitos clínicos adversos (RIZZARI *et al.*, 2013; SHRIVASTAVA *et al.*, 2015; AHMAD *et al.*, 2016; FERNANDES *et al.*, 2016). Pacientes com anticorpos anti-asparaginase apresentam quantidades insuficientes da enzima para a efetiva depleção de asparagina e, em alguns casos, a atividade da asparaginase pode ser indetectável (EGLER *et al.*, 2016; LOPES *et al.*, 2017; RIZZARI *et al.*, 2013). Portanto, estratégias para reduzir a hipersensibilidade através da exploração de mutações que diminuam o potencial imunogênico, enquanto preservam o potencial terapêutico e a estabilidade proteica são de extrema importância para assegurar a eficácia e segurança no tratamento da Leucemia Linfoblástica Aguda.

A desimunização de proteínas corresponde à estratégia de localizar e remover ou camuflar epítomos de células T através do uso combinado de técnicas de biologia molecular e imunologia, de modo a reduzir a imunogenicidade total de uma proteína terapêutica (PARKER *et al.*, 2010). O objetivo final desse processo é a mitigação das respostas de anticorpos antifármacos de modo a melhorar a eficácia da proteína terapêutica (ZHAO *et al.*, 2015).

Algoritmos computacionais, bancos de dados de epítomos e avanços nos métodos estatísticos têm possibilitado estratégias para mitigar o potencial de imunogenicidade de proteínas terapêuticas (BRYSON *et al.*, 2010). Diferentes estratégias de redução de imunogenicidade têm sido testadas, e incluem modificação direta da proteína terapêutica através de peguilação ou glicosilação na tentativa de mascarar epítomos imunogênicos (ETTINGER *et al.*, 1994; GROOT; SCOTT, 2007), modificações de resíduos imunodominantes de MHC classe II para inibir sua

apresentação a linfócitos (PARKER *et al.*, 2010; PARKER *et al.*, 2013) e estratégias que ajudem o sistema imune a tolerar a proteína terapêutica (CANTOR *et al.*, 2011; JAWA *et al.*, 2013).

A pegulação e glicosilação são os dois métodos mais comuns de modificações pós-traducionais de proteínas que não impactam sua eficácia terapêutica, e consistem na conjugação da proteína a moléculas de Polietilenoglicol (PEG) e glicose, respectivamente. PEG não é tóxico, não é imunogênico, é altamente solúvel em água e possui a aprovação do FDA para aplicação em biofármacos nos Estados Unidos (GROOT; SCOTT, 2007). Proteínas terapêuticas conjugadas ao PEG ou à glicose possuem uma série de vantagens, como a diminuição de sua degradação por enzimas metabólicas, o que resulta no aumento de sua meia-vida e na redução da imunogenicidade ao diminuir a proteólise, interferindo no processamento da proteína e consequentemente na apresentação de epítomos às moléculas de MHC (GRISWOLD; BAILEY-KELLOGG, 2016). Apesar disso, o uso desses métodos pode resultar na produção de anticorpos anti-PEG e criação de novos epítomos para anticorpos neutralizantes, respectivamente (ZHANG *et al.*, 2014). Além disso, a conjugação a açúcares incomuns podem resultar em um efeito contrário e aumentar a imunogenicidade (SO *et al.*, 1996).

Uma outra abordagem corresponde à humanização, que aumenta a similaridade da proteína terapêutica com a proteína humana, levando à redução parcial da imunogenicidade (LIU *et al.*, 2015). Entretanto, trata-se de um processo dispendioso que muitas vezes não é justificado pela imunogenicidade remanescente, que pode variar de uma imunogenicidade baixa à intolerável (SAFDARI *et al.*, 2013).

Finalmente, é possível recorrer à estratégia utilizada neste trabalho, na qual a sequência da proteína de interesse é diretamente manipulada via mutagênese sítio-dirigida. (PARKER *et al.*, 2010; MEHTA *et al.*, 2014). Mutações na sequência podem reduzir o número de epítomos produzidos pela clivagem da proteína como também diminuir a afinidade dos epítomos pelas moléculas de MHC II dificultando seu reconhecimento pelo sistema imune, o que resulta em uma redução na imunogenicidade geral da proteína (GROOT; SCOTT, 2007).

Para a obtenção de variantes desimunizadas da L-asparaginase de *Escherichia coli*, seria necessária a alteração de regiões reconhecidas como epítomos pelo sistema imune por meio de um conjunto específico de mutações sítio-dirigidas. Contudo, encontrar as exatas posições da sequência e identidade dos resíduos para os quais se deve mutar se torna um desafio quando se leva em conta que é possível mutar em cada uma das posições da sequência para todos os 20 aminoácidos de ocorrência natural. Considerando que cada uma das subunidades da

L-asparaginase possui 326 resíduos temos 20^{326} possíveis mutantes que devem ser avaliados para que o melhor seja encontrado.

Para sanar este problema, é possível se utilizar de algoritmos de otimização que possibilitem encontrar a solução de maneira mais eficiente, além de possibilitar a economia de recursos que seriam necessários em ensaios e experimentos para avaliar a eficiência de mutantes criados aleatoriamente.

1.5.1 Algoritmo Genético

Algoritmos de otimização buscam alcançar a solução ideal para um determinado problema, de forma que se possa alcançar eficientemente os objetivos estabelecidos. A qualidade de uma solução é dada pelo valor que a solução apresenta diante de uma função de avaliação, e sua viabilidade é dada por um conjunto de restrições. Algoritmos evolucionários são procedimentos de otimização que buscam as melhores soluções possíveis para o problema proposto partindo de um conjunto de soluções iniciais e gerando, sucessivamente, novas soluções mais adequadas (YU; GEN, 2010).

Um dos algoritmos de otimização mais utilizados é o Algoritmo Genético (AG), inspirado na teoria da evolução pela seleção natural proposta por Darwin e Wallace, na qual indivíduos mais adaptados possuem maior probabilidade de transmitir seus caracteres para a geração seguinte (HOLLAND, 1975). Em um algoritmo genético, cada indivíduo representa uma solução candidata do problema em questão e o processo de evolução executado corresponde a uma busca multi-direcional nesse espaço de soluções.

Os indivíduos são aprimorados por meio de transformações realizadas por operadores genéticos, permitindo a troca da orientação da direção representada por cada solução. A cada geração, as soluções são avaliadas por uma função de aptidão, de forma que os indivíduos mais adaptados sejam selecionados. Para fazer a distinção entre diferentes soluções, é empregada uma função objetivo que simula o papel da pressão exercida pelo ambiente sobre o indivíduo.

As soluções melhor adaptadas se reproduzem, enquanto que soluções pior adaptadas são eliminadas, permitindo que os indivíduos acumulem, ao longo do tempo evolutivo, um maior número de caracteres de interesse e idealmente se tornem mais aptos a solucionar o problema. Após um número de gerações, a condição de parada deve ser atendida, a qual geralmente indica a existência, na população, de um indivíduo que represente uma solução satisfatória para o problema.

Dessa forma, é possível obter variantes desimunizadas da L-asparaginase de *Escherichia coli* de maneira muito mais eficiente através de uma combinação ótima de mutações na sequência que não impliquem na criação de novos e sejam estruturalmente estáveis.

2 OBJETIVOS

2.1 Objetivo Geral

Desimunizar uma sequência de L-asparaginase de *Escherichia coli*, enzima terapêutica utilizada no tratamento de Leucemia Linfoblástica Aguda, através de mutações em epítomos determinados experimentalmente utilizando-se alelos de HLA relatados na literatura relacionados à resposta imune contra L-asparaginase.

2.2 Objetivos Específicos

- Calcular o escore de imunogenicidade de epítomos de L-asparaginase de *Escherichia coli*;
- Encontrar mutações para a sequência de L-asparaginase de *Escherichia coli* que reduzam seu potencial imunogênico enquanto preservam a estabilidade;
- Selecionar, dentre os mutantes estruturalmente estáveis, aqueles que representem o melhor compromisso entre menor imunogenicidade, menor número de mutações.

3 METODOLOGIA

A estratégia de desimunização da sequência de L-asparaginase de *Escherichia coli* se iniciou com a escolha das regiões da proteína a serem alvos de mutagênese (Seção 3.1.1). O método TEPITOPE foi utilizado para a obtenção dos escores de imunogenicidade, resultando em um valor correspondente a cada complexo epítipo-alelo de MHC II analisado (Seção 3.1.2).

Um Algoritmo Genético foi utilizado para investigar as variantes ideais, por meio da (i) realização de mutações, (ii) recombinação de mutações, e (iii) seleção dos indivíduos de menor imunogenicidade (Seção 3.2).

Finalmente, os mutantes obtidos pelo Algoritmo Genético foram avaliados frente a uma fronteira de Pareto para a seleção dos mutantes que representam o melhor compromisso entre a menor imunogenicidade e o menor número de mutações (Seção 3.3).

3.1 Determinação da Imunogenicidade de Epítomos Lineares

3.1.1 Identificação de epítomos de L-asparaginase de *Escherichia coli* e seleção de alelos de HLA-II

O primeiro estágio do trabalho consistiu na identificação dos epítomos lineares de linfócitos T CD4 na sequência de L-asparaginase de *Escherichia coli*.

Um estudo experimental realizado por Cantor *et al.* (2011) identificou dois epítomos de L-asparaginase de *E. coli*, sendo o primeiro epítipo de 10 aminoácidos, começando no resíduo 111 da sequência, e o outro com nove aminoácidos de comprimento, iniciando no resíduo de número 217 (CANTOR *et al.*, 2011).

O epítipo com 10 resíduos de comprimento foi separado em dois epítomos sobrepostos (111-119 e 112-120) para que tivessem seu escore de imunogenicidade calculado pelo TEPITOPE (Seção 3.1.2), visto que o mesmo analisa apenas peptídeos com nove resíduos de comprimento (nonâmeros), tamanho correspondente à fenda de ligação do MHC (SANCHEZ-TRINCADO *et al.*, 2017). Isso resultou em três epítomos com nove resíduos de comprimento, cada (Tabela 5).

Os epítomos identificados foram investigados com relação à sua propensão em se ligar a cada uma das proteínas de HLAs definidas neste trabalho. Para isso, foi realizada uma busca na literatura de alelos de HLA considerando apenas aqueles (i) relacionados a hipersensibilidade

Tabela 5 – Epítomos de L-asparaginase de *Escherichia coli* determinados experimentalmente que foram alvos de mutagênese

Epítopo	Sequência	Posição
M ¹¹¹	MVGAMPST	111-119
V ¹¹²	VGAMPSTS	112-120
V ²¹⁷	VYNYANASD	217-225

Fonte: Elaborado pela autora.

Tabela 6 – Alelos de HLA utilizados na predição da imunogenicidade de epítomos de L-asparaginase

Alelo	Relevância	Referência
DRB1*0701	Hipersensibilidade	(FERNANDEZ <i>et al.</i> , 2014)
DRB1*0401	LLA	(DORAK <i>et al.</i> , 1999)
DRB1*0101 DRB1*1302 DRB1*1501 DRB5*0101	Representam mais de 80% da população mundial	(SOUTHWOOD <i>et al.</i> , 1998)

Fonte: Elaborado pela autora.

à L-asparaginase (FERNANDEZ *et al.*, 2014), (ii) ligados à LLA (DORAK *et al.*, 1999), ou que (iii) juntos representam mais de 80% da população mundial, independentemente da origem étnica (SOUTHWOOD *et al.*, 1998) (Tabela 6).

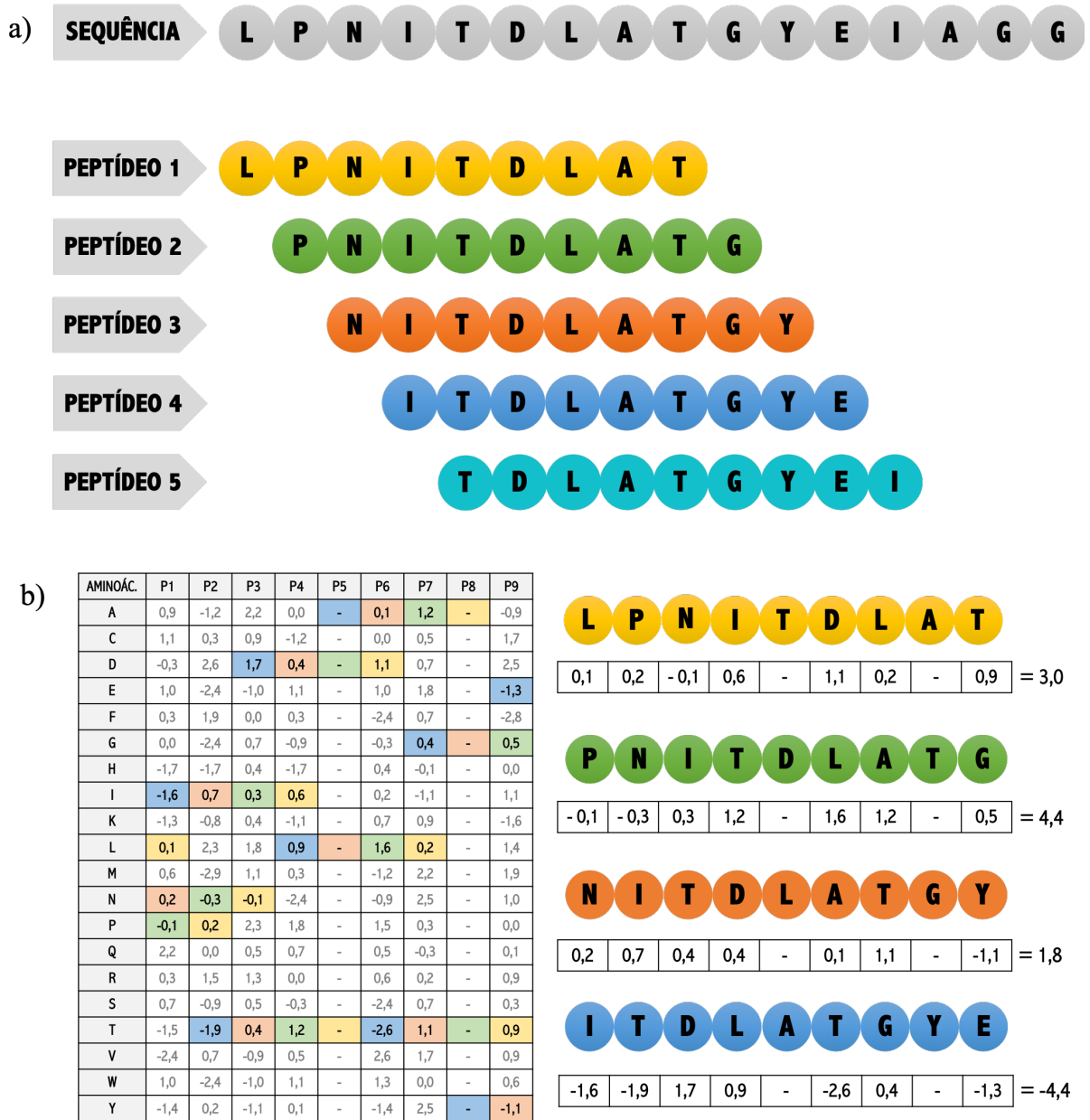
3.1.2 Cálculo da imunogenicidade dos epítomos

A determinação da imunogenicidade dos epítomos foi realizada pelo método TEPITOPE (STURNIOLO *et al.*, 1999). O arquivo no formato FASTA contendo a sequência de L-asparaginase de *Escherichia coli* obtida no *Protein Data Bank* (RCSB PDB, 1993) foi submetido ao TEPITOPE para a obtenção dos escores.

O TEPITOPE simulou a clivagem da proteína considerando todos os nonâmeros possíveis, iniciando do primeiro aminoácido ao nono, do segundo ao décimo e assim por diante, abrangendo todos os peptídeos nonaméricos sobrepostos ao longo da sequência de L-asparaginase (Figura 7 a).

O valor de imunogenicidade para cada nonâmero foi obtido pela soma da contribuição individual dos aminoácidos em diferentes posições das matrizes de HLA (Figura 7 b), que foram definidos experimentalmente através das variações da afinidade peptídeo-MHC (STURNIOLO

Figura 7 – Representação do cálculo dos escores de imunogenicidade de epítomos de L-asparaginase de *Escherichia coli* através do TEPITOPE



a) Os cinco primeiros peptídeos formados a partir da sequência original são mostrados na figura. O TEPITOPE simula a clivagem de não números das posições 1 a 9, 2 a 10, 3 a 11, e assim por diante, até que toda a sequência seja abrangida. b) Cada não número obtido pela clivagem sucessiva da sequência original é então submetido à análise da sua afinidade à molécula de MHC, utilizando-se as matrizes quantitativas para tal cálculo. Para cada posição da fenda de MHC (P1 a P9), existe um valor de afinidade correspondente ao aminoácido que a ocupa, com exceção das posições 5 e 8. Os valores de afinidade de cada posição são somados com os valores correspondentes aos demais resíduos de um mesmo peptídeo, e o resultado dessa operação indica a propensão de que aquele peptídeo específico se ligue ao MHC. O cálculo dos escores de afinidade é realizado para todos os peptídeos obtidos através da sequência proteica submetida para análise.

Fonte: Elaborado pela autora.

et al., 1999). As matrizes de afinidade não apresentam valores de contribuição para as posições 5 e 8 porque as cadeias laterais dos resíduos que poderiam interagir com os aminoácidos nestas posições estão voltadas para a direção oposta à fenda de ligação do MHC II, não influenciando na afinidade final da ligação.

O arquivo de saída apresentou os escores de imunogenicidade para cada nonâmero em ordem decrescente, de forma que os maiores escores representaram maior afinidade pelo MHC, indicando maior propensão do peptídeo ser um epítopo e, portanto, maior probabilidade de desencadear uma resposta imune. Epítomos promíscuos, isto é, presentes em mais de uma entrada, indicaram que o nonâmero foi identificado como epítopo por mais de um alelo de MHC II.

