

ANDERSON APARECIDO DO ESPIRITO SANTO

**Algoritmos genéticos aplicados ao planejamento
racional termodinâmico de anticorpos miméticos
baseados no domínio GB1 da proteína G estreptocócica:
Um estudo de simulação molecular e validação
experimental**

Tese apresentada ao Instituto de Química
da Universidade Estadual Paulista, como
parte dos requisitos para obtenção do título
de doutor em Biotecnologia, sob a orienta-
ção do Prof. Dr. Gustavo Troiano Feliciano

Araraquara

2024

S237a

Santo, Anderson Aparecido do Espírito

Algoritmos genéticos aplicados ao planejamento racional termodinâmico de anticorpos miméticos baseados no domínio GB1 da proteína G estreptocócica : um estudo de simulação molecular e validação experimental / Anderson Aparecido do Espírito Santo. -- Araraquara, 2024
135 f. : il.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Instituto de Química, Araraquara

Orientador: Gustavo Troiano Feliciano

1. Algoritmos genéticos. 2. Reconhecimento molecular. 3. Expressão gênica. 4. Computação evolutiva. 5. Dinâmica molecular. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Química, Araraquara. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Impacto potencial desta pesquisa

Neste estudo apresentamos uma nova metodologia que permite a reprogramação da estrutura e função das proteínas, projetando moléculas bioativas inteiramente originais. Esta nova abordagem redefine técnicas computacionais amplamente utilizadas na pesquisa acadêmica como agente ativo na criação de novos produtos biotecnológicos.

Potential impact of this research

In this work, we present a novel methodology that enables the reprogramming of protein structure and function, designing entirely original bioactive molecules. This new approach redefines widely used computational techniques in academic research as an active agent in the creation of new biotechnological products.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: Algoritmos genéticos aplicados ao planejamento racional termodinâmico de anticorpos miméticos baseados no domínio GB1 da proteína G estreptocócica: Um estudo de simulação molecular e validação experimental

AUTOR: ANDERSON APARECIDO DO ESPIRITO SANTO

ORIENTADOR: GUSTAVO TROIANO FELICIANO

Aprovado como parte das exigências para obtenção do Título de Doutor em Biotecnologia, pela Comissão Examinadora:

Documento assinado digitalmente
 **GUSTAVO TROIANO FELICIANO**
Data: 18/03/2024 07:44:06-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. GUSTAVO TROIANO FELICIANO (Participação Virtual)
Departamento de Engenharia, Física e Matemática / Instituto de Química de Araraquara - Unesp

Prof. Dr. ALEXANDRE SUMAN DE ARAÚJO (Participação Virtual)
Departamento de Física / Unesp/Câmpus de São José do Rio Preto

Prof. Dr. MAURÍCIO DOMINGUES COUTINHO NETO (Participação Virtual)
Universidade Federal do ABC - UFABC - Santo André

Prof. Dr. RAFAEL PLANA SIMÕES (Participação Virtual)
Depto. de Bioprocessos e Biotecnologia / FCA/Botucatu - Unesp

Prof. Dr. EUDES ETERNO FILETI (Participação Virtual)
Instituto de Ciência e Tecnologia - UNIFESP - São José dos Campos

Araraquara, 08 de março de 2024

DADOS CURRICULARES

IDENTIFICAÇÃO

Nome: Anderson Aparecido do Espírito Santo

Nome em citações bibliográficas: ESPIRITO SANTO, Anderson A.

FORMAÇÃO ACADÊMICA/TITULAÇÃO:

Doutorado em Biotecnologia

Universidade Estadual Paulista "Júlio de Mesquita Filho"(UNESP)

Instituto de Química, Campus de Araraquara

2019-2024 Tese: Algoritmos genéticos aplicados ao planejamento racional termodinâmico de anticorpos miméticos baseados no domínio GB1 da proteína G estreptocócica: Um estudo de simulação molecular e validação experimental

Orientador: Professor Dr. Gustavo Troiano Feliciano

Mestrado em Biotecnologia

Universidade Estadual Paulista "Júlio de Mesquita Filho"(UNESP)

2017-2019 Instituto de Química, Campus de Araraquara

Dissertação: Estudo sobre a variação da energia livre de ligação em complexos lectina-peroxidase através da dinâmica molecular

Orientador: Professor Dr. Gustavo Troiano Feliciano

Bacharelado em Ciências Físicas e Biomoleculares

2012-2016 Universidade de São Paulo (USP)

Instituto de Física de São Carlos (IFSC)

ARTIGOS COMPLETOS PUBLICADOS EM PERIÓDICOS

1. ESPIRITO SANTO, Anderson A.; FELICIANO, Gustavo T. Thermodynamic and structural aspects of molecular recognition in mannose-binding protein complexes: a theoretical study over hrp-artinm association. **Journal of Molecular Modeling**, v. 27, n. 4, 15 mar. 2021.
2. ESPIRITO SANTO, Anderson A.; FELICIANO, Gustavo T. Genetic Algorithms Applied to Thermodynamic Rational Design of Mimetic Antibodies Based on the GB1 Domain of Streptococcal Protein G: an atomistic simulation study. **The Journal of Physical Chemistry B**, v. 125, n. 29, p. 7985-7996, 15 jul. 2021.
3. ESPIRITO SANTO, Anderson A.; LAZAROTI, Vitor Hugo R.; FELICIANO, Gustavo T. Multidimensional redox potential/pKa coupling in multicopper oxidases from molecular dynamics: implications for the proton transfer mechanism. **Physical Chemistry Chemical Physics**, v. 23, n. 48, p. 27348-27354, 2021.
4. FREITAS, Luis Paulo M.; ESPIRITO SANTO, Anderson A.; LOURENÇO, Tuanan C.; SILVA, Juarez L. F. da; FELICIANO, Gustavo Troiano. Steric and Electrostatic Effects on the Diffusion of CH₄/CH₃OH in Copper-Exchanged Zeolites: insights from enhanced sampling molecular dynamics and free energy calculations. **Langmuir**, v. 37, n. 26, p. 8014-8023, 21 jun. 2021.
5. MACEDO, Lucyano J.A.; ESPIRITO SANTO, Anderson A.; SEDENHO, Graziela C.; HASSAN, Ayaz; IOST, Rodrigo M.; FELICIANO, Gustavo T.; CRESPILO, Frank N.. Three-dimensional catalysis and the efficient bioelectrocatalysis beyond surface chemistry. **Journal of Catalysis**, v. 401, p. 200-205, set. 2021.
6. MENDONÇA, João Paulo A. de; LOURENÇO, Tuanan C.; FREITAS, Luis Paulo M.; ESPIRITO SANTO, Anderson A.; FELICIANO, Gustavo T.; SILVA, Juarez L. F. da. Molecular dynamics investigation of the structural and energetic properties of CeO₂-MO_x (M = Gd, La, Ce, Zr) nanoparticles. **Materials Advances**, v. 2, n. 23, p. 7759-7772, 2021.