Os resultados gerados apresentaram todos os possíveis epítomos da sequência de L-asparaginase associados aos escores de imunogenicidade calculados pelo TEPITOPE, mas apenas os epítomos identificados previamente por Cantor *et al.* (2011) foram considerados na etapa seguinte para desimunização.

3.2 Desimunização da L-asparaginase de *Escherichia coli*

O Algoritmo Genético se iniciou com um conjunto de soluções candidatas ao problema, em que cada solução foi denominada indivíduo, e o conjunto de indivíduos, população inicial (Seção 3.2.1). Uma nova solução era formada através da aplicação de operadores genéticos de mutação (Seção 3.2.2), recombinação (Seção 3.2.3) e de um mecanismo de seleção que dava preferência aos indivíduos mais adaptados ao ambiente (Seção 3.2.3.1). Caracterizava-se assim um processo evolutivo, onde a aptidão de um indivíduo a cada geração apresentava uma tendência de crescimento continuado (Seção 3.2.5). Nas seções seguintes foram descritos os conceitos básicos de AG e a forma como foram implementados neste trabalho.

3.2.1 Indivíduo e População

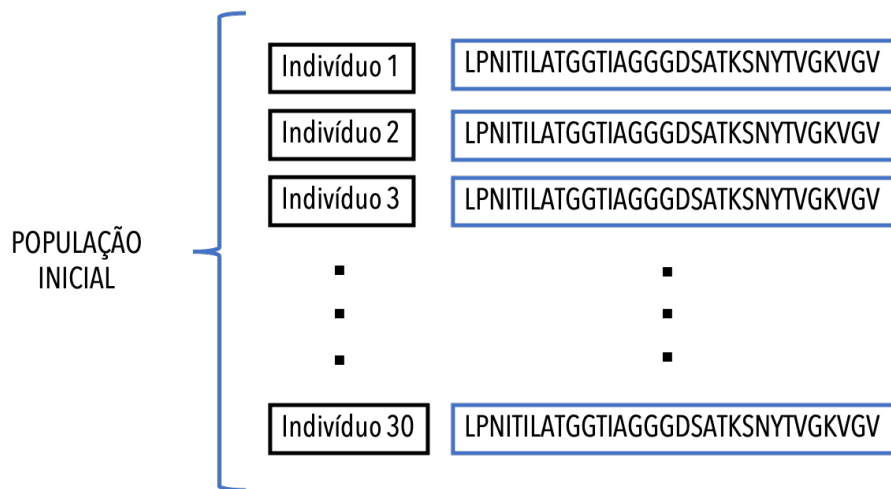
Um indivíduo (também denominado cromossomo ou genótipo) é o conjunto de parâmetros que representa uma possível solução para o problema. Neste trabalho, cada indivíduo representou uma cópia da sequência da proteína a ser desimunizada.

O conjunto de indivíduos é denominado população, que representa a totalidade de soluções candidatas. O tamanho máximo da população foi mantido, mas a composição da

população mudou ao longo das gerações conforme indivíduos mais adaptados foram criados.

O conjunto de sequências que foram usadas para inicializar o algoritmo consistiu em 30 cópias iguais da sequência de L-asparaginase de *Escherichia coli* que compuseram a denominada **população inicial** (Figura 8). Durante a execução do AG, mutações, recombinações e substituições foram feitas de modo a criar variabilidade na população visando obter indivíduos estáveis e menos imunogênicos, que, ao fim do processo evolutivo convergiram na **população final**.

Figura 8 – Representação dos indivíduos e da população inicial no Algoritmo Genético



População de tamanho 30, em que cada indivíduo corresponde a uma sequência de aminoácidos. Na população inicial, todos os indivíduos são iguais e, conforme o Algoritmo Genético insere os operadores genéticos a cada geração, a variabilidade é introduzida na população.

Fonte: Elaborado pela autora.

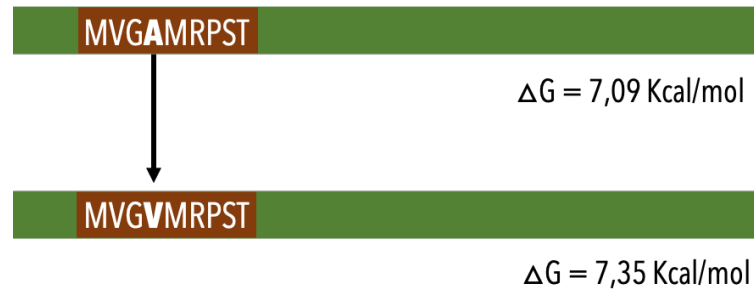
3.2.2 Operador de Mutação

Mutação é a forma de criar e manter diversidade na população de forma a experimentar novas soluções e investigar o espaço de busca de maneira não explorada. O operador de mutação utilizado neste AG modificava um único indivíduo, alterando sua sequência através da modificação aleatória de um aminoácido (Figura 9).

A probabilidade de ocorrência de mutação em um indivíduo, denominada taxa de mutação, foi de 0,016, a fim de criar variabilidade extra na população sem prejudicar o possível progresso obtido em gerações anteriores.

As mutações eram feitas sob as condições de que a posição da mutação estivesse na

Figura 9 – Representação do operador de mutação utilizado pelo Algoritmo Genético para geração de novas sequências desimunizadas de L-asparaginase de *Escherichia coli*



A mutação ocorre aleatoriamente em qualquer posição dos epítomos-alvo para outro aminoácido. Porém, apenas mutações que preservem a estabilidade da proteína serão mantidas como possíveis candidatos para compor a próxima geração.

Fonte: Elaborado pela autora.

lista de posições imunogênicas (Seção 3.1.1, Tabela 5) e que o aminoácido inserido nessa posição não prejudicasse a estabilidade da proteína nem gerasse novos epítomos. Dentre as posições que compõem os epítomos da L-asparaginase, uma era escolhida aleatoriamente. Todos os resíduos, com exceção da Prolina, eram testados para substituição na posição escolhida e, dentre os que não gerassem novos epítomos e não prejudicassem a estabilidade da proteína (Seção 3.2.4), um era escolhido aleatoriamente para ser o substituto.

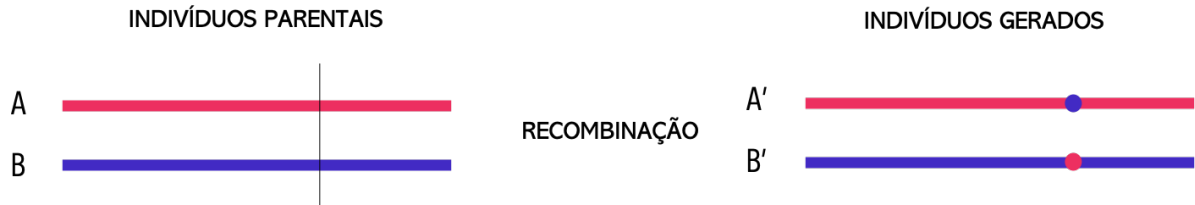
Devido ao fato da Prolina possuir uma cadeia lateral fechada que a impede de realizar ligações de hidrogênio com outro aminoácido, Prolinas já existentes não foram mutadas, assim como nenhum resíduo foi mutado para Prolina.

3.2.3 Operador de Recombinação

O operador de recombinação cria novos indivíduos através da combinação de resíduos de indivíduos parentais, permitindo a troca de informação entre soluções candidatas diferentes.

O operador de recombinação empregado foi o de recombinação de um ponto, em que apenas um resíduo da sequência era trocado. Para a aplicação deste operador, eram selecionados dois indivíduos-pais (Seção 3.2.3.1) e a partir de suas sequências eram gerados dois novos indivíduos-filhos. Para gerar os filhos, o algoritmo selecionava um ponto de corte aleatoriamente nas sequências dos pais, e os aminoácidos presentes nesta posição eram trocados (Figura 10). A taxa de recombinação, que indica a probabilidade de recombinação nos indivíduos, foi de 0,032.

Figura 10 – Representação do operador de recombinação utilizado pelo Algoritmo Genético para geração de novas sequências desimunizadas de L-asparaginase de *Escherichia coli*



O operador de recombinação começa com a seleção de dois indivíduos A e B da população, denominados parentais. Uma posição da sequência é escolhida aleatoriamente e os resíduos correspondentes a essa posição são trocados resultando em dois novos indivíduos, chamados filhos.

Fonte: Elaborado pela autora.

3.2.3.1 Seleção Parental

Para a obtenção de filhos mais adaptados, era favorável que os pais também possuísem melhores valores de aptidão. De forma a favorecer a seleção de indivíduos com menor imunogenicidade para serem selecionados pelo operador de recombinação, utilizou-se um método denominado **roleta**. Dessa forma, os indivíduos mais adaptados apresentavam maiores probabilidades de serem selecionados como parentais. Portanto, quanto menor a imunogenicidade de um mutante, maiores as suas chances de ser selecionado para ser submetido ao operador de recombinação (Figura 11).

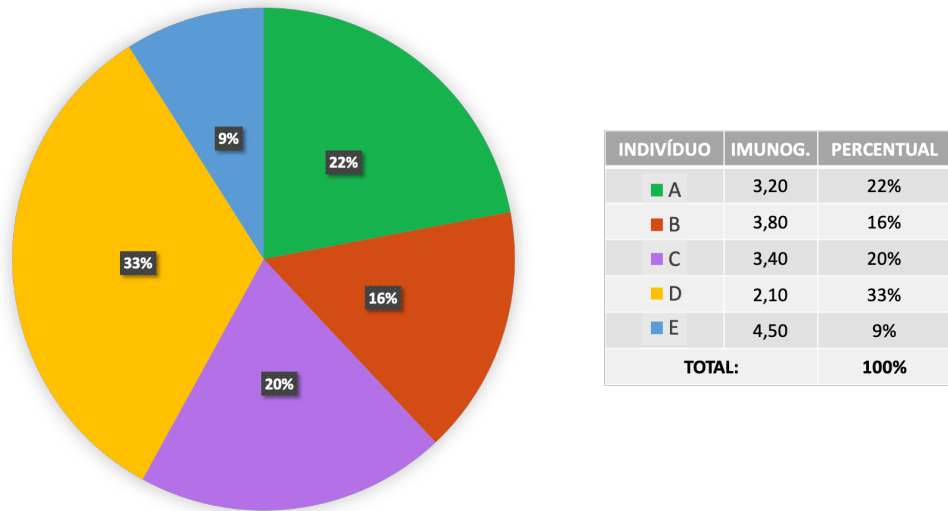
3.2.4 Funções de Restrição

3.2.4.1 Estabilidade Estrutural

Cada aplicação dos operadores de mutação (Seção 3.2.2) e recombinação (Seção 3.2.3) alterava a sequência de resíduos que compõem a proteína, o que potencialmente comprometeria sua conformação original. Embora o intuito da aplicação do Algoritmo Genético fosse a obtenção de sequências de menor imunogenicidade, um mutante de baixa imunogenicidade não seria relevante se sua estabilidade termodinâmica fosse menor que a da proteína original. De forma a garantir a estabilidade termodinâmica de cada novo indivíduo gerado, foi introduzido um critério que limitou a busca a sequências cuja estabilidade estrutural fossem próximas às da proteína selvagem.

Para a verificação da estabilidade proteica foi utilizado o *software* FoldX (Apêndice

Figura 11 – Representação do método de seleção de indivíduos parentais para recombinação utilizado pelo Algoritmo Genético para geração de novas sequências desimunizadas de L-asparaginase de *Escherichia coli*



Representação da roleta (esquerda) indicando a probabilidade inversamente proporcional ao escore de imunogenicidade (direita) de que o indivíduo seja escolhido como um parental para recombinação.

Fonte: Elaborado pela autora.

A) para cálculo da energia livre do enovelamento ΔG a partir da estrutura da proteína selvagem. A cada alteração de um indivíduo, o ΔG de sua estrutura era calculado pelo FoldX e comparado com o ΔG da estrutura selvagem. Para avaliar o efeito da mutação sobre a estabilidade da proteína, utilizou-se a diferença entre os valores de energia livre do indivíduo mutante com o selvagem, conforme a Equação 3.1:

$$\Delta\Delta G = \Delta G_{mutante} - \Delta G_{selvagem} \quad (3.1)$$

A estrutura terciária de cada mutante gerado era avaliada em função de sua energia livre ΔG , calculada utilizando o campo de forças empírico FoldX, a fim de verificar a viabilidade das mutações propostas em relação à estabilidade proteica. Mutações que apresentassem $|\Delta\Delta G| \leq 0,5kcal/mol$ eram estáveis e, portanto, consideradas viáveis (CAPRIOTTI *et al.*, 2008).

3.2.4.2 Criação de Novos Epítomos

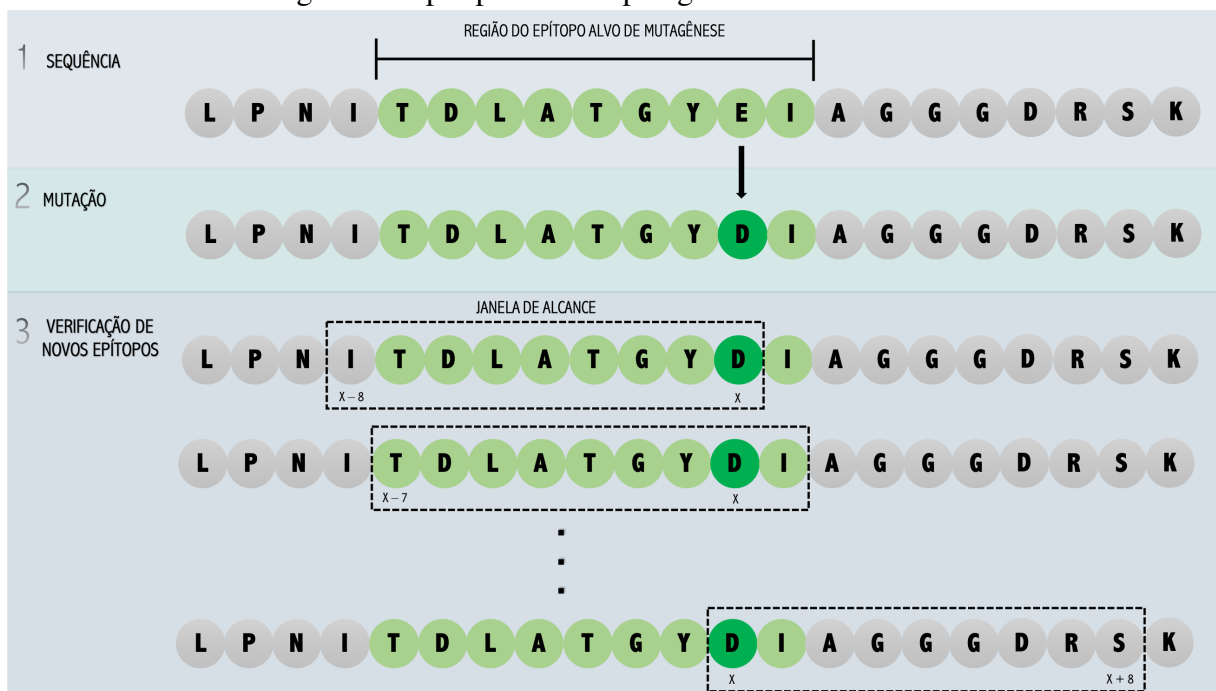
A cada operação de mutação ou recombinação realizada, o algoritmo verificava se havia redução da imunogenicidade do epítopo em questão. Entretanto, por tratar-se de uma sequência de nonâmeros sobrepostos contíguos, cada alteração que gerava redução do escore de

imunogenicidade em uma região poderia resultar em um efeito colateral de criação de epítomos em um trecho diferente da sequência.

Concretamente, para uma determinada mutação ou recombinação realizada na posição x , avaliava-se o efeito em toda a janela de peptídeos delimitada por $x - 8$ a $x + 8$ resíduos (Figura 12). A partir do ponto onde o novo resíduo havia sido inserido na sequência, o algoritmo percorria uma janela de posições de $x - 8$ a $x + 8$ a fim de verificar se a alteração proposta implicaria na geração de um novo epítomo.

Considerava-se que um epítomo novo havia sido criado se o escore TEPITOPE correspondente a qualquer um dos nonômeros entre $x - 8$ a $x + 8$ fosse maior que o original. Caso um novo epítomo tivesse sido criado, o indivíduo que apresentava esta alteração era descartado do conjunto de soluções possíveis.

Figura 12 – Função de restrição que verifica a criação de novos epítomos gerados pela alteração em uma região dos epítomos de L-asparaginase de *Escherichia coli*



Fonte: Elaborado pela autora.

3.2.5 Função de Aptidão

A cada indivíduo associava-se um valor denominado **função de aptidão** que indicava o quão adequado um indivíduo era para solucionar do problema. A função de aptidão foi, portanto, o critério que determinou a qualidade de um indivíduo frente aos demais da população, permitindo encontrar as melhores soluções.

Os valores de imunogenicidade obtidos através do TEPITOPE foram usados como Função de Aptidão. O escore de cada aminoácido do epítipo foi dado pelo valor de sua entrada na matriz de cada um dos HLAs escolhidos (Seção 3.1.2). O escore de cada epítipo foi, por sua vez, o valor da soma das contribuições individuais de cada aminoácido, enquanto a imunogenicidade total da proteína selvagem correspondeu à soma dos escores dos epítipos. Caso um aminoácido estivesse em regiões de sobreposições entre epítipos, os valores dos escores dos dois epítipos seriam somados.

Neste AG, uma solução era considerada apta se apresentasse imunogenicidade menor que a dos demais indivíduos da geração correspondente. Ao longo do tempo evolutivo, o valor de imunogenicidade era reduzido até chegar a uma solução ótima de imunogenicidade igual a zero.

3.2.6 *Substituição Parental*

A cada geração, x novos indivíduos eram gerados graças à aplicação dos operador de mutação (capaz de gerar um novo indivíduo) e recombinação (capaz de gerar até 2 novos indivíduos), e eram avaliados com relação aos seus respectivos valores de aptidão (Seção 3.2.5).

Para cada um dos novos indivíduos, o escore TEPITOPE foi avaliado frente aos valores dos demais presentes na população. Caso o valor do escore TEPITOPE do novo indivíduo i fosse menor do que o de algum indivíduo pré-existente p na população, o indivíduo p seria substituído pelo novo indivíduo i .

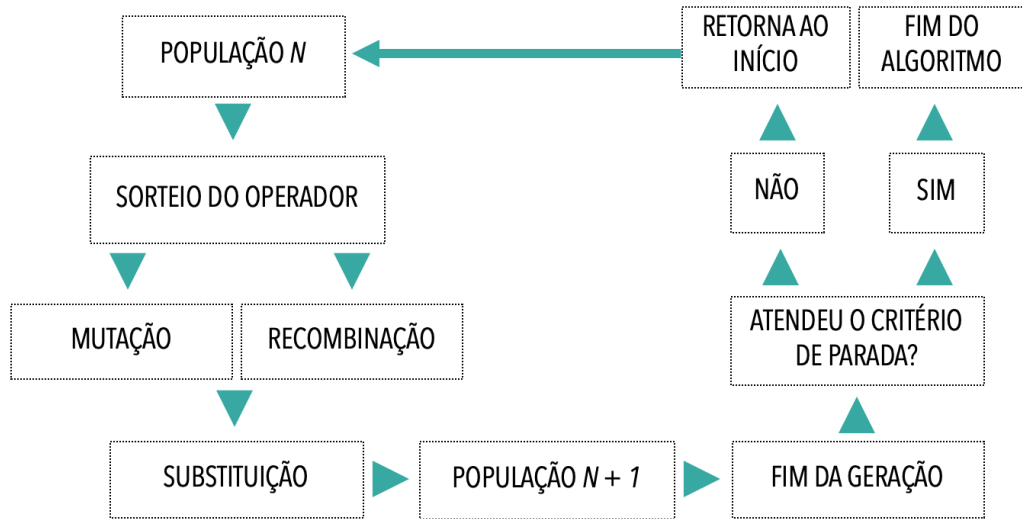
A substituição parental foi utilizada como critério de parada do algoritmo: caso uma geração se passasse sem que houvesse nenhuma mutação estável possível, ou seja, se nenhum indivíduo melhor fosse gerado, considerava-se que houve convergência da população.

3.2.7 *Fluxograma do Algoritmo Genético*

O Algoritmo Genético teve início com uma população inicial N de 30 indivíduos iguais que correspondiam à sequência de L-asparaginase de *E. coli*. Um dos operadores genéticos (mutação ou recombinação) era selecionado aleatoriamente e aplicado. Na mutação, apenas um indivíduo era sorteado e um aminoácido da região dos epítipos mutado. Na recombinação, dois indivíduos eram selecionados pelo AG via roleta e a escolha do ponto de corte a partir do qual as sequências seriam trocadas era aleatória.

Os novos indivíduos eram verificados pela função de restrição, que permitia que apenas indivíduos com $|\Delta\Delta G| \leq 0,5kcal/mol$ fossem mantidos para substituição de seus pa-

Figura 13 – Fluxograma de Operação do Algoritmo Genético para obtenção de sequências desimunizadas de L-asparaginase de *Escherichia coli*



Fonte: Elaborado pela autora.

Tabela 7 – Resumo dos parâmetros utilizados no Algoritmo Genético para obtenção de sequências desimunizadas de L-asparaginase de *Escherichia coli*

Parâmetro	Valor
Otimização	Redução da Imunogenicidade
Indivíduo	Sequência da L-asparaginase
Tamanho da População	$N = 30$
População Inicial	N cópias de indivíduos
Operador de Mutação	1 Ponto
Taxa de Mutação	0,016
Operador de Recombinação	1 Ponto
Taxa de Recombinação	0,032
Seleção Parental	Roleta
Substituição Parental	Substituição do pior indivíduo
Critério de Parada	Geração sem mutações estáveis possíveis
Função de Aptidão	Escore TEPITOPE
Função de Restrição	$\Delta\Delta G \leq 0,5 kcal/mol$; Não-criação de novos epítomos

Fonte: Elaborado pela autora.

rentais na população $N + 1$. Indivíduos de imunogenicidade menor que a dos parentais eram selecionados para compor a geração seguinte, substituindo os piores indivíduos da população. Dessa forma, a cada nova geração, indivíduos menos aptos eram substituídos por aqueles com melhor aptidão. O algoritmo foi encerrado quando não houve mais nenhuma mutação estável possível (Figura 13). Para obtenção de um maior número de sequências desimunizadas, o algoritmo foi aplicado 30 vezes seguindo os mesmos parâmetros citados. Os indivíduos obtidos em cada aplicação representaram o conjunto total de soluções obtidas pelo AG seguindo os parâmetros aplicados, resumidos na Tabela 7.

3.3 Análise e Seleção dos Mutantes Obtidos

3.3.1 Fronteira de Pareto

A fronteira de Pareto foi utilizada como uma ferramenta tática que permitiu a completa compreensão situacional das vantagens e desvantagens entre a quantidade de mutações e os diferentes níveis de diminuição da imunogenicidade que estas acarretaram.

Os objetivos considerados para esta análise foram a maior redução da imunogenicidade com o menor número de mutações. Para que ambos objetivos fossem atendidos igualmente dentro do conjunto de soluções propostas pelo Algoritmo Genético, todas as sequências mutantes obtidas foram analisadas segundo a Pareto-dominância. As sequências eficientes foram as que apresentaram a menor imunogenicidade possível dentro do menor número de mutações, formando o conjunto Pareto-ótimo de mutantes gerados. Os indivíduos que representaram o conjunto Pareto-ótimo foram plotados em um gráfico indicando a Fronteira de Pareto das sequências obtidas pelo Algoritmo Genético.

3.3.2 Avaliação da Similaridade de Sequência

Para verificar as variações causadas nas sequências mutantes produzidas quando comparadas à proteína selvagem, foram usados os valores da matriz BLOSUM62 (Tabela 8), que revela uma semelhança estatisticamente significativa entre as proteínas analisadas (HENIKOFF; HENIKOFF, 1992). Informações adicionais sobre a matriz foram descritas no Apêndice B.

As proteínas mutantes obtidas pelo Algoritmo Genético foram avaliadas em sua similaridade de sequência com a L-asparaginase de *Escherichia coli* selvagem, de acordo com os valores da BLOSUM62 (Tabela 8) a fim de comparar o grau de identidade das sequências

Tabela 8 – Matriz de Pontuação BLOSUM62 utilizada para avaliar a diferença entre as sequências de L-asparaginase de *Escherichia coli* selvagem e as sequências mutantes propostas pelo Algoritmo Genético

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4	-1	-2	-2	0	-1	-1	0	-2	-1	-1	-1	-1	-2	-1	1	0	-3	-2	0
R	-1	5	0	-2	-3	1	0	-2	0	-3	-2	2	-1	-3	-2	-1	-1	-3	-2	-3
N	-2	0	6	1	-3	0	0	0	1	-3	-3	0	-2	-3	-2	1	0	-4	-2	-3
D	-2	-2	1	6	-3	0	2	-1	-1	-3	-4	-1	-3	-3	-1	0	-1	-4	-3	-3
C	0	-3	-3	-3	9	-3	-4	-3	-3	-1	-1	-3	-1	-2	-3	-1	-1	-2	-2	-1
Q	-1	1	0	0	-3	5	2	-2	0	-3	-2	1	0	-3	-1	0	-1	-2	-1	-2
E	-1	0	0	2	-4	2	5	-2	0	-3	-3	1	-2	-3	-1	0	-1	-3	-2	-2
G	0	-2	0	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-2	-3	-3
H	-2	0	1	-1	-3	0	0	-2	8	-3	-3	-1	-2	-1	-2	-1	-2	-2	2	-3
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4	2	-3	1	0	-3	-2	-1	-3	-1	3
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4	-2	2	0	-3	-2	-1	-2	-1	1
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5	-1	-3	-1	0	-1	-3	-2	-2
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5	0	-2	-1	-1	-1	-1	1
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6	-4	-2	-2	1	3	-1
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7	-1	-1	-4	-3	-2
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4	1	-3	-2	-2
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5	-2	-2	0
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11	2	-3
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	-1
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4

Fonte: Henikoff e Henikoff (1992).

mutantes obtidas com a proteína original. Os valores obtidos através da matriz de substituição associados às sequências mutantes produzidas foram plotados em um gráfico para analisar o decaimento da similaridade de sequência de acordo com a redução da imunogenicidade.

3.3.3 Avaliação da Variação Estrutural

Mutações pontuais podem alterar de maneira significativa a estrutura conformacional de uma proteína, o que pode resultar no comprometimento de sua função (SCHAEFER; ROST, 2012). A métrica mais comumente usada para avaliação da diferença estrutural entre proteínas é o *Root-Mean-Square Deviation* (RMSD) ou desvio quadrático médio das posições atômicas, que corresponde à medida da distância média entre os átomos da cadeia principal de proteínas sobrepostas.

O RMSD calcula a raiz quadrática média da distância entre os resíduos correspondentes após o alinhamento entre duas proteínas (PETROVA, 2009), conforme a Equação 3.2, em que a média é realizada sobre os n pares de átomos correspondentes e d_i é a distância entre os

dois átomos no i -ésimo par entre as estruturas a e b :

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n d(a_i, b_i)^2} \quad (3.2)$$

Valores de RMSD que variam entre de 0 a 1,2 Å correspondem à flexibilidade inerente à proteína, ao passo que valores de RMSD maiores que 4 Å representam proteínas com homologia distante (KUFAREVA; ABAGYAN, 2011). Portanto, ao comparar a estrutura de L-asparaginase de *Escherichia coli* selvagem com as estruturas mutantes produzidas pelo AG, valores de RMSD menores do que 1,2 Å foram utilizados como parâmetro para indicar mutações que não impactam a conformação estrutural, enquanto valores maiores que 1,2 Å indicam que as mutações alteraram a estrutura tridimensional localmente.

Para verificar se as mutações propostas pelo AG nos melhores mutantes levaram a mudanças conformacionais da proteína, o cálculo do RMSD foi realizado através do programa Pymol, utilizando-se a estrutura de PDB da L-asparaginase de *Escherichia coli* selvagem e os PDBs das estruturas mutantes.

4 RESULTADOS E DISCUSSÃO

A Seção 4.1 descreve os resultados dos escores de imunogenicidade calculados para cada um dos epítomos e analisa a contribuição individual dos aminoácidos associada às posições da fenda de ligação de MHC II para a imunogenicidade da proteína. A Seção 4.2 indica os resultados da desimunização, apresentando a população de mutantes obtida pelo Algoritmo Genético e o padrão de redução da imunogenicidade. A Seção 4.3 detalha os melhores mutantes escolhidos e analisa suas características estruturais.

4.1 Determinação da Imunogenicidade de Epítomos Lineares

Os escores de imunogenicidade dos epítomos M^{111} , V^{112} e V^{217} de L-asparaginase de *Escherichia coli* são mostrados na Tabela 9, assim como a imunogenicidade total da proteína selvagem, resultante da soma dos escores de seus epítomos.

Tabela 9 – Escores de imunogenicidade calculados para os epítomos de L-asparaginase de *Escherichia coli*

Epítomo	Posição	Sequência	Alelo	Escore
M^{111}	111-119	MVGAMRPST	HLA-DRB1*1501	3,4
			HLA-DRB1*1302	1,3
			HLA-DRB1*0701	0,4
V^{112}	112-120	VGAMRPSTS	HLA-DRB5*0101	1,9
			HLA-DRB1*1501	1,3
			HLA-DRB1*1302	1,0
			HLA-DRB1*0401	0,9
V^{217}	217-225	VYNYANASD	HLA-DRB1*1501	2,7
Imunogenicidade			Total	12,9

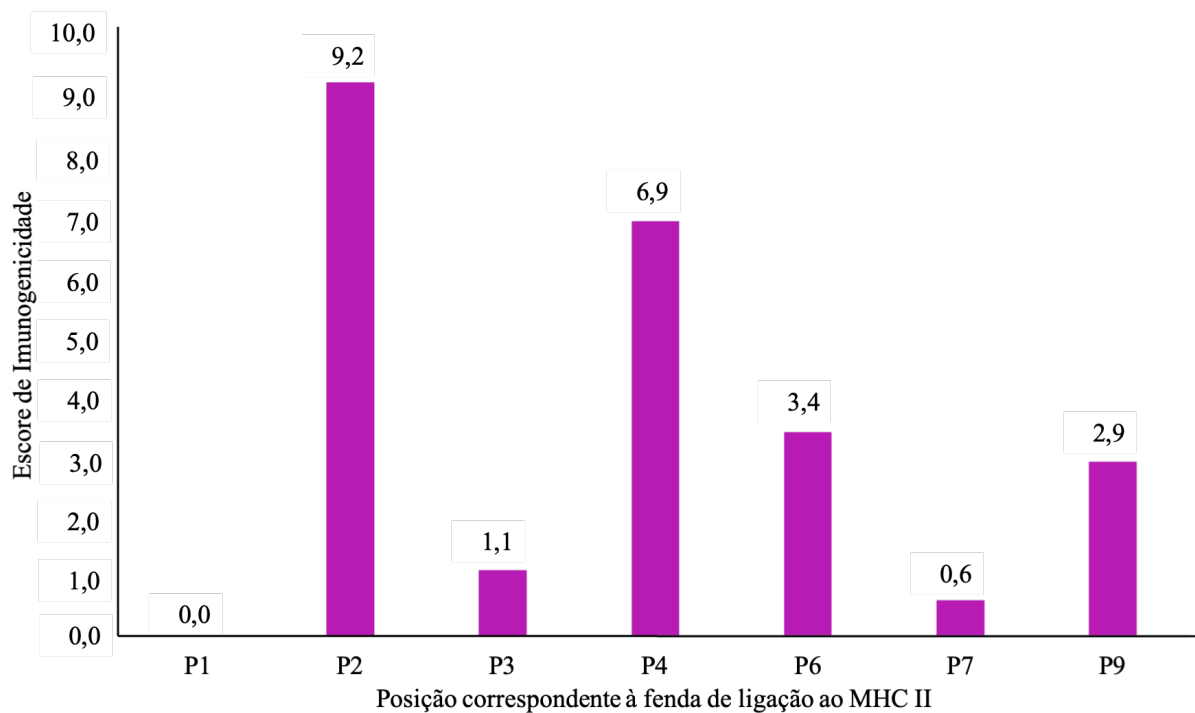
Fonte: Elaborado pela autora.

O epítomo M^{111} apresentou afinidade por três alelos, resultando em uma contribuição igual a 5,1 para o escore de imunogenicidade da proteína. O epítomo V^{112} , diferentemente, apresentou afinidade por quatro alelos e contribuiu com o mesmo escore de imunogenicidade do epítomo M^{111} . O epítomo V^{217} apresentou afinidade por apenas um alelo, com escore igual a 2,7, que junto com a contribuição dos demais epítomos, resultou em uma imunogenicidade total de

12,9 para a L-asparaginase de *E. coli* (Tabela 9). O valor de imunogenicidade calculado para a L-asparaginase correspondeu a um escore teórico, servindo somente como valor comparativo diante das estratégias de desimunização, que objetivaram a redução do escore imunogênico.

Foi feita uma análise dos aminoácidos que mais contribuíram para a imunogenicidade da proteína de acordo com a posição que ocuparam na fenda de ligação a MHC II, revelando que as posições que apresentaram os maiores escores foram P2, P4, P6 e P9 (Figura 14). Em estudos anteriores, as posições P4, P6 e P9 foram associadas às maiores contribuições para a afinidade à molécula de MHC II (JARDETZKY *et al.*, 1990; O'SULLIVAN *et al.*, 1991; HAMMER *et al.*, 1994; HILL *et al.*, 1994; DAVENPORT *et al.*, 1995), sendo chamadas de âncoras secundárias, que ocupam bolsas rasas na molécula de MHC II (MADDEN, 1995; FREMONT *et al.*, 1996). São consideradas secundárias porque qualquer uma delas pode ser substituída sem necessariamente haver perda completa da afinidade de ligação (CHELVANAYAGAM, 1997). Os aminoácidos que se ligam a posições de ancoragem secundárias contribuem de maneira mais significativa para a afinidade de ligação dos peptídeos, modulando a intensidade da afinidade.

Figura 14 – Contribuições para a afinidade de ligação de epítomos de L-asparaginase de *E. coli* associadas à cada posição da fenda de MHC II



Fonte: Elaborado pela autora.

As posições de ancoragem podem refletir dois efeitos principais distintos relacionados a (i) aminoácidos que refletem restrições estéricas ou (ii) contribuição dos aminoácidos para

a afinidade de ligação do peptídeo (HAMMER *et al.*, 1994). O aminoácido que se localiza na posição P1 do peptídeo reflete restrições estéricas, o que torna a P1 uma posição de importância primária para a avaliação da ligação entre o MHC e o epítopo (CHELVANAYAGAM, 1997). Todas as sequências que se ligam a HLA-DR apresentam um resíduo aromático-hidrofóbico na posição P1, cuja troca por aminoácido de outra natureza leva a perda completa da afinidade de ligação do peptídeo a qualquer um dos HLA-DR (Tabelas 10, 11, 12, 13, 14, 15). A importância da posição P1 foi apontada em diversos estudos (JARDETZKY *et al.*, 1990; HILL *et al.*, 1991; O'SULLIVAN *et al.*, 1991) e indica que a âncora P1 se encaixa no fenda hidrofóbica identificada do sítio de ligação DR (BROWN *et al.*, 1993). Apesar de sua contribuição primária, a presença de aminoácidos aromáticos ou hidrofóbicos em P1 é necessária, mas não suficiente para garantir a interação com o MHC (FREMONT *et al.*, 1996).

Os aminoácidos correspondentes à âncora P1 permitiram que os epítopos M^{111} , V^{112} e V^{217} se ligassem à fenda de MHC II porque não refletiam restrições estéricas. No entanto, esses mesmos resíduos não apresentaram contribuição alguma para a imunogenicidade dos epítopos (Figura 14).