Agradecimentos

Gostaria de expressar minha sincera gratidão ao meu estimado orientador, Prof. Dr. Gustavo Troiano Feliciano, pela sua notável orientação ao longo do meu doutorado. Sua experiência, sabedoria e apoio foram fundamentais para o sucesso deste projeto. Seus ensinamentos foram inestimáveis, moldando não apenas minha pesquisa, mas também meu crescimento acadêmico e profissional.

Gostaria de expressar minha profunda gratidão a Aline Reis e Anderson Amendola Pinheiro pela inestimável colaboração na parte experimental do meu projeto de doutorado. Sua dedicação, competência e empenho foram elementos fundamentais para o sucesso das investigações realizadas. Agradeço sinceramente por sua contribuição significativa, que não apenas enriqueceu os resultados, mas também aprimorou a qualidade e a validade dos experimentos.

Quero expressar minha profunda gratidão ao Instituto de Química da UNESP de Araraquara e a todos os seus dedicados funcionários pelo excepcional suporte que recebi ao longo do meu doutorado. Agradeço sinceramente pela infraestrutura de pesquisa de alta qualidade e pelas instalações que proporcionaram um ambiente propício ao desenvolvimento do meu trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 pela concessão da bolsa durante o período de realização deste doutorado.

Resumo

O desenvolvimento de anticorpos miméticos (AM), capazes de combinar a elevada afinidade e seletividade dos anticorpos com o pequeno tamanho dos peptídeos, tem um enorme potencial de aplicações na biotecnologia. Neste estudo, demonstramos que o projeto de AM pode ser realizado por meio de algoritmos genéticos (AG) criados a partir da combinação de *softwares* normalmente usados em simulação molecular. Nossos resultados demonstram que, usando o domínio GB1 da proteína G estreptocócica como modelo, é possível otimizar a capacidade de reconhecimento molecular de uma população de AM em poucas gerações. Nos três casos de estudo apresentados, o AG foi capaz de otimizar a energia livre de ligação (EL) das populações de AM para valores melhores do que os ligantes nativos. A convergência do AG acontece rapidamente, demonstrando que a seleção criteriosa das populações iniciais, baseada nas interações intermoleculares das superfícies dos antígenos, tem potencial de enfrentar um dos principais desafios do projeto de moléculas bioativas. Os testes experimentais imunoenzimáticos, realizados em um dos AM para avaliação da afinidade antigênica, apresentaram resultados positivos. Estes resultados comprovam que o AG foi capaz de otimizar a capacidade de reconhecimento molecular de um dos AM. Uma das consequências significativas deste estudo é a descoberta de novos motivos estruturais, que podem ser concebidos de forma original e inovadora a partir da própria estrutura de AM, dispensando a necessidade de softwares de predição estrutural e bancos de dados preexistentes. Por meio do AG desenvolvido neste estudo, demonstramos a aplicação de um novo protocolo capaz de orientar métodos experimentais no desenvolvimento de novas moléculas bioativas.

Palavras-chaves: algoritmos genéticos; anticorpos miméticos; simulação molecular; energia livre de ligação.

Abstract

The development of mimetic antibodies (AM), capable of combining the high affinity and selectivity of antibodies with the small size of peptides, has enormous potential applications in biotechnology. In this study, we demonstrate that the design of AM can be achieved through genetic algorithms (AG) created from the combination of software commonly used in molecular simulation. Our results show that, using the GB1 domain of the streptococcal G protein as a model, it is possible to optimize the molecular recognition ability of a population of AM in a few generations. In the three presented case studies, the AG was able to optimize the binding free energy (EL) of the AM populations to values better than native ligands. AG convergence occurs rapidly, demonstrating that careful selection of initial populations, based on intermolecular interactions of antigen surfaces, has the potential to address one of the main challenges of bioactive molecule design. Experimental immunoenzymatic tests, conducted in one of the AM to assess antigenic affinity, yielded positive results. These results demonstrate that the AG was able to optimize the molecular recognition capability of one of the AM. One significant consequence of this study is the discovery of new structural motifs, which can be conceived in an original and innovative manner from the AM structure itself, eliminating the need for structural prediction software and preexisting databases. Through the AG developed in this study, we demonstrate the application of a new protocol capable of guiding experimental methods in the development of new bioactive molecules.

Key-words: genetic algorithms; mimetic antibodies; molecular simulation; binding free energy.

Lista de ilustrações

Figura 1 – O domínio GB1 é composto por 56 resíduos de aminoácidos que se enovelam em quatro fitas- β e uma α -hélice. As quatro fitas- β formam uma enorme área de superfície com pelo menos 15 aminoácidos (Em destaque na ilustração) que podem ser modificados para modelagem de um AM.	21
Figura 2 – Ilustração dos epítomos alvos dos AM. A) Receptor tirosina quinase (VEGFR) (PDB ID: 1FLT). B) Envelope viral HC-E2 (PDB ID: 5KZP). C) Domínio RBD (PDB ID: 7BZ5).	23
Figura 3 – Ilustração do complexo formado pelas estruturas dos peptídeos GB1 e VEGFR após a minimização da energia do campo de força Amber.	28
Figura 4 – Ilustração da estrutura do molde PDB da superfície do VEGFR. As regiões em vermelho e azul são aceptoras de ligação de hidrogênio. As regiões em branco são doadoras de ligação de hidrogênio. As regiões em ciano são hidrofóbicas.	30
Figura 5 – Ilustração do complexo formado pela estrutura do GB1 e o molde VEGFR utilizado pela função de pontuação.	30
Figura 6 – Ilustração da estrutura do mutante mGB1-1 melhor classificado na função de pontuação do rotâmeros para VEGFR. Os resíduos são projetados com seus rotâmeros na direção de maior pontuação.	31
Figura 7 – Ilustração do cruzamento inicial realizado pelo AG. Os dois pares de fitas- β de dois mutantes são trocados resultando em dois novos mutantes filhos.	33
Figura 8 – Fluxograma representando a metodologia completa utilizada no AG.	35
Figura 9 – Representação da ordenação dos resíduos conforme a função de pontuação para o antígeno VEGFR.	36
Figura 10 – Representação da ordenação dos resíduos conforme a função de pontuação para o antígeno HC-E2.	37
Figura 11 – Representação da ordenação dos resíduos conforme a função de pontuação para o antígeno RBD.	38
Figura 12 – Ilustração mostrando o sistema base para a DM no projeto de AM. A) Sistema formado pelo VEGFR e o VEGF. B) Sistema base formado pelo GB1 e o VEGFR.	39
Figura 13 – Ilustração da conformação final dos mutantes mGB1 da população inicial do AG após 100ns em cada DM.	40