É interessante observar que a posição P2, que apresentou a maior contribuição para a imunogenicidade dos epítopos de L-asparaginase, não foi relatada como uma âncora secundária em estudos sobre o padrão de ligação de peptídeos a MHC II. Analisando a natureza dos aminoácidos que ocupam a posição P2 nos epítopos M^{111} , V^{112} e V^{217} , temos que todos se enquadram na categoria de resíduo aromático ou hidrofóbico, que é o requisito de ligação à P1. Pode-se associar essa contribuição maior para a afinidade de ligação ao efeito hidrofóbico causado pela sinergia das interações dos resíduos que ocupam P1 e P2, uma interação que gera grande estabilidade em proteínas (MOTIEJUNAS; WADE, 2007; SCHAEFFER, 2008).

Vários estudos independentes apontaram para um padrão de ligação comum a todos os HLA-DR denominado motivo P1-P6, que aponta o papel crucial para a ligação a alelos do tipo HLA-DR de um grande resíduo aromático ou hidrofóbico (F, W, Y, L, I, V, ou M) na posição P1, e de resíduos curtos e/ou hidrofóbicos (S, T, C, A, P, V, I, L ou M) na posição P6 (O'SULLIVAN *et al.*, 1991; HAMMER *et al.*, 1993; SETTE *et al.*, 1993; HAMMER *et al.*, 1994; HAMMER *et al.*, 1994; HILL *et al.*, 1994; MARSHALL *et al.*, 1995). Resíduos com cadeia lateral maior na posição P6 diminuem a afinidade de ligação, sugerindo restrições estéricas nesta região do sítio de ligação (HAMMER *et al.*, 1993).

Não há preferências listadas para as posições P5 e P8, confirmando com as posições

que geralmente se projetam para a direção oposta à fenda de ligação (STERN *et al.*, 1994; GHOSH *et al.*, 1995; FREMONT *et al.*, 1996). Além disto, a região correspondente a P8 não apresenta muitos polimorfismos, o que reforça o fato de que esse sítio interage principalmente com o TCR (GARBOCZI *et al.*, 1996). O alto valor de polimorfismo da posição P5 está associado à necessidade de permitir pequenas mudanças conformacionais da cadeia lateral e flexibilidade da cadeia principal nas posições âncora P4 e P6 (FREMONT *et al.*, 1996) e reconhecimento por TCR (REAY *et al.*, 1994; SCHILD *et al.*, 1995). Similarmente, a região da posição P7 também é bastante diversa e mutações nesta região também podem afetar o reconhecimento por TCR (HSU *et al.*, 1995; EVAVOLD *et al.*, 1993; SPAIN J L JORGENSEN; BERG, 1994).

Apesar de os alelos de HLA-DR apresentarem, em geral, um padrão de ligação a peptídeos semelhante, existem algumas características estruturais na fenda de ligação de cada um deles que lhes são específicas, resultando em preferências de ligação distintas.

O alelo **DRB1*1501** demonstrou alta afinidade por todos epítomos de L-asparaginase de *E. coli*, sendo o responsável pelo maior escore de imunogenicidade dentre todos os complexos epítomo-HLA. Além disso, o alelo DRB1*1501 foi o único que apresentou afinidade pelo epítomo V²¹⁷.

Os resultados mostraram que apenas 1 ou 2 resíduos de cada epítomo apresentaram contribuições negativas para a afinidade de ligação ao alelo DRB1*1501 (Tabela 10). Do total de 21 resíduos relevantes para a ligação peptídeo-MHC, apenas 4 representam valores negativos para a afinidade, sendo que 3 se encontram na posição de ancoragem secundária P9 e 1 na posição P6 (Tabela 10). Isso está de acordo com o padrão de ligação reportado para o alelo HLA-DRB1*1501, que revela grande promiscuidade nas posições de ancoragem P1, P4, P6 e P9 onde múltiplos aminoácidos foram observados com valores similares de afinidade (NIELSEN *et al.*, 2010).

O alelo **DRB5*0101** apresentou afinidade alta e específica pelo epítomo V¹¹², representando o terceiro maior escore de imunogenicidade obtido dentre os demais epítomos (Tabela 9). A maior contribuição para a imunogenicidade do epítomo V¹¹² em ligação a DRB5*0101 correspondeu à Met em P4, acompanhada da Ser em P9 (Tabela 11). Os epítomos M¹¹¹ e V²¹⁷ apresentaram contribuições negativas ou nulas nas principais posições de ancoragem deste alelo, P1, P4, P6 e P9, e por isso não demonstraram afinidade pelo alelo DRB5*0101, ao contrário do epítomo V¹¹², que apresentou aminoácidos favoráveis à ligação ao MHC nessas posições, que foram reportados em estudos anteriores (SOUTHWOOD *et al.*, 1998; SCHOLZ *et al.*, 2017).

Tabela 10 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*1501

Posição na fenda de MHC II	M^{111}		V^{112}		V^{217}	
	AA	Cont. Ind.	AA	Cont. Ind.	AA	Cont. Ind.
P1	M	0,0	V	0,0	V	0,0
P2	V	2,1	G	0,5	Y	0,9
P3	G	0,2	A	0,0	N	0,5
P4	A	0,0	M	1,0	Y	2,5
P5	M	-	R	-	A	-
P6	R	1,0	P	-0,2	N	0,7
P7	P	0,3	S	0,3	A	0,0
P8	S	-	T	-	S	-
P9	T	-0,2	S	-0,3	D	-1,9
Contribuição Total		3,4		1,3		2,7

AA: aminoácido. Cont. Ind.: Contribuição Individual.

Fonte: Elaborado pela autora.

O padrão de ligação ao alelo DRB5*0101 sugere que a posição P1 se liga preferencialmente a resíduos aromáticos grandes, P4 a resíduos alifáticos menores e P6 a resíduos polares ou de cadeia lateral reduzida, ao passo que a posição P7 não revelou preferências, diferente de P8 e P9 que preferem resíduos positivamente carregados e aminoácidos básicos, respectivamente (SOUTHWOOD *et al.*, 1998; SCHOLZ *et al.*, 2017).

O alelo **DRB1*1302** demonstrou afinidade pelos epítomos M^{111} e V^{112} , representando a segunda maior contribuição por alelo para a imunogenicidade de L-asparaginase. A afinidade por este alelo se deve principalmente ao epítopo M^{111} , atribuindo as maiores contribuições à Val em P2 e à Arg em P6 (Tabela 12). A maior contribuição para a afinidade apresentada pelo epítopo V^{112} não pode ser atribuída a um aminoácido específico, mas a uma soma de pequenas contribuições individuais (Gly em P2, Met em P4 e Pro em P6) que resultaram em um escore positivo e próximo do escore de imunogenicidade do epítopo M^{111} , indicando afinidades semelhantes por DRB1*1302.

Motivos de ligação para DRB1*1302 foram relatados em três estudos diferentes (BOITEL *et al.*, 1995; DAVENPORT *et al.*, 1995; DAVENPORT *et al.*, 1996) apresentando resultados similares entre si, com exceção das preferências em P9. Assim como em todos os

Tabela 11 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB5*0101

Posição na fenda de MHC II	M^{111}		V^{112}		V^{217}	
	AA	Cont. Ind.	AA	Cont. Ind.	AA	Cont. Ind.
P1	M	-1,0	V	-1,0	V	-1,0
P2	V	2,1	G	0,5	Y	0,9
P3	G	0,2	A	0,0	N	0,5
P4	A	0,0	M	1,7	Y	-0,6
P5	M	-	R	-	A	-
P6	R	-1,3	P	0,2	N	-1,3
P7	P	-0,6	S	-0,2	A	0,0
P8	S	-	T	-	S	-
P9	T	-0,2	S	0,7	D	-1,5
Contribuição Total		-0,8	1,9		-3,0	

AA: aminoácido. Cont. Ind.: Contribuição Individual.

Fonte: Elaborado pela autora.

HLA-DR, peptídeos que se ligaram ao DRB1*1302 apresentaram aminoácidos aromáticos ou hidrofóbicos na posição P1, onde foi constatada interação com a fenda hidrofóbica da molécula de HLA-DR (BROWN *et al.*, 1993), aromáticos ou hidrofóbicos em P4, Tyr (DAVENPORT *et al.*, 1995), Gly ou Ala (BOITEL *et al.*, 1995) e aminoácidos polares ou positivamente carregados em P9 (DAVENPORT *et al.*, 1996).

O alelo **HLA-DRB1*0401** demonstrou afinidade apenas pelo epítopo V^{112} , que não apresentou nenhum aminoácido com contribuição significativa para a imunogenicidade (Tabela 13).

O alelo **DRB1*0401** se liga a peptídeos caracterizados por um resíduo de cadeia lateral aromática ou hidrofóbica na posição P1 e um resíduo de cadeia lateral pequena não-carregada na posição P6 (HAMMER *et al.*, 1993; O'SULLIVAN *et al.*, 1991). Os epítomos M^{111} , V^{112} e V^{217} possuem um aminoácido de cadeia lateral hidrofóbica (M e V) na posição P1 em comum. Contudo, somente o epítopo V^{112} , único que demonstrou afinidade pelo alelo 0401, apresenta também um aminoácido não-carregado na posição P6 (P) (Tabela 13).

Embora o padrão das posições P1 e P6 fosse necessário para a ligação entre o peptídeo e DRB1*0401, ele não foi suficiente, como relatado por Sette *et al.* (1993). A maioria

Tabela 12 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*1302

Posição na fenda de MHC II	M^{111}		V^{112}		V^{217}	
	AA	Cont. Ind.	AA	Cont. Ind.	AA	Cont. Ind.
P1	M	-1,0	V	-1,0	V	-1,0
P2	V	2,1	G	0,5	Y	0,9
P3	G	0,2	A	0,0	N	0,5
P4	A	0,0	M	0,8	Y	0,4
P5	M	-	R	-	A	-
P6	R	1,0	P	0,5	N	-0,6
P7	P	-0,5	S	-0,9	A	0,0
P8	S	-	T	-	S	-
P9	T	-0,5	S	1,1	D	-0,6
Contribuição Total		1,3		1,0		-0,4

AA: aminoácido. Cont. Ind.: Contribuição Individual.

Fonte: Elaborado pela autora.

dos efeitos secundários que contribuíram para a ligação de um peptídeo ao 0401 foi associada às âncoras secundárias P4, P7 e P9 (MADDEN, 1995).

É interessante observar que, apesar de M^{111} e V^{112} serem essencialmente parte de um só epítopo maior de 10 aa (111 a 120), eles não apresentaram o mesmo grau de afinidade pelos mesmos alelos. Apesar de haver sobreposição de 8 aa entre os epítomos citados, a soma das contribuições dos aminoácidos nas posições que ocupam foram diferentes, pois, a exemplo da Arg em P6, um mesmo aa em uma posição diferente da fenda de ligação resultou em um valor de imunogenicidade bastante distinto.

O alelo **DRB1*0701** demonstrou afinidade apenas pelo epítopo M^{111} , com o escore de imunogenicidade mais baixo dentre os observados na Tabela 9. Ao analisar a matriz de ligação dos epítomos M^{111} e V^{112} ao alelo DRB1*0701 (Tabela 14), podemos verificar que a Val na posição P2 foi a maior responsável pelo escore imunogênico atribuído ao epítopo M^{111} , e que as contribuições associadas à P7 não foram representativas.

Uma análise das sequências que se ligam a DRB1*0701 revelou que mais de 90% dos ligantes de alta afinidade apresentavam o motivo P1-P6 e que 99% dos ligantes de alelos DR também apresentavam este padrão (SOUTHWOOD *et al.*, 1998). As contribuições secundárias

Tabela 13 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*0401

Posição na fenda de MHC II	M^{111}		V^{112}		V^{217}	
	AA	Cont. Ind.	AA	Cont. Ind.	AA	Cont. Ind.
P1	M	-1,0	V	-1,0	V	-1,0
P2	V	2,1	G	0,5	Y	0,9
P3	G	0,2	A	0,0	N	0,5
P4	A	0,0	M	0,9	Y	-1,6
P5	M	-	R	-	A	-
P6	R	-2,4	P	0,0	N	1,3
P7	P	-0,7	S	-0,2	A	0,0
P8	S	-	T	-	S	-
P9	T	-1,2	S	0,7	D	-1,7
Contribuição Total		-3,0		0,9		-1,6

AA: aminoácido. Cont. Ind.: Contribuição Individual.

Fonte: Elaborado pela autora.

de ligação ao alelo DRB1*0701 também se concentraram nas posições P4, P7 e P9, sendo que a posição de maior contribuição para DRB1*0701 relatada foi a P7 (SOUTHWOOD *et al.*, 1998).

Nenhum dos epítomos considerados apresentou afinidade por **HLA-DRB1*0101**, mesmo com padrão de ligação semelhante aos demais alelos (Tabela 15). As contribuições individuais dos aminoácidos mostram valores negativos concentrados em P6 para M^{111} , em P1 para V^{112} e em P4, P6 e P9 para V^{217} , correspondendo às posições de ancoragem primária e secundárias.

Foi relatado um motivo de ligação P1-P6 bem evidente para o alelo DRB1*0101 (SOUTHWOOD *et al.*, 1998). Em relação às demais posições de ancoragem secundárias, P4 demonstrou maior afinidade por resíduos alifáticos de cadeia lateral média ou resíduos polares, P7 com preferência por aminoácidos carregados e P9 maior afinidade por resíduos polares de cadeia lateral reduzida ou alifáticos (JAMES *et al.*, 2010).

Diante desses resultados, os epítomos M^{111} e V^{112} juntos demonstraram afinidade por 5 alelos diferentes do total de 6 que foram usados neste estudo, sendo considerados, portanto, promíscuos. Mutações nestes epítomos seriam mais abrangentes do que em epítomos não-

Tabela 14 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*0701

Posição na fenda de MHC II	M^{111}		V^{112}		V^{217}	
	AA	Cont. Ind.	AA	Cont. Ind.	AA	Cont. Ind.
P1	M	-1,0	V	-1,0	V	-1,0
P2	V	2,1	G	0,5	Y	0,9
P3	G	0,2	A	0,0	N	0,5
P4	A	0,0	M	-0,4	Y	-0,9
P5	M	-	R	-	A	-
P6	R	-1,1	P	-0,5	N	-0,6
P7	P	-0,2	S	0,4	A	0,0
P8	S	-	T	-	S	-
P9	T	0,4	S	-0,3	D	-1,2
Contribuição Total		0,4		-1,3		-2,3

AA: aminoácido. Cont. Ind.: Contribuição Individual.

Fonte: Elaborado pela autora.

promíscuos, o que os tornariam mais vantajosos em abordagens de desimunização por mutação. Dessa forma, mutações pontuais na região das posições 111 a 120 da L-asparaginase de *E. coli* são promissoras para reduzir a imunogenicidade total da proteína de maneira mais significativa do que mutações na região das posições 217 a 225.

4.2 Desimunização da L-asparaginase de *Escherichia coli*

Após a obtenção do escore imunogênico da proteína selvagem, a sequência de L-asparaginase de *Escherichia coli* foi submetida à mutagênese através do algoritmo genético, produzindo 51.224 sequências mutantes sem que novos epítomos fossem criados. Dessas, um total de 953 variantes estáveis ($|\Delta\Delta G| \leq 0,5kcal/mol$) e menos imunogênicas foram selecionadas pelo AG, apresentando entre 1 a 13 mutações por indivíduo (Figura 15). O algoritmo genético reduziu o valor da imunogenicidade calculada pelo TEPITOPE por meio da troca de aminoácidos presentes em epítomos.

A Figura 15 *a* representa o número de indivíduos desimunizados e estáveis obtidos após a aplicação do AG. Conforme o esperado, não foram observados muitos indivíduos com

Tabela 15 – Contribuição dos aminoácidos para a afinidade de ligação dos epítomos M^{111} , V^{112} e V^{217} com o alelo HLA-DRB1*0101

Posição na fenda de MHC II	M^{111}		V^{112}		V^{217}	
	AA	Cont. Ind.	AA	Cont. Ind.	AA	Cont. Ind.
P1	M	-1,0	V	-1,0	V	-1,0
P2	V	2,1	G	0,5	Y	0,9
P3	G	0,2	A	0,0	N	0,5
P4	A	0,0	M	0,8	Y	-1,1
P5	M	-	R	-	A	-
P6	R	-1,8	P	-0,2	N	-1,1
P7	P	0,1	S	-0,2	A	0,0
P8	S	-	T	-	S	-
P9	T	-0,2	S	-0,3	D	-1,9
Contribuição Total		-0,6		-0,4		-3,7

AA: aminoácido. Cont. Ind.: Contribuição Individual.