Figura 14 – Energias calculadas em todas as gerações do AG para o sistema VEGFR. O critério de convergência é a EL entre o VEGF e o receptor VEGFR.	41
Figura 15 – Desvio padrão da energia calculada para o VEGFR, utilizado como critério de convergência, e a energia do mutante de melhor energia mGB1-11.	45
Figura 16 – Contribuições individuais dos resíduos do mutante mGB1-11 na EL total.	46
Figura 17 – Ilustração mostrando o sistema base para a DM no projeto de AM. A) Sistema formado pelo fragmento Fab do anticorpo e o envelope HC-E2. B) Sistema base formado pelo GB1 e o envelope HC-E2.	47
Figura 18 – Energias calculadas em todas as gerações do AG para o sistema HC-E2. O critério de convergência é a EL entre o Grupo Fab do anticorpo e o peptídeo HC-E2.	48
Figura 19 – Desvio padrão da energia calculada para o HC-E2, utilizado como critério de convergência, e a energia do mutante de melhor energia HCmGB1-81.	49
Figura 20 – Contribuições individuais dos resíduos do mutante HCmGB1-81 na EL total.	50
Figura 21 – Ilustração da região do domínio RBD (Azul) neutralizada por anticorpos. A) Estrutura do anticorpo B38 (vermelho) neutralizando a região de reconhecimento ACE2. B) Estrutura do domínio GB1 (vermelho) na região de otimização do sistema base.	51
Figura 22 – Energias calculadas em todas as gerações do AG para o sistema RBD. O critério de convergência é a EL entre o fragmento Fab do anticorpo B38 e o Domínio RBD da proteína spike.	52
Figura 23 – Desvio padrão da energia calculada para o RBD, utilizado como critério de convergência, e as energias do mutante de melhor energia SGB1-121.	54
Figura 24 – Valores da EL dos mutantes SGB1-121 _{MD} e SGB1-121 _{his-tag} com adição do termo entrópico dos modos normais de vibração.	55
Figura 25 – Contribuições individuais dos resíduos do mutante SGB1-121 na EL total.	56
Figura 26 – Ilustração da previsão estrutural da ferramenta de MH. Valores de QMEAN próximos a zero indicam grande similaridade com estruturas experimentais.	57
Figura 27 – Ilustração do modelo de ancoragem molecular realizado pelo servidor online HDOCK.	58

Figura 28 – Ilustração da sobreposição dos <i>frames</i> durante 10 ns finais das DMs do mutante SGB1-121. A) Sobreposição dos <i>frames</i> da trajetória do mutante SGB1-121 _{MD} . B) Sobreposição dos <i>frames</i> da trajetória do mutante SGB1-121 _{his-tag}	59
Figura 29 – Análise estrutural de todos os mutantes SGB1-121. A) Gráfico de RMSD dos mutantes com e sem his-tag. B) Gráfico de RMSF dos mutantes com e sem his-tag.	60
Figura 30 – Análise estrutural da estrutura secundária do mutante SGB1-121 durante a DM de 100 ns.	61
Figura 31 – Ilustração da sobreposição dos <i>frames</i> da trajetória de DM do mutante SGB1-121 _{MD} . Moléculas de água são coloridas em vermelho no interior do arcabouço proteico.	62
Figura 32 – Gráfico mostrando o espectro de dicroísmo circular das soluções contendo GB1-Ref e SGB1-121 (diluídas 150 vezes), com a previsão teórica fornecida pelo servidor online PDBMD2CD.	64
Figura 33 – Gráfico mostrando a porcentagem de inibição do extrato bruto não purificado de <i>E. Coli</i> e dos extratos purificados contendo GB1-Ref e SGB1-121. Valores acima de 30% são considerados positivos para inibição.	66
Figura 34 – Script shell usado na função de pontuação.	121
Figura 35 – Script shell usado para gerar a população inicial de mutantes do AG.	122
Figura 36 – Script shell usado para gerar a população de mutantes do AG.	123
Figura 37 – Script shell usado na DM do algoritmo genético.	124

Lista de tabelas

Tabela 1 – Parâmetros usados na calibração da função de pontuação de rotâmeros.	78
Tabela 2 – Ordem dos resíduos de aminoácidos mais bem classificados em cada posição na função de pontuação do rotâmeros	78
Tabela 3 – Validação Experimental do AG. Todos os valores de ΔG estão em quilocaloria por mol.	79
Tabela 4 – Primeira geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	83
Tabela 5 – Mutantes filhos da primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	83
Tabela 6 – Segunda geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	84
Tabela 7 – Mutantes filhos da segunda geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	84
Tabela 8 – Terceira geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	85
Tabela 9 – Mutantes filhos da terceira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	85
Tabela 10 – Quarta geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	86
Tabela 11 – Mutantes filhos da quarta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	86
Tabela 12 – Quinta geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	87
Tabela 13 – Mutantes filhos da quinta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	87
Tabela 14 – Sexta geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	88
Tabela 15 – Mutantes filhos da sexta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	88
Tabela 16 – Sétima geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	89
Tabela 17 – Mutantes filhos da sétima geração de AM. A Todos os valores de ΔG estão em quilocaloria por mol.	89

Tabela 18 – Oitava geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	90
Tabela 19 – Mutantes filhos da oitava geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	90
Tabela 20 – Nona geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	91
Tabela 21 – Mutantes filhos da nona geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	91
Tabela 22 – Décima geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	92
Tabela 23 – Mutantes filhos da décima geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	92
Tabela 24 – Décima primeira geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	93
Tabela 25 – Mutantes filhos da décima primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	93
Tabela 26 – Décima segunda geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	94
Tabela 27 – Primeira geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	95
Tabela 28 – Mutantes filhos da primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	95
Tabela 29 – Segunda geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	96
Tabela 30 – Mutantes filhos da segunda geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	96
Tabela 31 – Terceira geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	97
Tabela 32 – Mutantes filhos da terceira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	97
Tabela 33 – Quarta geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	98
Tabela 34 – Mutantes filhos da quarta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	98
Tabela 35 – Quinta geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	99
Tabela 36 – Mutantes filhos da quinta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	99

Tabela 37 – Sexta geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	100
Tabela 38 – Mutantes filhos da sexta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	100
Tabela 39 – Sétima geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	101
Tabela 40 – Mutantes filhos da sétima geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	101
Tabela 41 – Oitava geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	102
Tabela 42 – Mutantes filhos da oitava geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	102
Tabela 43 – Nona geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.	103
Tabela 44 – Primeira geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	104
Tabela 45 – Mutantes filhos da primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	104
Tabela 46 – Segunda geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	105
Tabela 47 – Mutantes filhos da segunda geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	105
Tabela 48 – Terceira geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	106
Tabela 49 – Mutantes filhos da terceira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	106
Tabela 50 – Quarta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	107
Tabela 51 – Mutantes filhos da quarta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	107
Tabela 52 – Quinta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	108
Tabela 53 – Mutantes filhos da quinta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	108
Tabela 54 – Sexta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	109
Tabela 55 – Mutantes filhos da sexta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	109

Tabela 56 – Sétima geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	110
Tabela 57 – Mutantes filhos da sétima geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	110
Tabela 58 – Oitava geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	111
Tabela 59 – Mutantes filhos da oitava geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	111
Tabela 60 – Nona geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	112
Tabela 61 – Mutantes filhos da nona geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	112
Tabela 62 – Décima geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	113
Tabela 63 – Mutantes filhos da Décima geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	113
Tabela 64 – Décima primeira geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	114
Tabela 65 – Mutantes filhos da Décima primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	114
Tabela 66 – Décima segunda geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	115
Tabela 67 – Mutantes filhos da Décima segunda geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	115
Tabela 68 – Décima terceira geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	116
Tabela 69 – Mutantes filhos da Décima terceira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	116
Tabela 70 – Décima quarta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	117
Tabela 71 – Mutantes filhos da Décima quarta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	117
Tabela 72 – Décima quinta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	118
Tabela 73 – Mutantes filhos da Décima quinta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.	118
Tabela 74 – Décima sexta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.	119

Lista de abreviaturas e siglas

AG	Algoritmo genético
AM	Anticorpos miméticos
CDR	Regiões determinantes complementares
EL	Energia livre de ligação
Fab	Fragmento de ligação ao antígeno
Fc	Região do fragmento cristalizável
HC	Hepatite C
MH	Modelagem por homologia
DM	Dinâmica molecular
MM-GBSA	Molecular Mechanics-Generalized Born Surface Area
RMSD	Raiz quadrada do desvio quadrático médio
RMSF	Raiz quadrada da flutuação quadrática média
VEGF	Fator de crescimento vascular
VEGFR	Receptor tirosina quinase do tipo VEGF-FMS

Lista de símbolos

α	Letra grega minúscula Alfa
β	Letra grega minúscula Beta
γ	Letra grega minúscula Gama
Δ	Letra grega maiúscula Delta
ϵ	Letra grega minúscula Epsilon
θ	Letra grega minúscula Theta
μ	Letra grega minúscula Mi
π	Letra grega minúscula Pi
ω	Letra grega minúscula Omega

Sumário

1	Introdução	20
2	Objetivos	25
3	Metodologia	26
3.1	<i>Softwares</i> usados	26
3.2	Campo de força	26
3.3	MM-GBSA	27
3.4	Construção do sistema base	28
3.5	Função de pontuação	29
3.6	Dinâmica molecular	32
3.7	Algoritmo Genético	32
3.8	Síntese e expressão	33
3.9	Dicroísmo circular	34
3.10	Deteção de anticorpos	34
4	Resultados	36
4.1	Função de pontuação	36
4.2	Algoritmo genético do sistema VEGF	39
4.2.1	Análise do desvio padrão mGB1-11	45
4.2.2	Decomposição de energia do mGB1-11	46
4.3	Algoritmo genético do sistema HC-E2	47
4.3.1	Análise do desvio padrão HC-mGB1-81	49
4.3.2	Decomposição de energia do mutante HCmGB1-81	50
4.4	Algoritmo genético do sistema RBD	51
4.4.1	Análise do desvio padrão SGB1-121	54
4.4.2	Análise de termos entrópicos	55
4.4.3	Decomposição de energia do mutante SGB1-121	56
4.4.4	Modelagem por homologia SGB1-121	57
4.4.5	Teste de ancoragem molecular	58
4.4.6	Teste de His-tag	59
4.4.7	Análise da estabilidade estrutural	60
4.4.8	Análise de estrutura secundária	61
4.4.9	Moléculas estruturais de água	62
5	Síntese e validação experimental	63
5.1	Análise de dicroísmo circular	64
5.2	Avaliação de afinidade antigênica	66
6	Observações gerais	67

7	Conclusão	70
	Referências	71
	APÊNDICE A – PARÂMETROS DA FUNÇÃO DE PONTUAÇÃO . .	78
	APÊNDICE B – VALIDAÇÃO DO AG	79
	APÊNDICE C – PROTOCOLOS EXPERIMENTAIS	80
	APÊNDICE D – ENERGIAS DO AG PARA OS MUTANTES MGB1 ANTI-FLT-1	83
	APÊNDICE E – ENERGIAS DO AG PARA OS MUTANTES MGB1 ANTI-HC-E2	95
	APÊNDICE F – ENERGIAS DO AG PARA OS MUTANTES SGB1 ANTI-RBD	104
	APÊNDICE G – ARQUIVOS DE TRAJETÓRIAS SGB1-121	120
	APÊNDICE H – CÓDIGOS USADOS PELO AG	121

1 Introdução

No cenário atual, com o surgimento de novas doenças e epidemias, novas estratégias são necessárias para otimizar o desenvolvimento de moléculas bioativas. Nesse contexto, as terapias baseadas em anticorpos são uma das abordagens mais bem-sucedidas (Lu *et al.*, 2020), devido à sua formidável capacidade de reconhecimento molecular. Essa propriedade é modulada pelas alças variáveis nas Regiões Determinantes de Complementaridade (CDR), localizadas no fragmento de ligação ao antígeno (Fab), dentro das cadeias leve e pesada dos anticorpos (Nelson; Cox, 2014). Ademais, os anticorpos têm a capacidade de ativar a resposta imune por meio de marcadores de opsonina presentes na região do fragmento cristalizável (Fc). Do ponto de vista farmacodinâmico, os anticorpos são o agente terapêutico ideal, capazes de neutralizar um único antígeno, evitando problemas de efeitos colaterais comuns em medicamentos, devido à baixa seletividade. No entanto, encontrar um anticorpo relacionado a cada doença é uma das tarefas mais desafiadoras (Scott; Wolchok; Old, 2012). Além disso, os altos custos de pesquisa e desenvolvimento tornam as terapias baseadas em anticorpos extremamente caras.