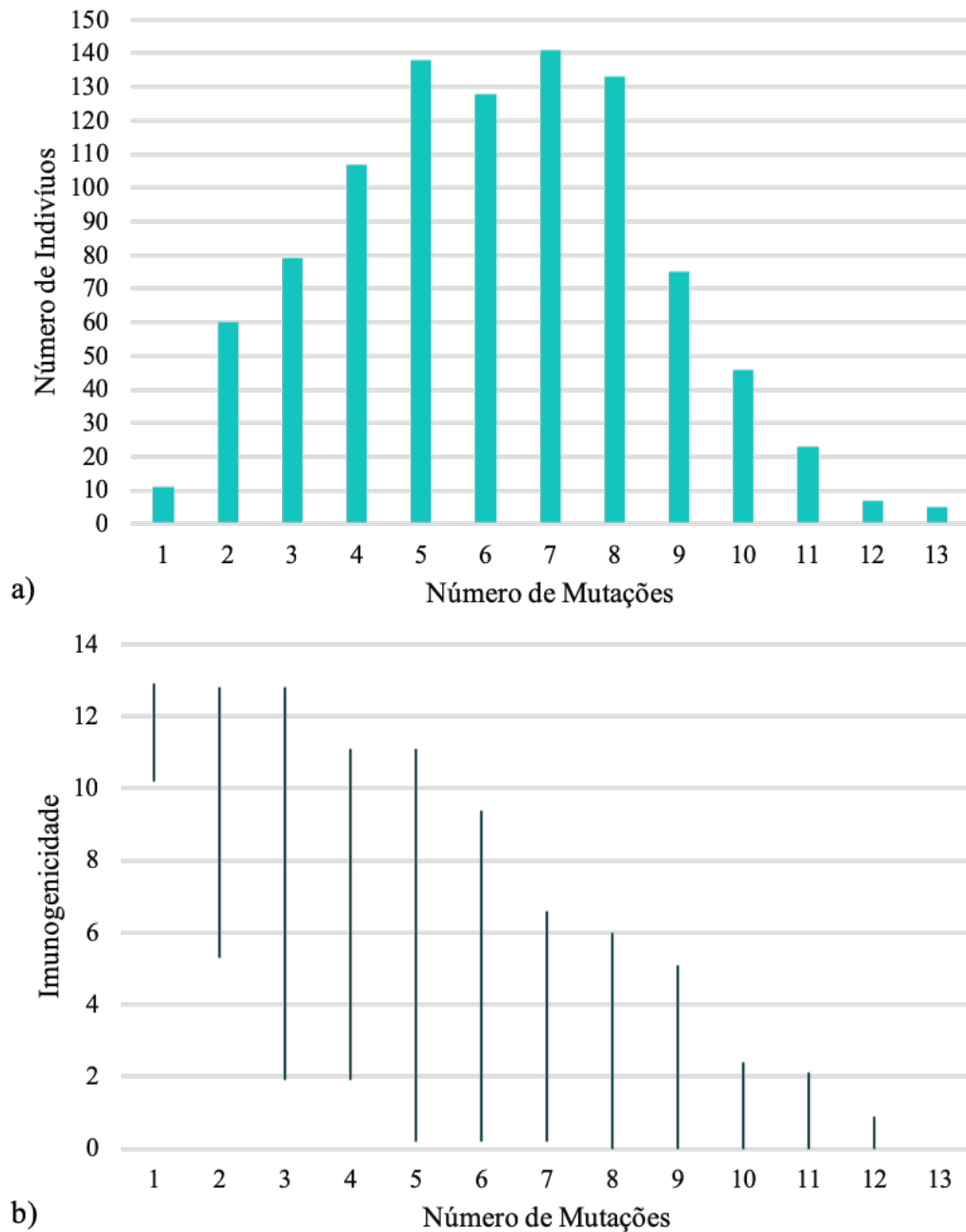
Fonte: Elaborado pela autora.

apenas 1 mutação na população final, porque a aplicação de somente 1 mutação permite a desimunização apenas do epítopo V^{217} ou, na melhor hipótese, dos epítomos sobrepostos M^{111} e V^{112} . Isso deixa a imunogenicidade da proteína alta em relação aos demais indivíduos (Figura 15 b), e conseqüentemente os elimina da população ao longo do tempo evolutivo.

Como se tratam de três epítomos da L-asparaginase de *E. coli*: M^{111} , V^{112} , V^{217} , com uma faixa de sobreposição de oito resíduos entre epítomos M^{111} e V^{112} o que indica que o mínimo de mutações esperada para se desimunizar a proteína seriam dois. Aumentar de uma para duas mutações permitiu um aumento em seis vezes no número de mutantes na população final (Figura 15 a) e a obtenção de indivíduos de imunogenicidade aproximadamente 2,5 vezes menor (Figura 15 b). Como há sobreposição de 8 aminoácidos entre os epítomos M^{111} e V^{112} , é provável que uma mutação nessa região seja capaz de reduzir a imunogenicidade de ambos. Ao aplicar três mutações, as possibilidades de indivíduos aumentaram em 1/3 e a imunogenicidade caiu em até 85%.

A faixa de mutações que concentrou o maior número de variantes foi de cinco a oito mutações (Figura 15 a), somando juntos 540 indivíduos, mais da metade da população obtida pelo AG. Em contrapartida, a diversidade de valores para a imunogenicidade foi mais extensa

Figura 15 – Indivíduos gerados pelo Algoritmo Genético a partir de mutações nos epítomos M^{111} , V^{112} e V^{217} de L-asparaginase de *E. coli*



a) Número de indivíduos produzidos pelo Algoritmo Genético para cada número de mutações realizadas na sequência. b) Distribuição dos escores de imunogenicidade obtidos para cada número de mutações representada pela linha vertical. Reduções significativas de imunogenicidade podem ser observadas em indivíduos a partir de apenas 2 mutações. Indivíduos com imunogenicidade igual a zero foram primeiramente obtidos com 8 mutações. A partir de 13 mutações, todos os indivíduos apresentaram imunogenicidade igual a zero.

Fonte: Elaborado pela autora.

em indivíduos de três e cinco mutações, onde alguns indivíduos apresentaram imunogenicidade similar à nativa (12,8) enquanto outros, valores próximos de zero (Figura 15 b). Considerando os

três epítomos que foram alvos de mutagênese, foi observado que mesmo com um número baixo de mutações por região foi possível alcançar reduções satisfatórias da imunogenicidade.

Porém, a partir de oito mutações, a quantidade de indivíduos gerados pelo AG foi reduzida pela metade do valor original. Isso possivelmente ocorreu porque um aumento do número de mutações em um espaço próximo na estrutura resulta em alterações na estabilidade da proteína, reduzindo o número de possibilidades de variantes com imunogenicidade menor. Apesar disso, foi possível obter indivíduos com um número grande de mutações (acima de oito mutações por indivíduo), sendo que a imunogenicidade se manteve decrescendo continuamente. O aumento de nove para dez mutações permitiu a redução da imunogenicidade pela metade (de 5 para cerca de 2,5 nos piores casos).

As mutações inseridas têm o potencial de prejudicar a atividade catalítica da L-asparaginase, podendo resultar em epítomos conformacionais, ou até mesmo no comprometimento da estabilidade termodinâmica de forma não-contemplada perfeitamente pelo modelo. Por isso, a partir dos indivíduos mutantes produzidos pelo AG, foi elaborada a fronteira de Pareto, para escolher os mutantes que representam o melhor compromisso entre número de mutações e a imunogenicidade.

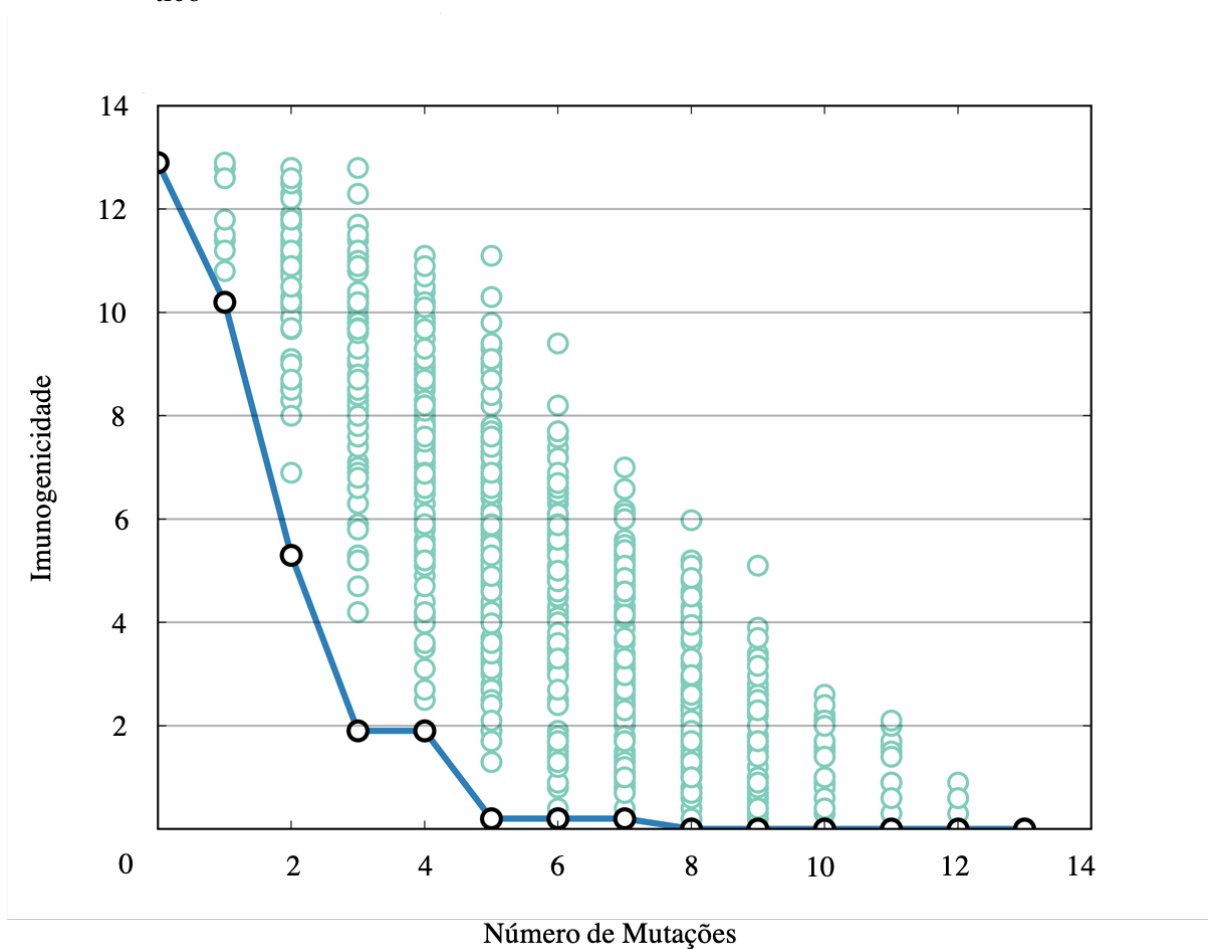
4.3 Análise e Seleção dos Mutantes Obtidos

Dentre as variantes propostas pelo algoritmo, os indivíduos da fronteira de Pareto representam, coletivamente, as soluções de melhor compromisso entre menor imunogenicidade menor número de mutações (Figura 16). Todos os pontos na fronteira de Pareto são qualitativamente equivalentes.

Os indivíduos que compõem a fronteira de Pareto apresentaram uma redução de 59% na imunogenicidade a partir de duas mutações, de 85% a partir de apenas três mutações e de 98,5% a partir de 5 mutações. O mutante com 4 mutações possui, aproximadamente, a mesma imunogenicidade que o indivíduo de 3 mutações, e, na prática não pode ser considerado melhor que o mutante de 3 mutações. Logo, os mutantes que apresentaram de 3 a 5 mutações foram selecionados como as opções mais relevantes.

O aumento do número de mutações na população resulta em melhores valores de imunogenicidade, pois mais mutações implicam em um distanciamento da similaridade de sequência em relação à proteína selvagem, permitindo a obtenção de peptídeos de menor imunogenicidade nas regiões dos epítomos. Entretanto, este distanciamento da similaridade da

Figura 16 – Fronteira de Pareto disposto todos os indivíduos produzidos pelo Algoritmo Genético



Cada círculo no gráfico corresponde a um indivíduo produzido pelo Algoritmo Genético. Os indivíduos estão distribuídos de acordo com o número de mutações que apresentam associado ao seu escore de imunogenicidade correspondente. A linha traçada em azul representa a Fronteira de Pareto, onde os indivíduos representam o melhor compromisso entre o menor escore imunogênico e menor número de mutações. Portanto, os indivíduos representados em círculos pretos são considerados as melhores soluções produzidas.

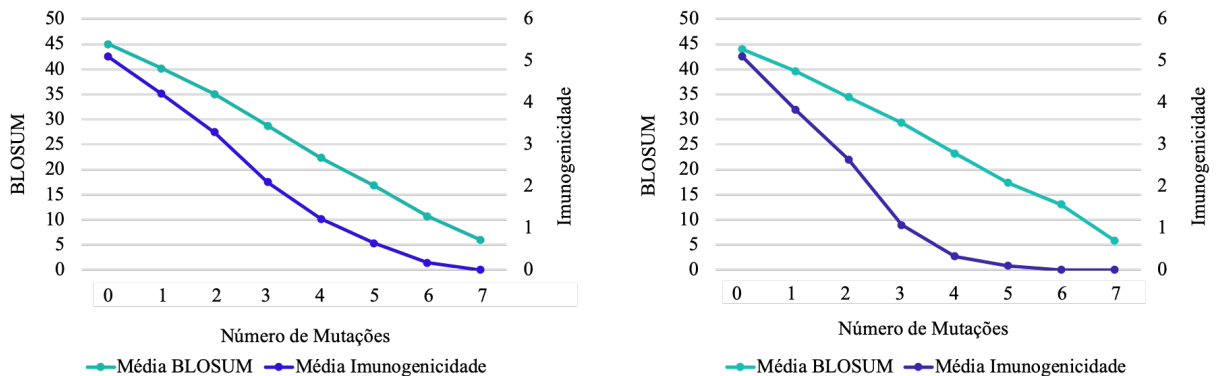
Fonte: Elaborado pela autora.

sequência original pode resultar em uma proteína comprometida do ponto de vista estrutural, funcional, farmacodinâmico ou farmacocinético.

Por isso, é interessante que as soluções apresentadas possuam a menor imunogenicidade possível com o menor número possível de mutações. Obter a menor imunogenicidade com o menor número de mutações configura interesses conflitantes porque quanto maior o número de mutações, maior o número de variantes possíveis. Isso é confirmado pela análise da população final, que revela que os indivíduos de menor imunogenicidade possuem um alto número de mutações (Figura 15).

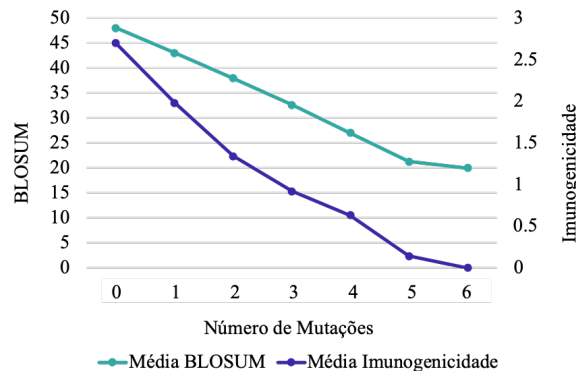
O aumento do número de mutações implicou no distanciamento da similaridade

Figura 17 – Decaimento da similaridade de sequência em comparação à redução da imunogenicidade de mutantes obtidos pelo AG a partir de mutações nos epítomos de L-asparinase de *E. coli*



a) Epítipo 111

b) Epítipo 112



c) Epítipo 217

Fonte: Elaborado pela autora.

em relação à sequência inicial, o que, evidentemente, foi responsável pela diminuição da imunogenicidade. Idealmente, as mutações inseridas diminuiriam a imunogenicidade mantendo a sequência o mais próximo o possível da original, o que pode ser obtido com a redução do número de mutações ou com a aplicação de mutações por aminoácidos similares. De forma a compreender o quanto a similaridade de sequência diminui conforme o aumento de mutações, os valores dados pela matriz de similaridade BLOSUM62 foram comparados com os valores do TEPITOPE para os indivíduos da população final (Figura 17).

Observa-se que com o aumento das mutações ocorre uma redução da imunogenicidade mais rápida do que a redução da imunogenicidade para V^{112} e V^{217} (Figura 17 b e c). Isso indica que as contribuições das mutações têm efeito maior na redução da imunogenicidade do que na redução da similaridade da sequência, e que o algoritmo foi capaz de reduzir o potencial imunogênico com menor impacto na similaridade de sequência, possivelmente em função da verificação do impacto na estabilidade (Seção 3.2.4).

Como os epítomos M^{111} e V^{112} apresentam parte de sua sequência compartilhada, mutações propostas na região de sobreposição beneficiaram um epítomo em detrimento de outro. Logo, a similaridade da sequência decaiu para os dois peptídeos mas isso não se refletiu na queda da imunogenicidade para o epítomo M^{111} (Figura 17 a).

Como a análise feita considera o comportamento dos epítomos separadamente, a diferença de resposta observada entre os epítomos pode ser atribuída ao fato de que as mutações feitas em posições relevantes do epítomo V^{112} coincidem com posições nulas do M^{111} . Adicionalmente, as posições 5 e 8 não influenciam a afinidade de ligação ao MHC (Seção 3.1.2). A mutação em P7 de V^{112} coincide com a mutação em P8 de M^{111} (analisado adiante, na Seção 4.3.1). Isto é, M^{111} é mutado em uma posição neutra e tem sua similaridade de sequência diminuída enquanto não se beneficia do efeito da redução da imunogenicidade.

4.3.1 Triplo-Mutante

O melhor triplo mutante obtido após a análise da fronteira de Pareto (Seção 4.3) possui alterações nos aminoácidos de posição G-113, S-118 e V-217 (Escore TEPITOPE = 1,9; $|\Delta\Delta G| = 0,45$ kcal/mol; RMSD = 0,018).

A posição G-113 corresponde à P3 do epítomo M^{111} e à P2 do epítomo V^{112} . Esta mutação permitiu que estes dois epítomos sobrepostos fossem desimunizados. Embora a G-113 possua baixa contribuição para a afinidade (0,2 em P3 de M^{111} e 0,5 em P2 de V^{112} , Tabelas 10 a 15), sua mutação para um aminoácido de valor de imunogenicidade negativo permitiu a redução do escore dos epítomos sobrepostos de 10,2 para 2,0. Após a mutação, o epítomo M^{111} demonstrou afinidade apenas pelo HLA-DRB1*1501, embora reduzida de 3,4 para 1,9, e o epítomo V^{112} apresentou afinidade apenas pelo HLA-DRB5*0101, também com redução de 1,9 para 0,1. Essa mutação representa uma solução bastante eficiente porque com a mudança de apenas um aminoácido foi possível reduzir o escore dos epítomos de maior imunogenicidade.

A mutação proposta no resíduo V-217, P1 do epítomo V^{217} , correspondeu à substituição de um resíduo essencial para a ligação a HLAs por um aminoácido não-hidrofóbico (Escore: -999), o que exclui completamente a possibilidade de que o epítomo se ligue à molécula de MHC (Seção 4.1). A mutação em V-217 ressalta a importância do aminoácido em P1, a âncora primária cuja mutação permite, segundo o critério do TEPITOPE, a desimunização completa do epítomo. P1 é um ótimo alvo para mutações, mas foi ignorada pelo algoritmo no caso dos epítomos M^{111} e V^{112} porque resultavam em indivíduos instáveis (Tabela 16).

Existe uma diferença no padrão de mutações do epítopo V^{217} em relação às dos epítopos M^{111} e V^{112} . É evidente que a mutação proposta em P1 para V^{217} não era possível para M^{111} e V^{112} porque elas são instáveis, mas este não é o único fator a ser considerado. V^{217} possui afinidade apenas pelo HLA-DRB1*1501, enquanto que M^{111} e V^{112} apresentam afinidade por diversos alelos (Tabela 9). Esta afinidade por um número maior de alelos impediu a realização de mutações em posições de alto valor de escore mas que atendessem apenas um HLA. Ao invés disso, foram priorizadas mutações em posições que contribuiriam com escores baixos para todos os HLAs e representassem o melhor compromisso para a desimunização da proteína inteira.

Outro fato interessante é de que as matrizes dos alelos considerados neste trabalho apresentam exatamente os mesmos valores de afinidade correspondentes às posições P2 e P3. Uma mutação que tenha valor negativo para a afinidade em uma dessas posições implica no mesmo valor negativo para a afinidade por todos os alelos. Mutações nessa região têm o potencial de diminuir ou extinguir a afinidade dos epítopos por uma grande diversidade de alelos através de uma mutação pontual estratégica, como foi o caso da mutação em G-113. Apesar da mutação em G-113 permitir a redução do escore de 10,2 para 2,0, nenhum indivíduo da população possui uma mutação nesta posição apenas. A análise do ΔG dos mutantes revela porque indivíduos com mutações apenas em G-113 não integram a população: qualquer mutação em 113 possui $|\Delta\Delta G| > 0,5$, e, portanto, não atende ao requisito de estabilidade exigido (Tabela 16). Por se tratar de uma Gly na posição 113, aminoácido de menor cadeia lateral, substituições nessa posição possivelmente resultam em colisões estéricas que desestabilizam a proteína.

A mutação na posição S-118 coincide com uma posição nula P8 no epítopo M^{111} , e, portanto, não contribui para a redução da imunogenicidade deste epítopo. Ademais, para o epítopo V^{112} a redução foi de apenas 0,1. Em conjunto, seria possível concluir que a mutação em S-118 possui baixa contribuição para a desimunização da proteína e hipotetizar que sua presença no indivíduo triplo-mutante se deve à natureza estocástica do algoritmo genético. Entretanto, foi observado que todos os indivíduos da população que apresentavam mutações no resíduo G-113 continham também a mutação em S-118.

Nos indivíduos que apresentaram a mutações adicional em S-118, o $\Delta\Delta G$ se encontrava dentro dos requisitos de estabilidade, o que justifica sua aplicação para que o mutante possa atender aos critérios de redução da imunogenicidade e manutenção da estabilidade. Conclui-se, portanto, que a mutação em S-118 é uma importante mutação "acessória" que possui baixa contribuição para a desimunização sozinha, mas permite a obtenção de variantes de baixa

Tabela 16 – $\Delta\Delta G$ das mutações propostas nos resíduos correspondentes às posições 111, 112 e 113 da sequência de L-asparaginase de *E. coli*

Mutação	$\Delta\Delta G$ (kcal/mol)		
	Posição 111	Posição 112	Posição 113
A	3,31	1,4	1,03
D	5,18	4,98	1,45
E	4,2	3,52	0,77
F	0,03	0,85	2,35
G	4,72	2,79	0,00
H	3,02	1,11	8,64
I	0,12	0,38	3,09
K	5,2	3,59	4,85
L	0,45	0,05	2,26
M	0,42	0,17	6,16
N	3,54	3,18	1,55
Q	2,05	2,81	9,58
R	5,52	0,73	9,7
S	4,47	1,71	3,98
T	2,06	1,66	2,99
V	0,16	0,69	2,09
W	1,91	3,59	4,84
Y	1,52	0,6	1,7

Fonte: Elaborado pela autora.

imunogenicidade em G-113.

4.3.2 *Quádruplo-Mutantes*

Os três melhores indivíduos com 4 mutações possuem imunogenicidades iguais (Escore TEPITOPE = 1,9) e foram mutados nas seguintes posições:

1. G-113, M-115, S-118 e V-217 ($|\Delta\Delta G| = 0,49$ kcal/mol; RMSD = 0,027)
2. G-113, S-118, V-217 e N-222 ($|\Delta\Delta G| = 0,20$ kcal/mol; RMSD = 0,045)
3. G-113, S-118, V-217 e S-224 ($|\Delta\Delta G| = 0,40$ kcal/mol; RMSD = 0,020)

Todos os indivíduos apresentaram mutações nas mesmas posições do indivíduo

triplo-mutante (G-113, S-118 e V-217, Seção 4.3.1), com uma mutação adicional em posições diferentes.

No indivíduo 1, a nova mutação foi realizada na posição 115, que é compartilhada pelos epítomos M^{111} e V^{112} . Para o epítopo M^{111} , o resíduo 115 corresponde à posição nula P5 e, portanto, sua mutação não afeta a imunogenicidade. Similarmente, a mutação em M-115 não afeta o escore do epítopo V^{112} porque as mutações em G-113 e S-118 são suficientes para reduzir o escore a zero (Seção 4.3.1).

Nos indivíduos representados em 2 e 3, as novas mutações também ocorreram em um epítopo que já estava totalmente desimunizado (escore TEPITOPE = 0) por única mutação em V-217. Apesar da mutação de Asn em P6 diminuir o escore imunogênico em 5 dos 6 alelos considerados, ela não foi relevante do ponto de vista da imunogenicidade, pois como mencionado anteriormente (Seção 4.1), a mutação em P1 deste epítopo já foi suficiente para que não houvesse nenhuma interação com os HLAs.

4.3.3 *Quíntuplo-Mutante*

O indivíduo de 5 mutações com melhor imunogenicidade (Escore TEPITOPE = 0,2; $|\Delta\Delta G| = 0,4$; RMSD = 0,051) apresenta mutações nos resíduos G-113, S-118, T-119, V-217 e N-222.

Essas são as mesmas mutações do indivíduo quádruplo-mutante n^o 2 (Seção 4.3.2) acrescido de uma mutação no resíduo 119, correspondente à P9 do epítopo M^{111} , que permitiu a redução do escore da proteína de 1,9 para 0,2. Como se trata da região de sobreposição entre os epítomos M^{111} e V^{112} , essa mutação pode ter efeitos negativos sobre V^{112} . Como a posição 119 coincide com a posição neutra P8 do epítopo V^{112} e V^{112} já se encontrava totalmente desimunizado (escore TEPITOPE = 0) pelas mutações em G-113 e S-118, a mutação em T-119 reduziu em 1,7 a imunogenicidade do epítopo M^{111} , mas não teve influência alguma sobre o epítopo V^{112} .

Na Seção 4.3.1 foi apresentado o triplo mutante com as seguintes mutações: G-113, S-118, V-217. Na Seção 4.3.2, foi apresentado o quádruplo mutante ao qual a mutação em 222 foi adicionada sem que houvesse, no entanto, redução na imunogenicidade. O quádruplo mutante aqui apresentado consiste neste mesmo quádruplo mutante acrescido de uma mutação em T-119, que permitiu a redução da imunogenicidade. Isso motiva a questionar porque o algoritmo reteve a mutação em N-222, que não reduz a imunogenicidade, e não sugeriu um quádruplo mutante

com mutações em G-113, S-118, T-119 e V-217.

O quádruplo mutante hipotetizado acima G-113, S-118, T-119 e V-217 foi construído e teve sua estabilidade calculada. O valor encontrado foi de $|\Delta\Delta G| = 0,88$ kcal/mol, *i.e.* fora do permitido, o que constitui uma estrutura conformacional com estabilidade comprometida. Em contrapartida, o quádruplo mutante apresenta $|\Delta\Delta G| = 0,4$ kcal/mol. Dessa forma, conclui-se que apesar da mutação em N-222 não apresentar efeito sobre a redução da imunogenicidade, ela se faz necessária para que a mutação em T-119 possa ser inserida e a proteína se mantenha estável.

Finalmente, têm-se que as mutações mais relevantes para a desimunização de L-asparaginase de *Escherichia coli* foram nos resíduos G-113, S-118, T-119 e V-217. As mutações em G-113 e S-118 reduziram o escore imunogênico do epítopo M^{111} de 5,1 para 1,9 e desimunizaram completamente o epítopo V^{112} . A mutação em V-217 também eliminou a afinidade do epítopo V^{217} por qualquer alelo. A mutação em T-119 contribuiu, junto com G-113 e S-118, para a desimunização do epítopo M^{111} , restando um escore de apenas 0,2, resultando em uma redução total da imunogenicidade de 12,7 com apenas 4 mutações. Apesar da mutação em S-118 ter contribuído para a redução da imunogenicidade (redução de 0,1), seu principal papel foi o de tornar a mutação em G-113 estável, resultando na mutação que mais desimunizou a L-asparaginase (redução de 8,2). Adicionalmente, a mutação em N-222 não demonstrou alteração alguma na imunogenicidade da proteína, porém permitiu que a mutação em T-119 (redução de 1,7) fosse estável.

5 CONCLUSÃO

Os escores de imunogenicidade obtidos para epítomos de L-asparaginase de *Escherichia coli*, utilizando-se alelos de HLA relacionados à resposta imune serviram como referência para a abordagem de desimunização da proteína por mutação *in silico*.

Com base nesses escores, o algoritmo genético utilizado foi eficiente na geração de mutantes desimunizados e estruturalmente estáveis, dentre os quais foram observados indivíduos contendo entre 3 a 5 mutações com reduções de 85,3% a 98,5% de imunogenicidade.

Em trabalhos futuros, as mutações propostas poderão ser realizadas em laboratório, para validação de sua imunogenicidade *in vitro* e posterior aplicação *in vivo*. A metodologia empregada pode ser utilizada para demais sequências proteicas, apenas adaptando os parâmetros computacionais.

REFERÊNCIAS

- ABBAS, A.; LICHTMAN, A. H.; PILLAI, S. In: **Imunologia Celular e Molecular**. [S.l.]: Elsevier, 2015. v. 8.
- AHMAD, T. A.; EWEIDA, A. E.; EL-SAYED, L. H. T-cell epitope mapping for the design of powerful vaccines. **Vaccine Reports**, Elsevier, p. 13–22, 2016.
- AHMED, R. K.; MAEURER, M. J. T-cell epitope mapping. In: **Methods in Molecular Biology**. [S.l.]: Humana Press, 2009. v. 524, p. 427–438.
- AMYLON, M. D.; SHUSTER, J.; PULLEN, J.; BERARD, C.; LINK, M. P.; WHARAM, M.; KATZ, J.; YU, A.; LAVER, J.; RAVINDRANATH, Y.; KURTZBERG, J.; DESAI, S.; CAMITTA, B.; MURPHY, S. B. Intensive high-dose asparaginase consolidation improves survival for pediatric patients with T cell acute lymphoblastic leukemia and advanced stage lymphoblastic lymphoma: a Pediatric Oncology Group study. **Leukemia**, Nature, v. 13, p. 335–342, 1999.
- BACKERT, L.; KOHLBACHER, O. Immunoinformatics and epitope prediction in the age of genomic medicine. **Genome Medicine**, BioMed Central, p. 1–12, 2015.
- BAKER, M. P.; REYNOLDS, H. M.; LUMICISI, B.; BRYSON, C. J. Immunogenicity of protein therapeutics: The key causes, consequences and challenges. **Self/Nonself**, Landes Bioscience, v. 1, p. 314–322, 2010.
- BHASIN, M.; RAGHAVA, G. P. S. SVM based method for predicting HLA-DRB1*0401 binding peptides in an antigen sequence. **Bioinformatics**, Oxford University Press, v. 20, p. 421–423, 2004.
- BOITEL, B.; BLANK, U.; MÈGE, D.; CORRADIN, G.; SIDNEY, J.; SETTE, A.; ACUTO, O. Strong similarities in antigen fine specificity among DRB1*1302-restricted tetanus toxin tt830-843-specific TCRs in spite of highly heterogeneous CDR3. **The Journal of Immunology**, American Association of Immunologists, v. 154, p. 3245–3255, 1995.
- BOUVIER, M.; WILEY, D. Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. **Science**, American Association for the Advancement of Science, v. 265, p. 398–402, 1994.
- BRENNAN, F. R.; MORTON, L. D.; SPINELDREHER, S.; KIESSLING, A.; ALLENSPACH, R.; HEY, A.; MULLER, P. Y.; FRINGS, W.; SIMS, J. Safety and immunotoxicity assessment of immunomodulatory monoclonal antibodies. **MAbs**, Taylor and Francis Group, v. 2, p. 233–255, 2010.
- BROOME, J. D. Evidence that the L-asparaginase of guinea pig serum is responsible for its antilymphoma effects. **Nature**, v. 191, p. 1114–1115, 1961.
- BROOME, J. D. Evidence that the L-asparaginase of guinea pig serum is responsible for its antilymphoma effects: Properties of the L-asparaginase of guinea pig serum in relation to those of the antilymphoma substance. **Journal of Experimental Medicine**, v. 118, n. 1, p. 99–120, 1963.

BROWN, J. H.; JARDETZKY, T. S.; GORGA, J. C.; STERN, L. J.; URBAN, R. G.; STROMINGER, J. L.; WILEY, D. C. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. **Nature**, v. 364, p. 33–39, 1993.

BRYSON, C. J.; JONES, T. D.; BAKER, M. P. Prediction of immunogenicity of therapeutic proteins. **BioDrugs**, Springer, v. 24, n. 1, p. 1–8, 2010.

BUI, H.-H.; SIDNEY, J.; PETERS, B.; SATHIAMURTHY, M.; SINICHI, A.; PURTON, K.-A.; MOTHÉ, B. R.; CHISARI, F. V.; WATKINS, D. I.; SETTE, A. Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. **Immunogenetics**, Springer, v. 57, p. 304–314, 2005.

CAMPBELL, H. A.; MASHBURN, L. T. L-asparaginase EC-2 from *Escherichia coli*: Some substrate specificity characteristics. **Biochemistry**, American Chemical Society, v. 8, p. 3768–3775, 1969.

CAMPBELL, H. A.; MASHBURN, L. T.; BOYSE, E. A.; OLD, L. J. Two L-asparaginases from *Escherichia coli* B. Their separation, purification, and antitumor activity. **Biochemistry**, American Chemical Society, v. 6, p. 721–730, 1967.

CANTOR, J. R.; YOO, T. H.; DIXIT, A.; IVERSON, B. L.; FORSTHUBER, T. G.; GEORGIU, G. Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 108, p. 1272–1277, 2011.

CAPRIOTTI, E.; FARISELLI, P.; ROSSI, I.; CASADIO, R. A three-state prediction of single point mutations on protein stability changes. **BMC Bioinformatics**, BioMed Central, v. 9, p. S6–S14, 2008.

CHELVANAYAGAM, G. A Roadmap for HLA-DR Peptide Binding Specificities. **Human Immunology**, Elsevier, v. 58, p. 61–69, 1997.

CLEMENTI, D. A. La désamidation enzymatique de L'asparagine chez les différentes espèces animales et la signification physiologique de sa présence dans l'organisme. **Archives Internationales de Physiologie**, Taylor and Francis Group, v. 19, p. 369–398, 1922.

CRISTIANO, M. V. de M. B. **Sensibilidade e Especificidade na Curva ROC: Um Caso de Estudo**. Dissertação (Mestrado) — Universidade do Porto, 2017.

D'AMARO, J.; HOUBIERS, J. G.; DRIJFHOUT, J. W.; M.P.BRANDT, R.; SCHIPPER, R.; BAVINCK, J. N. B.; MELIEF, C. J.; KAST, W. M. A computer program for predicting possible cytotoxic T lymphocyte epitopes based on HLA class I peptide-binding motifs. **Human Immunology**, Elsevier, v. 43, p. 13–18, 1995.

DAVENPORT, M. P.; QUINN, C. L.; CHICZ, R. M.; GREEN, B. N.; WILLIS, A. C.; LANE, W. S.; BELL, J. I.; HILL, A. V. S. Naturally processed peptides from two disease-resistance-associated HLA-DR13 alleles show related sequence motifs and the effects of the dimorphism at position 86 of the HLA-DRI3 chain. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 92, p. 6567–6571, 1995.

DAVENPORT, M. P.; QUINN, C. L.; VALSASNINI, P.; SINIGAGLIA, F.; HILL, A. V. S.; BELL, J. I. Analysis of peptide-binding motifs for two disease associated HLA-DR13 alleles using an M13 phage display library. **Immunology**, Wiley, v. 88, p. 482–486, 1996.

DAY, A. **The Source of Stability in Proteins**. 1996. Disponível em: <<http://www.cryst.bbk.ac.uk/PPS2/projects/day/TDayDiss/index.html>>. Acesso em: 16 fev. 2019.

DEEHAN, M.; GARCÊS, S.; KRAMER, D.; BAKER, M. P.; RAT, D.; ROETTGER, Y.; KROMMINGA, A. Managing unwanted immunogenicity of biologicals. **Autoimmunity Reviews**, Elsevier, p. 1–6, 2015.

DHANDA, S. K.; GUPTA, S.; VIR, P.; RAGHAVA, G. P. Prediction of IL4 inducing peptides. **Clinical and Developmental Immunology**, Hindawi, v. 2013, p. 1–10, 2013.

DIMITROV, I.; GARNEV, P.; FLOWER, D. R.; DOYTCHINOVA, I. EpiTOP—a proteochemometric tool for MHC class II binding prediction. **Bioinformatics**, Oxford University Press, v. 26, p. 2066–2068, 2010.