Nessa perspectiva, os pesquisadores buscam descobrir moléculas mais simples que possam reproduzir, ainda que parcialmente, as propriedades dos anticorpos, dentre as quais, destacam-se os anticorpos miméticos (AM). A principal característica de um AM é a existência de uma região constante, que estabiliza o enovelamento, e de regiões variáveis, que medeiam a capacidade de reconhecimento molecular (Simeon; Chen, 2017). Embora os AM sejam incapazes de estimular respostas imunes, eles podem reproduzir a capacidade de neutralização dos anticorpos, ligando-se fortemente aos antígenos. Devido à sua maior simplicidade e origem peptídica, os AM são muito mais fáceis de sintetizar e armazenar, não necessitando de uma rota sintética completamente diferente para cada nova molécula, como as moléculas de fármacos tradicionais (Bruce; Ta; Mcnaughton, 2016). Além disso, os AM se adaptam melhor a diferentes organismos e possuem maior penetrabilidade que os anticorpos convencionais.

Um dos casos de maior sucesso na engenharia de anticorpos miméticos são os *affibodies* baseados no domínio Z da proteína A, cuja estrutura é suficientemente estável para modular a capacidade de reconhecimento molecular através da randomização de 13 aminoácidos (Frejd; Kim, 2017). Analogamente aos *affibodies*, a estrutura do domínio B1 da proteína G estreptocócica (GB1) (PDB ID: 3GB1) é conhecida na literatura por sua alta estabilidade frente a mudanças de aminoácidos (Voet; Voet, 2013, p. 282). Devido às suas propriedades, o GB1 já serviu de base para o desenvolvimento de uma metaloproteína artificial (Bozkurt *et. al*, 2018) e como modelo para estudar a

compatibilidade estérica de ligação entre proteínas (Nerattini *et al.*, 2018).

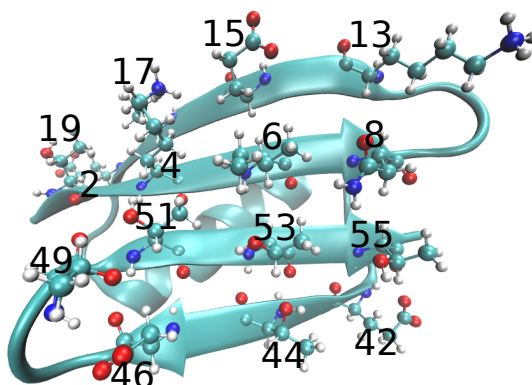


Figura 1 – O domínio GB1 é composto por 56 resíduos de aminoácidos que se enovelam em quatro fitas- β e uma α -hélice. As quatro fitas- β formam uma enorme área de superfície com pelo menos 15 aminoácidos (Em destaque na ilustração) que podem ser modificados para modelagem de um AM.

Fonte: elaborado pelo Autor (2024).

Como pode ser visto na Figura 1, o domínio GB1 é composto por 56 resíduos de aminoácidos, com uma estrutura tridimensional muito bem definida, que se enovela em quatro fitas- β e uma α -hélice. Essa estrutura confere ao domínio GB1 as características de uma proteína globular, onde as quatro fitas- β formam uma área superficial exposta na qual os resíduos 2, 4, 6, 8, 13, 15, 17, 19, 42, 44, 46, 49, 51, 53 e 55 podem ser modificados para mediar a capacidade de reconhecimento molecular (Espírito Santo; Feliciano, 2021b). Todas essas propriedades fazem do domínio GB1 um candidato ideal para uma prova de conceito no desenvolvimento de AM.

A escolha dos candidatos a antígeno para um AM deve levar em consideração o tamanho e a estabilidade molecular. Diante disso, o mapeamento de epítomos é considerado uma etapa crucial no desenvolvimento de anticorpos terapêuticos (Ladner, 2007), pois identifica as menores porções de antígenos capazes de gerar a resposta imune. Isso é resultado da capacidade desses epítomos de manter a integridade estrutural, mesmo quando dissociadas da estrutura principal da proteína.

Neste caso, o receptor tirosina quinase do tipo VEGF-FMS (VEGFR), um indutor angiogênico em organismos vivos, composto por sete domínios de imunoglobulina (Wiesmann *et al.*, 1997), é uma excelente opção de antígeno para o projeto de AM. Como pode ser visto pela resolução da estrutura cristalográfica (PDB ID: 1FLT), o segundo domínio do VEGFR mostrado na Figura 2A é responsável pela ligação ao fator de crescimento vascular (VEGF). A neutralização do VEGF foi um dos casos de maior sucesso na redução da mortalidade de pacientes com câncer em terapias com anticorpos monoclonais (Ellis; Hicklin, 2008). No entanto, são necessários dois

anticorpos para neutralizar cada VEGF. Portanto, o desenvolvimento de AM capazes de neutralizar o VEGFR teria um enorme potencial para auxiliar na terapia do câncer.

Ademais, as doenças virais, como a Hepatite C (HC) e a síndrome respiratória aguda grave (COVID-19), também são alvo de grande interesse no desenvolvimento de novas moléculas terapêuticas. Ressalta-se que, apesar dos avanços na medicina e introdução de novos fármacos, a HC ainda é uma das doenças virais mais graves em todo o mundo (Geller *et al.*, 2016). Isto se deve principalmente à grande mutabilidade deste vírus, o que dificulta a produção de vacinas eficazes. Neste caso, a utilização de epítomos, como o envelope viral HC-E2 (PDB ID: 5KZP), é suficiente para induzir uma resposta imunitária eficaz. Como pode ser visto na Figura 2B, o envelope HC-E2 pode ser sintetizado na forma do peptídeo defensina cíclica. Este epítopo é formado por apenas 15 resíduos de aminoácidos, os quais constituem duas fitas- β estabilizadas por ligações de dissulfeto em suas extremidades (Pierce *et al.*, (2017). Apesar de seu tamanho extremamente pequeno, ainda é grande o suficiente para ser reconhecido por anticorpos, tornando-o um potencial alvo para o projeto de AM.

Concomitantemente, a COVID-19, causada pelo coronavírus SARS-CoV-2, disseminou-se pelo mundo nos últimos anos, gerando milhares de vítimas (Chen; LI, 2020). Esse vírus tem uma rota de contaminação muito bem estudada, através do conhecimento vindo de outros vírus da mesma linhagem. Atualmente, sabe-se que esse vírus utiliza proteínas spike, presentes em sua superfície (Olukitibi *et al.*, 2023). Essa proteína é composta por duas subunidades S1 e S2 que ligam-se ao domínio peptidase do receptor ACE2 em células humanas. Esse receptor é responsável pela fusão das membranas celulares, atuando como porta de entrada do vírus nas células. O reconhecimento molecular ocorre quando a α -hélice do receptor ACE2 se liga ao domínio RBD da proteína spike (Cao *et al.*, 2020). Devido à sua importância na ligação vírus-célula, o epítopo RBD mostrado na Figura 2C representa um alvo promissor no desenvolvimento de anticorpos terapêuticos e vacinas contra o vírus. Sabe-se que a neutralização do domínio RBD por anticorpos bloqueia a capacidade do vírus de se ligar ao domínio ACE2 e, conseqüentemente, impede a sua replicação (Shuaib *et al.*, 2023). O conhecimento deste ciclo resulta que a maioria das vacinas desenvolvidas em todo o mundo são concebidas para estimular respostas imunitárias à proteína spike.