DONNES, P.; KOHLBACHER, O. SVMHC: a server for prediction of MHC binding peptides. **Nucleic Acids Research**, Oxford University Press, v. 34, p. W194–W197, 2006.

DORAK, M. T.; LAWSON, T.; MACHULLA, H. K.; DARKE, C.; MILLS, K. I.; BURNETT, A. K. Unravelling an HLA-DR association in Childhood Acute Lymphoblastic Leukemia. **Blood Journal**, The American Society of Hematology, v. 94, p. 694–700, 1999.

DROUIN, E. E.; SEWARD, R. J.; STRLE, K.; MCHUGH, G.; KATCHAR, K.; LONDONO, D.; YAO, C.; COSTELLO, C. E.; STEERE, A. C. A novel human autoantigen, endothelial cell growth factor, is a target of T and B cell responses in patients with lyme disease. **Arthritis and Rheumatology**, American College of Rheumatology, v. 65, p. 186–196, 2013.

DUVAL, M.; SUCIU, S.; FERSTER, A.; RIALLAND, X.; NELKEN, B.; LUTZ, P.; BENOIT, Y.; ROBERT, A.; MANEL, A.-M.; VILMER, E.; OTTEN, J.; PHILIPPE, N. Comparison of *Escherichia coli*-asparaginase with *Erwinia*-asparaginase in the treatment of childhood lymphoid malignancies: Results of a randomized European Organisation for Research and Treatment of Cancer-Children’s Leukemia Group phase 3 trial. **Blood Journal**, The American Society of Hematology, v. 99, p. 2734–2739, 2002.

EGLER, R. A.; AHUJA, S. P.; MATLOUB, Y. L-asparaginase in the treatment of patients with acute lymphoblastic leukemia. **Journal of Pharmacology and Pharmacotherapeutics**, Wolters Kluwer - Medknow, v. 7, p. 62–71, 2016.

ETTINGER, L. J.; KURTZBERG, J.; VOÛTE, P. A.; JURGENS, H.; HALPERN, S. L. An open-label, multicenter study of Polyethylene Glycol-L-Asparaginase for the treatment of Acute Lymphoblastic Leukemia. **Cancer**, Wiley, v. 75, p. 1176–1181, 1994.

EVAVOLD, B. D.; SLOAN-LANCASTER, J.; HSU, B. L.; ALLEN, P. M. Separation of T helper 1 clone cytolysis from proliferation and lymphokine production using analog peptides. **The Journal of Immunology**, American Association of Immunologists, v. 150, p. 3131–3140, 1993.

FERNANDES, H. S.; TEIXEIRA, C. S. S.; FERNANDES, P.; RAMOS, M.; CERQUEIRA, N. M. F. S. A. Amino acid deprivation using enzymes as a targeted therapy for cancer and viral infections. **Expert Opinion on Therapeutic Patents**, Taylor and Francis Group, 2016.

FERNANDEZ, C. A.; SMITH, C.; YANG, W.; DATÉ, M.; BASHFORD, D.; LARSEN, E.; BOWMAN, W. P.; LIU, C.; RAMSEY, L. B.; CHANG, T.; TURNER, V.; LOH, M. L.; RAETZ,

- E. A.; WINICK, N. J.; HUNGER, S. P.; CARROLL, W. L.; ONENGUT-GUMUSCU, S.; CHEN, W.-M.; CONCANNON, P.; RICH, S. S.; SCHEET, P.; JEHA, S.; PUI, C.-H.; EVANS, W. E.; DEVIDAS, M.; RELLING, M. V. HLA-DRB1 0701 is associated with a higher risk of asparaginase allergies. **Blood Journal**, The American Society of Hematology, v. 124, p. 1266–1276, 2014.
- FLOWER, D. R. Predicting immunogenicity in silico. In: **Methods in Molecular Biology - Immunoinformatics**. [S.l.]: Humana Press, 2007. v. 409.
- FREMONT, D. H.; HENDRICKSON, W. A.; MARRACK, P.; KAPPLER, J. Structures of an MHC Class II Molecule with Covalently Bound Single Peptides. **Nature**, v. 272, p. 1001–1004, 1996.
- GARBOCZI, D. N.; GHOSH, P.; UTZ, U.; FAN, Q. R.; BIDDISON, W. E.; WILEY, D. C. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. **Nature**, v. 384, p. 134–141, 1996.
- GESTO, D. S.; CERQUEIRA, N. M. F. S. A.; FERNANDES, P. A.; RAMOS, M. J. Unraveling the enigmatic mechanism of L-asparaginase II with QM/QM calculations. **Journal of the American Chemical Society**, American Chemical Society, v. 135, p. 7146–7158, 2013.
- GHOSH, P.; AMAYA, M.; MELLINS, E.; WILEY, D. C. The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. **Nature**, v. 378, p. 457–462, 1995.
- GRISWOLD, K. E.; BAILEY-KELLOGG, C. Design and engineering of deimmunized biotherapeutics. **Current Opinion in Structural Biology**, PMC, v. 39, p. 79–88, 2016.
- GROOT, A. S. D.; SCOTT, D. W. Immunogenicity of protein therapeutics. **Trends in Immunology**, Elsevier, v. 28, p. 482–490, 2007.
- GUAN, P.; DOYTCHINOVA, I. A.; ZYGOURI, C.; FLOWER, D. R. MHCpred: a server for quantitative prediction of peptide–MHC binding. **Nucleic Acids Research**, Oxford University Press, v. 31, p. 3621–3624, 2003.
- GUPTA, S. K.; SRIVASTAVA, M.; AKHOON, B. A.; SMITA, S.; SCHMITZ, U.; WOLKENHAUER, O.; VERA, J.; GUPTADE, S. K. Identification of immunogenic consensus T-cell epitopes in globally distributed influenza-A H1N1 neuraminidase. **Infection, Genetics and Evolution**, Elsevier, v. 11, p. 308–319, 2011.
- HAMMER, J.; BELUNIS, C.; BOLIN, D.; PAPADOPOULOS, J.; WALSKY, R.; HIGELIN, J.; DANHO, W.; SINIGAGLIA, F.; NAGY, Z. A. High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 91, p. 4456–4460, 1994.
- HAMMER, J.; BONO, E.; GALLAZZI, F.; BELUNIS, C.; NAGY, Z.; SINIGAGLIA, F. Precise prediction of major histocompatibility complex class II-peptide interaction based on peptide side chain scanning. **Journal of Experimental Medicine**, The Rockefeller University Press, v. 80, p. 2353–2358, 1994.
- HAMMER, J.; VALSASNINI, P.; TOLBA, K.; BOLIN, D.; HIGELIN, J.; TAKACS, B.; SINIGAGLIA, F. Promiscuous and allele-specific anchors in HLA-DR-binding peptides. **Nature**, v. 74, p. 197–203, 1993.

- HENIKOFF, S.; HENIKOFF, J. G. Amino acid substitution matrices from protein blocks. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 89, p. 10915–10919, 1992.
- HILL, C. M.; HAYBALL, J. D.; ALLISON, A. A.; ROTHBARD, J. B. Conformational and structural characteristics of peptides binding to HLA-DR molecules. **The Journal of Immunology**, American Association of Immunologists, v. 147, p. 189–197, 1991.
- HILL, C. M.; LIU, A.; MARSHALL, K. W.; MAYER, J.; JORGENSEN, B.; YUAN, B.; CUBBON, R. M.; NICHOLS, E. A.; WICKER, L. S.; ROTHBARD, J. B. Exploration of requirements for peptide binding to HLA DRB1*0101 and DRB1*0401. **The Journal of Immunology**, American Association of Immunologists, v. 152, p. 2890–2898, 1994.
- HO, D. H. W.; THETFORD, B.; CARTER, C. J. K.; III, E. F. Clinical pharmacologic studies of L-asparaginase. **Clinical Pharmacology and Therapeutics**, American Society for Clinical Pharmacology and Therapeutics, v. 11, p. 408–417, 1970.
- HOLLAND, J. H. In: **Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence**. [S.l.]: U Michigan Press, 1975. p. 183.
- HSU, B. L.; EVAVOLDS, B. D.; ALLEN, P. M. Modulation of T cell development by an endogenous altered peptide ligand. **The Journal of Experimental Medicine**, Rockefeller University Press, v. 181, p. 805–810, 1995.
- JAMES, E. A.; MOUSTAKAS, A. K.; BUI, J.; PAPADOPOULOS, G. K.; BONDINAS, G.; BUCKNER, J. H.; KWOK, W. W. DR1001 presents ‘altered-self’ peptides derived from joint associated proteins by accepting citrulline in three of its binding pockets. **Arthritis and Rheumatism**, v. 62, p. 2909–2918, 2010.
- JARDETZKY, T. S.; GORGA, J. C.; BUSCH, R.; ROTHBARD, J.; STROMINGER, J. L.; WILEY, D. C. Peptide binding to HLA-DR1: a peptide with most residues substituted to alanine retains MHC binding. **The Embo Journal**, EMBO Press, v. 9, p. 1797–1803, 1990.
- JAWA, V.; COUSENS, L. P.; AWWAD, M.; WAKSHULL, E.; KROPSHOFER, H.; GROOT, A. S. D. T-cell dependent immunogenicity of protein therapeutics: Preclinical assessment and mitigation. **Clinical Immunology**, Elsevier, v. 149, p. 534–555, 2013.
- JONES, T. D.; CROMPTON, L. J.; CARR, F. J.; BAKER, M. P. Deimmunization of monoclonal antibodies. In: **Methods in Molecular Biology**. [S.l.]: Springer, 2009. v. 525, p. 405–423.
- KIDD, J. G. Regression of transplanted lymphomas induced in vivo by means of normal guinea pig serum: Studies on the nature of the active serum constituent: Histological mechanism of the regression: Tests for effects of guinea pig serum on lymphoma cells in vitro: Discussion. **Journal of Experimental Medicine**, v. 98, n. 6, p. 583–606, 1953.
- KUFAREVA, I.; ABAGYAN, R. Methods of protein structure comparison. **Homology Modeling**, Methods and Protocols - Humana Press, v. 857, p. 231–257, 2011.
- LINDORFF-LARSEN, K.; PACI, E.; SERRANO, L.; DOBSON, C. M.; VENDRUSCOLO, M. Calculation of mutational free energy changes in transition states for protein folding. **Biophysical Journal**, Biophysical Society, v. 85, p. 1207–1214, 2003.

LIU, J.; WANG, L.; ZHAO, F.; TSENG, S.; NARAYANAN, C.; SHURA, L.; WILLINGHAM, S.; HOWARD, M.; PROHASKA, S.; VOLKMER, J.; CHAO, M.; WEISSMAN, I.; MAJETI, R. Pre-clinical development of a humanized anti-CD47 antibody with anti-cancer therapeutic potential. **Chemistry and Biology**, Public Library of Science, v. 10, p. 1–23, 2015.

LOPES, A. M.; OLIVEIRA-NASCIMENTO, L. de; RIBEIRO, A.; JR, C. A. T.; BREYER, C. A.; OLIVEIRA, M. A. de; MONTEIRO, G.; SOUZA-MOTTA, C. M. de; MAGALHÃES, P. de O.; AVENDAÑO, J. G. F.; CAVACO-PAULO, A. M.; MAZZOLA, P. G.; RANGEL-YAGUI, C. de O.; SETTE, A. C. L. D.; PESSOA, A. Therapeutic L-asparaginase: upstream, downstream and beyond. **Critical Reviews in Biotechnology**, Taylor and Francis Group, v. 37, n. 1, p. 82–99, 2017.

MA, J.; WANG, S. Chapter 5 - Algorithms, applications, and challenges of protein structure alignment. In: **Advances in Protein Chemistry and Structural Biology**. [S.l.]: Elsevier, 2014. v. 94, p. 121–175.

MADDEN, D. R. The three-dimensional structure of peptide-MHC complexes. **Annual Review of Immunology**, Annual Reviews, v. 13, p. 587–622, 1995.

MALHERBE, L. T-cell epitope mapping. **Annals of Allergy, Asthma and Immunology**, Elsevier, v. 103, p. 76–79, 2009.

MARSHALL, K. W.; WILSON, K. J.; LIANG, J.; WOODS, A.; ZALLER, D.; ROTHBARD, J. B. Prediction of peptide affinity to HLA DRB1*0401. **The Journal of Immunology**, American Association of Immunologists, v. 154, p. 5927–5933, 1995.

MASHBURN, L. T.; WRISTON, J. C. Tumor inhibitory effect of L-asparaginase from *Escherichia coli*. **Archives of Biochemistry and Biophysics**, Elsevier, v. 105, p. 450–453, 1964.

MEHTA, R. K.; VERMA, S.; PATI, R.; SENGUPTA, M.; KHATUA, B.; JENA, R. K.; SETHY, S.; KAR, S. K.; MANDAL, C.; ROEHM, K. H.; SONAWANE, A. Mutations in subunit interface and b-cell epitopes improve antileukemic activities of *Escherichia coli* asparaginase-II. **The Journal of Biological Chemistry**, The American Society for Biochemistry and Molecular Biology, v. 289, p. 3555–3570, 2014.

MOTIEJUNAS, D.; WADE, R. C. Chapter 4 - Structural, energetic, and dynamic aspects of ligand–receptor interactions. In: **Comprehensive Medicinal Chemistry II**. [S.l.]: Elsevier, 2007. v. 4, p. 193–213.

MULLER, H.; BOOS, J. Use of L-asparaginase in childhood ALL. **Critical Reviews in Oncology Hematology**, Elsevier, v. 28, p. 97–113, 1998.

NARTA, U. K.; KANWAR, S. S.; AZMI, W. Pharmacological and clinical evaluation of L-asparaginase in the treatment of leukemia. **Critical Reviews in Oncology Hematology**, Elsevier, v. 61, n. 3, p. 208–221, 2007.

NATIONAL CANCER INSTITUTE. **Childhood Acute Lymphoblastic Leukemia Treatment (PDQ) – Patient Version**. 2018. Disponível em: <<https://www.cancer.gov/types/leukemia/patient/child-all-treatment-pdq>>. Acesso em: 27 out. 2018.

NEUMAN, R. E.; MCCOY, T. A. Dual requirement of walker carcinoma 256 in vitro for asparagine and glutamine. **Science**, v. 124, n. 3212, p. 124–125, 1956.

NIELSEN, M.; LUND, O.; BUUS, S.; LUNDEGAARD, C. MHC Class II epitope predictive algorithms. **Immunology**, Wiley, v. 130, p. 319–328, 2010.

NIELSEN, M.; LUNDEGAARD, C.; BLICHER, T.; PETERS, B.; SETTE, A.; JUSTESEN, S.; BUUS, S.; LUND, O. Quantitative predictions of peptide binding to any HLA-DR molecule of known sequence: NetMHCIIpan. **PLOS Computational Biology**, Public Library of Science, v. 4, p. 421–423, 2008.

NIELSEN, M.; LUNDEGAARD, C.; LUND, O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. **BMC Bioinformatics**, BioMed Central, v. 8, p. 238–249, 2007.

NIELSEN, M.; LUNDEGAARD, C.; LUND, O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. **BMC Bioinformatics**, BioMed Central, v. 8, p. 1–12, 2007.

NIELSEN, M.; LUNDEGAARD, C.; WORNING, P.; HVID, C. S.; LAMBERTH, K.; BUUS, S.; BRUNAK, S.; LUND, O. Improved prediction of MHC class I and class II epitopes using a novel Gibbs sampling approach. **Bioinformatics**, Oxford University Press, v. 20, p. 1388–1397, 2004.

OFFMAN, M. N.; KROL, M.; PATEL, N.; KRISHNAN, S.; LIU, J.; SAHA, V.; BATES, P. A. Rational engineering of L-asparaginase reveals importance of dual activity for cancer cell toxicity. **Blood Journal**, The American Society of Hematology, v. 117, p. 1614–1621, 2011.

OHNUMA, T.; BERGEL, F.; BRAY, R. C. Enzymes in cancer: Asparaginase from chicken liver. **The Biochemistry Journal**, v. 103, p. 238–245, 1967.

OHNUMA, T.; HOLLAND, J. F.; FREEMAN, A.; SINKS, L. F. Biochemical and pharmacological studies with L-asparaginase in man. **Cancer Research**, American Association for Cancer Research, v. 30, p. 2297–2305, 1970.

O’SULLIVAN, D.; ARRHENIUS, T.; SIDNEY, J.; GUERCIO, M. F. D.; ALBERTSON, M.; WALL, M.; OSEROFF, C.; SOUTHWOOD, S.; COLON, S. M.; GAETA, F. C.; SETTE, A. On the interaction of promiscuous antigenic peptides with different DR alleles: identification of common structural motifs. **The Journal of Immunology**, American Association of Immunologists, v. 147, p. 2663–2669, 1991.

OYARZUN, P.; ELLIS, J. J.; BODEN, M.; KOBE, B. PREDIVAC: CD4+ T-cell epitope prediction for vaccine design that covers 95% of HLA class II DR protein diversity. **BMC Bioinformatics**, BioMed Central, v. 14, p. 52–62, 2013.

PARKER, A. S.; CHOI, Y.; GRISWOLD, K. E.; BAILEY-KELLOGG, C. Structure-guided deimmunization of therapeutic proteins. **Journal of Computational Biology**, Mary Ann Liebert, v. 20, p. 152–165, 2013.

PARKER, A. S.; ZHENG, W.; GRISWOLD, K. E.; BAILEY-KELLOGG, C. Optimization algorithms for functional deimmunization of therapeutic proteins. **BMC Bioinformatics**, BioMed Central, v. 11, p. 1–15, 2010.

PARKER, K. C.; BEDNAREK, M. A.; COLIGAN, J. E. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. **The Journal of Immunology**, American Association of Immunologists, v. 152, p. 163–175, 1994.

PETROVA, D. Protein structure comparison methods. **Information Technologies and Control**, The John Atanasoff Society of Automatics and Informatics, v. 2, p. 32–43, 2009.