Na Figura 2 podemos ver ilustrações das estruturas dos antígenos propostos como possíveis alvos para o projeto de AM. Neste caso, é evidente que o VEGFR e o HC-E2 apresentam tamanhos significativamente reduzidos em comparação com o RBD. Nos três casos, os antígenos são formados apenas por resíduos de aminoácidos, sem modificações pós-traducionais, mas com presença de algumas ligações dissulfeto, que auxiliam os epítomos a manter sua estabilidade estrutural. Além disso, todos os modelos propostos têm suas estruturas muito bem definidas experimentalmente através

da cristalografia de raios-X, o que reduz a incerteza sobre possíveis conformações em modelos puramente teóricos.

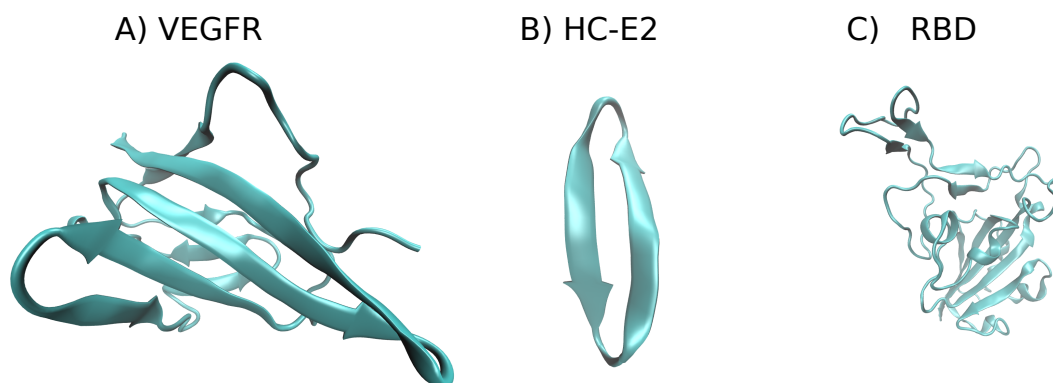


Figura 2 – Ilustração dos epítomos alvos dos AM. A) Receptor tirosina quinase (VEGFR) (PDB ID: 1FLT). B) Envelope viral HC-E2 (PDB ID: 5KZP). C) Domínio RBD (PDB ID: 7BZ5).

Fonte: elaborado pelo Autor (2024).

Nos casos especiais, quando existem bancos de dados de moléculas com atividade biológica conhecida, o projeto de novos AM pode ser baseado no estudo de mutações sitio-dirigidas (Valiente *et al.*, 2021) e transplantes de motivos estruturais (Kadonosono *et al.*, 2020), que visam a transferência e aprimoramento da capacidade de reconhecimento molecular. Essa tarefa pode ser realizada com uso de ferramentas computacionais. *Softwares*, como o Rosetta, usam bancos de dados estruturais para otimizar a capacidade de reconhecimento molecular de anticorpos (Adolf-Bryfogle *et al.*, 2018), mas a grande variabilidade das CDR tornam esse projeto um grande desafio. Na ausência de bancos de dados, esse problema pode ganhar dimensões imensuráveis. Apenas com 15 resíduos de aminoácidos da superfície do GB1 haveria um número gigantesco de combinações e regiões de reconhecimento molecular, impossibilitando a triagem de todas as possibilidades. No entanto, as interações intermoleculares que as cadeias laterais dos aminoácidos podem fazer são bem conhecidas. Além disso, o reconhecimento molecular não é um problema de uma única solução, pois existem muitos resíduos de aminoácidos com propriedades semelhantes (Gellman, 1997). Muitas combinações de aminoácidos podem ser soluções indistinguíveis de energia, capazes de desempenhar a mesma função de reconhecimento molecular.

Com base nesses princípios, as informações sobre a área superficial dos antígenos podem ser utilizadas para selecionar os resíduos com maior afinidade, modelando ligantes específicos para esta região, reduzindo o problema a dezenas de milhares de moléculas (Hunt *et al.*, 2022). Além disso, devido às inúmeras fases de desenvolvimento, desde a pesquisa experimental até ao produto final, onde moléculas com maior potencial avançam para fases posteriores, o projeto de moléculas bioativas pode ser visto essencialmente como um problema de evolução e seleção natural. Todas essas suposições podem servir de base para abordagens mais eficientes, que evitem a

necessidade de testar todas as combinações de aminoácidos possíveis, capazes de projetar racionalmente novas moléculas bioativas. Nessa situação, quando precisamos de soluções aproximadas para um problema multidimensional, podemos usar algoritmos genéticos para otimizar as propriedades de moléculas (Bozkurt et. al, 2014). Os AG são inspirados na biologia evolutiva, como hereditariedade, mutação, seleção natural e crossing over. *Softwares*, como o AutoGrow4 (Spiegel; Durrant, 2020), são capazes de usar a técnica de AG, em conjunto com modelos gerados por funções de pontuação de ancoragem molecular, para projetar novas moléculas. Contudo, essa metodologia é limitada a um banco de dados contendo algumas dezenas de pequenas moléculas. Além disso, ferramentas de ancoragem molecular geram modelos naturalmente direcionados para regiões com melhor afinidade.

Tendo em vista as limitações das técnicas usadas atualmente, uma nova abordagem de AG, que faça uso da dinâmica molecular (DM), poderia abrir um novo horizonte no planejamento de AM. A técnica de DM baseia-se na resolução numérica das equações do movimento de um sistema de moléculas (Hollingsworth; Dror, 2018). Essa abordagem permite explorar melhor o espectro conformacional das moléculas, revelando as conformações prováveis, onde o reconhecimento molecular pode ocorrer espontaneamente dentro do tempo de simulação. Aliado a DM, podemos utilizar o método *Molecular Mechanics-Generalized Born Surface Area* (MM-GBSA), popular para estimar a energia livre de ligação (EL). Esse método tem a capacidade de priorizar variedades de soluções no projeto de moléculas bioativas, atuando como um filtro de triagem virtual para bancos de dados muito grandes (Gupta; Chaudhary; Aparoy, 2018), facilitando estudos de descoberta de novos fármacos.