PIETERS, R.; APPEL, I.; KUEHNEL, H.-J.; TETZLAFF-FOHR, I.; PICHLMEIER, U.; VAART, I. van der; VISSER, E.; STIGTER, R. Pharmacokinetics, pharmacodynamics, efficacy, and safety of a new recombinant asparaginase preparation in children with previously untreated acute lymphoblastic leukemia: a randomized phase 2 clinical trial. **Blood Journal**, The American Society of Hematology, v. 112, p. 4832–4838, 2008.

PRAGER, M. D.; BACHYNSKY, N. Asparagine synthetase in normal and malignant tissues; correlation with tumor sensitivity to asparaginase. **Archives of Biochemistry and Biophysics**, Elsevier, v. 127, p. 645–654, 1968.

PUI, C.-H.; YANG, J. J.; HUNGER, S. P.; PIETERS, R.; SCHRAPPE, M.; BIONDI, A.; VORA, A.; BARUCHEL, A.; SILVERMAN, L. B.; SCHMIEGELOW, K.; ESCHERICH, G.; HORIBE, K.; BENOIT, Y. C.; IZRAELI, S.; YEOH, A. E. J.; LIANG, D.-C.; DOWNING, J. R.; EVANS, W. E.; RELLING, M. V.; MULLIGHAN, C. G. Childhood acute lymphoblastic leukemia: Progress through collaboration. **JOURNAL OF CLINICAL ONCOLOGY**, American Society of Clinical Oncology, v. 33, p. 2938–2948, 2015.

RAMMENSEE, H. G.; BACHMANN, J.; EMMERICH, N. P. N.; BACHOR, O. A.; STEVANOVIC, S. SYFPEITHI: database for MHC ligands and peptide motifs. **Immunogenetics**, Springer, v. 50, p. 213–219, 1999.

RCSB PDB. **Crystal structure of *Escherichia coli* L-asparaginase, an enzyme used in cancer therapy**. 1993. Disponível em: <<https://www.rcsb.org/structure/3eca>>. Acesso em: 6 set. 2018.

REAY, P. A.; KANTOR, R. M.; DAVIS, M. M. Use of global amino acid replacements to define the requirements for MHC binding and T cell recognition of moth cytochrome c (93-103). **The Journal of Immunology**, American Association of Immunologists, v. 152, p. 3946–3957, 1994.

RECHE, P. A.; GLUTTING, J. P.; ZHANG, H.; REINHERZ, E. L. Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. **Immunogenetics**, Springer, v. 56, p. 405–419, 2004.

RICHARDS, N. G. J.; SCHUSTER, S. M. Advances in enzymology and related areas of molecular biology. In: **Mechanistic issues in asparagine synthetase catalysis**. [S.l.]: Interscience, 1998. v. 72, p. 145–198.

RIZZARI, C.; CONTER, V.; STARY, J.; COLOMBINI, A.; MOERICKE, A.; SCHRAPPE, M. Optimizing asparaginase therapy for acute lymphoblastic leukemia. **Current Opinion in Oncology**, Wolters Kluwer Health, v. 25, p. S1–S9, 2013.

RIZZARI, C.; PUTTI, M. C.; COLOMBINI, A.; CASAGRANDA, S.; FERRARI, G. M.; PAPAYANNIDIS, C.; IACOBUCCI, I.; ABBENANTE, M. C.; SARTOR, C.; MARTINELLI, G. Rationale for a pediatric-inspired approach in the adolescent and young adult population with acute lymphoblastic leukemia, with a focus on asparaginase treatment. **Hematology Reports**, PAGEPress, v. 5554, n. 6, p. 51–57, 2014.

SAFDARI, Y.; FARAJNIA, S.; ASGHARZADEH, M.; KHALILI, M. Antibody humanization methods – a review and update. **Biotechnology and Genetic Engineering Reviews**, Taylor and Francis Group, v. 29, p. 175–186, 2013.

- SANCHEZ-TRINCADO, J. L.; GOMEZ-PEROSANZ, M.; RECHE, P. A. Fundamentals and methods for T- and B-cell epitope prediction. **Journal of Immunology Research**, Hindawi, v. 2017, p. 1–14, 2017.
- SCHAEFER, C.; ROST, B. Predict impact of single amino acid change upon protein structure. **BMC Genomics**, BioMed Central, v. 13, p. 1–10, 2012.
- SCHAEFFER, L. Chapter 14 - The role of functional groups in drug–Receptor interactions. In: **The Practice of Medicinal Chemistry**. [S.l.]: Academic Press, 2008. v. 4, p. 359–378.
- SHELLEKENS, H. Immunogenicity of Therapeutic Proteins: Clinical Implications and Future Prospects. **Clinical Therapeutics**, Elsevier, v. 24, p. 1720–1740, 2002.
- SCHILD, H.; GRÜNEBERG, U.; POUGIALIS, G.; KEILHOLZ, H.-J. W. W.; STEVANOVIĆ, S.; RAMMENSEE, H.-G. Natural ligand motifs of H-2E molecules are allele specific and illustrate homology to HLA-DR molecules. **International Immunology**, Oxford University Press, v. 7, p. 1957–1965, 1995.
- SCHMIEGELOW, K.; FORESTIER, E.; HELLEBOSTAD, M.; HEYMAN, M.; KRISTINSSON, J.; SÖDERHÄLL, S.; TASKINEN, M. Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. **Leukemia**, Nature, v. 24, p. 345–354, 2010.
- SCHOLZ, E. M.; MARCILLA, M.; DAURA, X.; ARRIBAS-LAYTON, D.; JAMES, E. A.; ALVAREZ, I. Human Leukocyte Antigen (HLA)-DRB1*15:01 and HLA-DRB5*01:01 Present Complementary Peptide Repertoires. **Frontiers in Immunology**, v. 8, p. 1–12, 2017.
- SCHYMKOWITZ, J.; BORG, J.; STRICHER, F.; NYS, R.; ROUSSEAU, F.; SERRANO, L. The foldx web server: an online force field. **Nucleic Acids Research**, Oxford University Press, v. 33, p. W382–W388, 2005.
- SETTE, A.; SIDNEY, J.; OSEROFF, C.; GUERCIO, M. F. del; SOUTHWOOD, S.; ARRHENIUS, T.; POWELL, M. F.; COLÓN, S. M.; GAETA, F. C.; GREY, H. M. HLA DR4w4-binding motifs illustrate the biochemical basis of degeneracy and specificity in peptide-DR interactions. **The Journal of Immunology**, American Association of Immunologists, v. 151, p. 3163–3170, 1993.
- SHARMA, G.; HOLT, R. A. T-cell epitope discovery technologies. **Human Immunology**, Elsevier, v. 75, p. 514–519, 2014.
- SHRIVASTAVA, A.; KHANA, A. A.; KHURSHIDB, M.; KALAMB, M. A.; JAIND, S. K.; SINGHAL, P. K. Recent developments in L-asparaginase discovery and its potential as anticancer agent. **Critical Reviews in Oncology Hematology**, Elsevier, v. 100, n. 2016, p. 1–10, 2015.
- SILVERTHORN, D. U. Capítulo 24 - O sistema imune. In: **Fisiologia Humana - Uma abordagem integrada**. [S.l.]: Artmed, 2017. v. 7. ISBN 9780321981226.
- SINGH, H.; RAGHAVA, G. P. S. ProPred: prediction of HLA-DR binding sites. **Bioinformatics**, Oxford University Press, v. 17, p. 1236–1237, 2001.
- SO, T.; ITO, H.-O.; KOGA, T.; UEDA, T.; IMOTO, T. Reduced immunogenicity of monomethoxypolyethylene glycol-modified lysozyme for activation of T cells. **Immunology Letters**, Elsevier, v. 49, p. 91–97, 1996.

- SOUTHWOOD, S.; SIDNEY, J.; KONDO, A.; GUERCIO, M.-F. del; APPELLA, E.; HOFFMAN, S.; KUBO, R. T.; CHESNUT, R. W.; GREY, H. M.; SETTE, A. Several common HLA-DR types share largely overlapping peptide binding repertoires. **The Journal of Immunology**, American Association of Immunologists, v. 160, p. 3363–3373, 1998.
- SPAIN J L JORGENSEN, M. M. D. L. M.; BERG, L. J. A peptide antigen antagonist prevents the differentiation of T cell receptor transgenic thymocytes. **The Journal of Immunology**, American Association of Immunologists, v. 152, p. 1709–1717, 1994.
- STAMS, W. A. G.; BOER, M. L. den; HOLLEMAN, A.; APPEL, I. M.; BEVERLOO, H. B.; WERING, E. R. van; JANKA-SCHAUB, G. E.; EVANS, W. E.; PIETERS, R. Asparagine synthetase expression is linked with L-asparaginase resistance in TEL-AML1-negative but not TEL-AML1-positive pediatric acute lymphoblastic leukemia. **Blood Journal**, The American Society of Hematology, v. 105, p. 4223–4225, 2005.
- STERN, L. J.; BROWN, J. H.; JARDETZKY, T. S.; GORGA, J. C.; URBAN, R. G.; STROMINGER, J. L.; WILEY, D. C. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. **Nature**, v. 368, p. 215–221, 1994.
- STRICKLER, N.; BALABANOV, S.; CASAURO, K.; SCHANZ, U.; MANZ, M. G.; GERBER, B. Acute central nervous system complications and ammonium levels in adult patients with acute lymphoblastic leukemia receiving L-asparaginase. **Leukemia and Lymphoma**, Taylor and Francis Group, v. 59, p. 855–862, 2018.
- STURNIOLO, T.; BONO, E.; DING, J.; RADDRIZZANI, L.; TUERECI, O.; SAHIN, U.; BRAXENTHALER, M.; GALLAZZI, F.; PROTTI, M. P.; SINIGAGLIA, F.; HAMMER, J. Generation of tissuespecific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. **Nature Biotechnology**, Nature America, v. 17, p. 555–561, 1999.
- SUTOW, W. W.; GEORGE, S.; LOWMAN, J. T.; STARLING, K. A.; HUMPHREY, G. B.; HAGGARD, M. E.; VIETTI, T. J. Evaluation of dose and schedule of L-asparaginase in multidrug therapy of childhood leukemia. **Medical Pediatric Oncology**, Wiley, v. 2, p. 387–395, 1976.
- THAKALLAPALLY, R.; KIBBE, W.; LANG, D.; KORBER, B. Motif Scan: A web-based tool to find HLA anchor residues in proteins or peptides. In: . [S.l.: s.n.], 2000.
- TRUEWORTHY, R. C.; SUTOW, W. W.; PULLEN, J.; KOMP, D.; BERRY, D. H. Repeated use of L-asparaginase in multi-drug therapy of childhood leukemia. **Medical Pediatric Oncology**, Wiley, v. 4, p. 91–97, 1978.
- WAN, H. An overall comparison of small molecules and large biologics in adme testing. **ADMET and DMPK**, IAPC Publishing, v. 4, p. 1–22, 2016.
- WAN, J.; LIU, W.; XU, Q.; REN, Y.; FLOWER, D. R.; LI, T. SVRMHC prediction server for MHC-binding peptides. **BMC Bioinformatics**, BioMed Central, v. 7, p. 1–5, 2006.
- WANG, P.; SIDNEY, J.; DOW, C.; MOTHÉ, B.; SETTE, A.; PETERS, B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. **PLOS Computational Biology**, Public Library of Science, v. 4, p. 1–10, 2008.
- WANG, W. Instability, stabilization, and formulation of liquid protein pharmaceuticals. **International Journal of Pharmaceutics**, Elsevier Science, v. 185, p. 129–188, 1999.

WEAVER, J. M.; LAZARSKI, C. A.; RICHARDS, K. A.; CHAVES, F. A.; JENKS, S. A.; MENGES, P. R.; SANT, A. J. Immunodominance of CD4 T Cells to foreign antigens is peptide intrinsic and independent of molecular context: Implications for vaccine design. **The Journal of Immunology**, v. 181, n. 5, p. 3039–3048, 2008.

XIANG, Z.; HE, Y. Vaxign: a web-based vaccine target design program for reverse vaccinology. **Procedia in Vaccinology**, Elsevier, v. 1, p. 23–29, 2009.

YU, X.; GEN, M. In: **Introduction to evolutionary algorithms**. [S.l.]: Springer, 2010.

ZHANG, F.; LIU, M.-R.; WAN, H.-T. Discussion about several potential drawbacks of PEGylated therapeutic proteins. **Biological and Pharmaceutical Bulletin**, The Pharmaceutical Society of Japan, v. 37, p. 335–339, 2014.

ZHANG, G. L.; DELUCA, D. S.; KESKIN, D. B.; CHITKUSHEV, L.; ZLATEVA, T.; LUND, O.; REINHERZ, E. L.; BRUSIC, V. MULTIPRED2: a computational system for large-scale identification of peptides predicted to bind to HLA supertypes and alleles. **Journal of Immunological Methods**, Elsevier, v. 374, p. 53–61, 2011.

ZHANG, Q.; WANG, P.; KIM, Y.; HASTE-ANDERSEN, P.; BEAVER, J.; BOURNE, P. E.; BUI, H.-H.; BUUS, S.; FRANKILD, S.; GREENBAUM, J.; LUND, O.; LUNDEGAARD, C.; NIELSEN, M.; PONOMARENKO, J.; SETTE, A.; ZHU, Z.; PETERS, B. Immune epitope database analysis resource (IEDB-AR). **Nucleic Acids Research**, Oxford University Press, v. 36, p. W513–W518, 2008.

ZHAO, H.; VERMA, D.; LI, W.; CHOI, Y.; NDONG, C.; FIERING, S. N.; BAILEY-KELLOGG, C.; ; GRISWOLD, K. E. Depletion of T cell epitopes in lysostaphin mitigates anti-drug antibody response and enhances antibacterial efficacy *in vivo*. **Chemistry and Biology**, Elsevier, v. 22, p. 629–639, 2015.

APÊNDICE A – TERMOS DE ENERGIA CONSIDERADOS PELO FOLDX

A estabilidade de uma proteína é determinada por um grande número de pequenas energias de interação positivas e negativas (DAY, 1996), tais como interações hidrofóbicas, ligações de hidrogênio, formação de pontes salinas, repulsão eletrostática, ligações dissulfeto, interações com componentes do solvente, e outras interações de natureza eletrostática, que são resultado de interações entre átomos pertencentes à proteína e entre átomos da proteína e do solvente (WANG, 1999). Para a termodinâmica, a estabilidade é entendida como a variação de energia livre (ΔG) resultante do processo de enovelamento (LINDORFF-LARSEN *et al.*, 2003).

O FoldX (SCHYMKOWITZ *et al.*, 2005) é um algoritmo de *design* de proteínas que usa um campo de força empírico desenvolvido para a estimativa de mudanças de energia livre mutacional sobre a estabilidade de uma proteína ou um complexo proteico. A análise da estabilidade é feita através de um cálculo de energia livre que produz a ΔG do enovelamento (em kcal/mol) decomposta nos diferentes termos de energia usados pelo FoldX: ligação de hidrogênio entre os átomos da cadeia principal e entre os átomos de cadeia lateral, energia dada pela formação de pontes dissulfeto, ligações específicas entre um íon metálico e átomos polares ou carregados, contribuição eletrostática devido à inclusão do dipolo helicoidal, custo de entropia que surge quando a liberdade rotacional e translacional de duas moléculas de proteína é restrita em relação uma à outra devido à sua ligação, além do custo de entropia adicional da formação de complexos causado pela restrição conformacional da cadeia principal ou das cadeias laterais.

O algoritmo também calcula penalidades diante da ocorrência de colisões estéricas, que ocorre quando os átomos se interpenetram, ou seja, quando a distância entre dois átomos é menor que a soma de seus raios de Van der Waals, resultando em uma forte interação repulsiva. Essas colisões são divididas em choque torsional, que surge quando uma cadeia lateral adota uma conformação que causa uma sobreposição entre seus próprios átomos ou sua cadeia principal e colisões de Van der Waals, que são os choques entre átomos de diferentes resíduos.

O FoldX se encontra disponível para uso público e possui precisão similar aos campos de força física na predição de variações de energia livre, mas gera resultados com maior rapidez, já que a estimativa das contribuições entrópicas para as interações de proteína é diretamente derivada da estrutura usando uma abordagem empírica. Por esta razão, o FoldX foi utilizado para o cálculo de $\Delta\Delta G$ após as mutações realizadas pelo algoritmo genético, garantindo que o mutante fosse tão estável quanto a proteína selvagem.

APÊNDICE B – MATRIZ DE PONTUAÇÃO BLOSUM62

A matriz BLOSUM62 foi desenvolvida analisando as frequências de substituições de aminoácidos em grupos de proteínas relacionadas, de forma que as sequências de aminoácidos em cada bloco eram pelo menos 62% idênticas quando duas proteínas estavam alinhadas (HENIKOFF; HENIKOFF, 1992). As frequências de todas as substituições de aminoácidos que ocorreram nesses blocos conservados de proteínas foram determinadas computacionalmente e esses dados foram usados para construir a matriz de pontuação BLOSUM62 para substituições de aminoácidos.

Para calcular a diferença estatística da substituição do aminoácido i pelo aminoácido j , a equação utilizada é:

$$S_{ij} = \frac{1}{\lambda} \log \frac{P_{ij}}{q_i q_j}$$

onde P_{ij} é a probabilidade da substituição dos aminoácidos i e j em proteínas homólogas, e q_i e q_j são as probabilidades de encontrar os aminoácidos i e j em qualquer sequência proteica. O termo $(1 / \lambda)$ é um fator de escala usado para que a matriz contenha valores inteiros facilmente computáveis.

BLOSUM62 é uma matriz triangular, o que significa que pontuação atribuída à substituição do aminoácido i pelo j é a mesma que a da substituição de j por i e os maiores valores são os da diagonal da matriz. No geral, pontuações positivas são menos comuns do que pontuações negativas, sugerindo que a maioria das substituições afeta negativamente a proteína (MA; WANG, 2014).