Dentro do contexto geral, este estudo procurou investigar uma nova abordagem computacional que combina DM e MM-GBSA em um algoritmo de busca e otimização, visando transformar o GB1 em um AM para os antígenos selecionados (Espírito Santo; Feliciano, 2021b). Como os AM gerados nesta abordagem são o resultado de um modelo baseado em uma representação simplificada, com várias aproximações usadas para viabilizar simulações de DM, eles precisam ser submetidos à validação experimental. Esta validação, baseada em testes de expressão gênica, análise estrutural e afinidade antigênica, é necessária para comprovar as propriedades para as quais foram concebidos. A proposta desenvolvida neste estudo visa repensar a aplicabilidade de técnicas amplamente utilizadas na pesquisa acadêmica, redefinindo-as não apenas como ferramentas para analisar sistemas biológicos mais complexos, mas como agentes ativos em sua transformação. Demonstra-se que essa transformação pode ser realizada através da reprogramação de funções biológicas, permitindo, assim, a criação de novos produtos biotecnológicos.

2 Objetivos

O principal objetivo deste estudo é desenvolver uma nova abordagem computacional para o desenho de anticorpos miméticos. Essa abordagem visa organizar os protocolos comumente utilizados em simulações moleculares em um AG, com a capacidade de aprimorar o reconhecimento molecular dos anticorpos miméticos. Dentro desse objetivo, estão os seguintes objetivos específicos:

- a) elaborar uma base peptídica estável para o desenho de um anticorpo mimético;
- b) desenvolver um algoritmo genético baseado em dinâmica molecular;
- c) otimizar a capacidade de reconhecimento molecular dos anticorpos miméticos;
- d) realizar a validação experimental de um dos anticorpos miméticos.

3 Metodologia

3.1 *Softwares* usados

No projeto, as ferramentas de visualização e edição molecular PyMOL (Delano, 2002), UCSF Chimera (Pettersen *et al.*, 2004) e VMD (Humphrey; Dalke; Schulten, 1996) foram utilizadas em conjunto com o *Software* de Dinâmica Molecular Amber 2018 (Case *et al.*, 2018). Os códigos auxiliares foram escritos em linguagem Shell, C e C++, e estão disponíveis no APÊNDICE H.

3.2 Campo de força

A DM é um método de simulação computacional que usa os princípios da mecânica clássica para descrever os movimentos dos átomos em um sistema modelo (Hollingsworth, Dror, 2018). Esses movimentos são descritos por meio de um campo de força que fornece informações sobre o comportamento dinâmico microscópico do sistema simulado. A forma básica do campo de força do Amber pode ser vista na equação 1 (Maier *et al.*, 2015):

$$\begin{aligned}
 V_{\vec{r}} = & \sum_{Ligado} k_r (r - r_0)^2 + \sum_{Angulo} k_{\theta} (\theta - \theta_0)^2 + \\
 & \sum_{Diedral} \sum_{i=1}^4 \frac{V_i}{2} (1 + \cos(i\phi - \phi_i)) + \\
 & \sum_{i>j} \left[C \left(\frac{A_{i,j}}{r_{i,j}^{12}} + \frac{B_{i,j}}{r_{i,j}^6} \right) + \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{i,j}} \right]
 \end{aligned} \tag{1}$$

A energia de ligação química entre dois átomos V_{Ligado} e a energia potencial angular V_{Angulo} são representadas pela equação clássica da Lei de Hooke. As energias relativas aos ângulos de torção entre os planos das ligações químicas são calculadas usando a soma das energias representadas no potencial diédrico $V_{diedral}$. As energias das interações entre átomos não ligados são calculadas usando as equações para os potenciais de Lennard-Jones e Coulomb (Espírito Santo; Feliciano, 2021a). As energias são calculadas numericamente a cada passo de tempo, sendo essa a parte mais demorada da simulação.

3.3 MM-GBSA

A EL é uma das grandezas termodinâmicas que permitem estimar a afinidade entre biomoléculas. Dentre os métodos computacionais para calcular a EL, o MM-GBSA ocupa um lugar de destaque (Genheden; Ryde, 2015). Esse método usa o fato da EL ser uma função de estado, o que permite fazer um caminho termodinâmico alternativo através do vácuo, simplificando os cálculos. A EL ($\Delta G_{ligacao}$) é calculada pela diferença entre a energia do conjugado receptor-ligante (ΔG_{RL}) e suas energias individuais (ΔG_R) e (ΔG_L) como podemos ver na equação 2:

$$\begin{aligned}
 \Delta G_{ligacao} &= \Delta G_{RL} - \Delta G_R - \Delta G_L \\
 \Delta G_i &= E_{mm} + \Delta G_{solv} - TS \\
 \Delta G_{Solv,Apolar} &\propto \gamma SASA + \beta \\
 G_{Solv,Polar} &= - \left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_{sol}} \right) \sum_{i,j} \frac{q_i q_j}{F_{GB}} \\
 F_{GB} &= \sqrt{r_{i,j}^2 + \alpha_{i,j}^2 + \exp\left(\frac{r_{i,j}^2}{4\alpha_{i,j}^2}\right)}
 \end{aligned} \tag{2}$$

No MM-GBSA a energia livre (ΔG_i) é decomposta em componentes de energia mecânica molecular (E_{MM}), energia de solvatação (ΔG_{solv}) e energia de componentes entrópicos ($T\Delta S$). A E_{MM} é obtida através do campo de força usado na DM. A ΔG_{solv} pode ser decomposta em energia de solvatação apolar ($\Delta G_{solv,apolar}$) e energia de solvatação polar ($\Delta G_{solv,polar}$). A $\Delta G_{solv,apolar}$ representa o trabalho necessário para gerar a cavidade no solvente onde o sistema está localizado, onde γ é a tensão superficial e β um termo de correção.

A $\Delta G_{solv,polar}$ pode ser interpretado como o potencial resultante da blindagem que os íons do meio exercem sobre a superfície do sistema (Espírito Santo; Feliciano, 2021a). O modelo GB busca maior eficiência ao descrever os átomos como esferas carregadas, cuja constante dielétrica interna da proteína ϵ_{in} é menor que a ϵ_{sol} externa. Além disso, utiliza apenas as cargas dos átomos q_i , as distâncias entre eles $r_{i,j}$ e sua média geométrica $\alpha_{i,j}$, não necessitando de métodos numéricos computacionalmente caros para resolver a equação de Poisson-Boltzmann (Wang *et al.*, 2019).

O termo entrópico vibracional é calculado a partir da diagonalização da matriz hessiana do sistema. Esta matriz representa as segundas derivadas da energia potencial em relação às coordenadas dos átomos no sistema. Quando diagonalizada, obtemos os autovalores e autovetores, os quais são usados para calcular a entropia dos modos normais de vibração (Genheden *et al.*, 2012).

3.4 Construção do sistema base

Ao projetar AM queremos otimizar o reconhecimento molecular entre duas superfícies que originalmente não têm afinidade e seletividade. Neste caso, não faz sentido utilizar qualquer tipo de técnica de ancoragem molecular (Arya; Kaur, 2022). Usando o software UCSF Chimera, no modo de visualização surface, realizamos manualmente uma ancoragem molecular entre as superfícies do peptídeo GB1 (PDB ID: 3GB1) e da região de interesse do antígeno alvo, procurando evitar espaços entre as áreas de contato. Existe um consenso de que a estrutura cristalográfica de uma proteína não corresponde perfeitamente à conformação em solução, devido ao próprio processo de cristalização (Conway *et al.*, 2013).

Desta forma, realizamos a minimização de energia com um campo de força de DM para relaxar as estruturas, possibilitando ao modelo obter condições de simulação conformacional mais realistas. Assim sendo, a estrutura resultante é inserida em uma caixa de simulação cúbica, com 10 nm de borda e solvatada em água TIP3P. Em seguida, é realizada uma minimização da energia, usando o campo de força ff14SB do software Amber (Maier *et al.*, 2015), através do algoritmo de descida mais íngreme, para relaxar a geometria dos peptídeos e evitar efeitos estéricos (Espírito Santo; Feliciano, 2021b). Após a minimização, as estruturas dos peptídeos são convertidas novamente para o formato PDB.

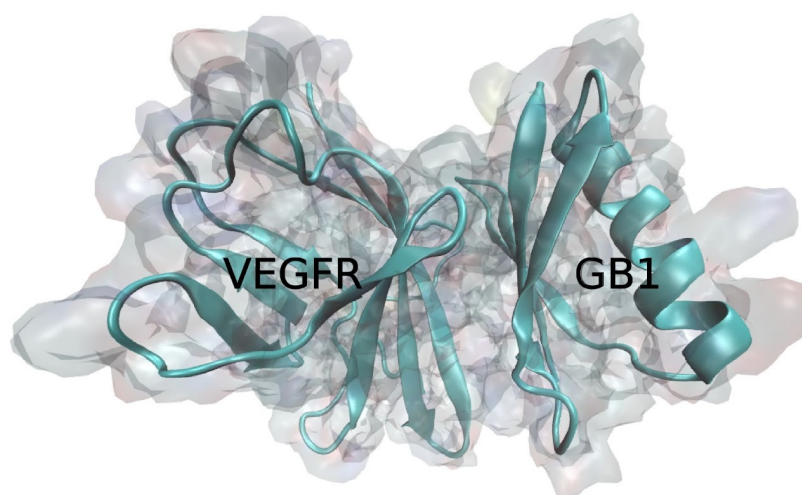


Figura 3 – Ilustração do complexo formado pelas estruturas dos peptídeos GB1 e VEGFR após a minimização da energia do campo de força Amber.

Fonte: elaborado pelo Autor (2024).

Na Figura 3 podemos ver a Ilustração do complexo formado pelas estruturas dos peptídeos GB1 e VEGFR após a minimização da energia do campo de força Amber. Nota-se que as estruturas dos dois peptídeos permanecem enoveladas corretamente

apesar da ancoragem molecular. Esse protocolo permite que as duas superfícies tenham a maior área de contato possível, mesmo que não tenham afinidade. A distância final entre as superfícies é dada pela minimização da energia, variando conforme o tamanho das cadeias laterais dos aminoácidos. Essa forma de ancoragem evita a desnaturação da estrutura secundária dos peptídeos, permitindo que as cadeias laterais se ajustem quando em contato com outro peptídeo.

A estrutura relaxada é usada somente pela função de pontuação. Para ser submetida a etapa de DM, a estrutura relaxada é espaçada com o Chimera, em aproximadamente 1 nm, na direção do vetor normal à superfície de contato. Assim, além de haver mais espaço para todos os tipos de resíduos, garantimos que o reconhecimento molecular possa ocorrer de forma espontânea durante a DM.

3.5 Função de pontuação

Os algoritmos genéticos normalmente usam populações iniciais criadas aleatoriamente, que são otimizadas a partir dessa condição inicial. No caso de reconhecimento molecular entre proteínas, essa estratégia seria computacionalmente proibitiva. No entanto, é possível usar as interações intermoleculares da superfície do antígeno para pré-selecionar configurações mais favoráveis (Chauhan; Pantazes, 2022). Essa tarefa é realizada por meio de um shell script, capaz de usar o software PyMOL dentro uma função de pontuação simples (APÊNDICE H), usando parâmetros empíricos de termos energéticos (Huang; Grinter; Zou, 2010), como ligação de hidrogênio, dipolo-dipolo e Lennard-Jones (APÊNDICE A - Tabela 1).

Os valores de energia de interação intermolecular podem variar na literatura em ordens de grandeza, dependendo da condição experimental. Para a energia de ligação de hidrogênio foi usado o valor de $1,5 \text{ kcal.mol}^{-1}$, que é a energia média para peptídeos em solução (Sheu et al., 2003). A energia usada no termo de Lennard-Jones foi de $0,17 \text{ kcal.mol}^{-1}$, um valor aproximado dos átomos orgânicos no campo de força Amber (Su et al., 2014). A energia do termo de repulsão polar, usado para penalizar átomos de polaridade similar, foi de $0,95 \text{ kcal.mol}^{-1}$, um valor típico para interação dipolo-dipolo (Whitten, 2014, p. 450).

O script usa o arquivo PDB do sistema relaxado para fazer uma varredura no espaço a cada $0,5 \text{ Å}$ ao redor do antígeno alvo. Todos os átomos da estrutura PDB são analisados e, quando um átomo é encontrado na superfície, a posição xyz e um átomo rotulado são armazenados em um arquivo molde de superfície no formato PDB, como vemos na Figura 4:

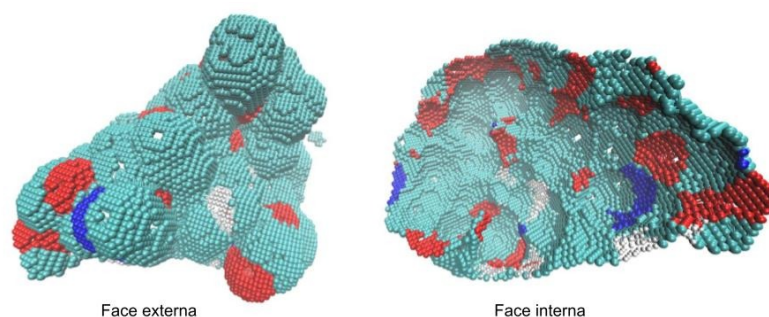


Figura 4 – Ilustração da estrutura do molde PDB da superfície do VEGFR. As regiões em vermelho e azul são aceptoras de ligação de hidrogênio. As regiões em branco são doadoras de ligação de hidrogênio. As regiões em ciano são hidrofóbicas.

Fonte: elaborado pelo Autor (2024).

Na Figura 4, os átomos marcados em vermelho e azul são aceptores de ligação de hidrogênio, brancos são doadores de ligação de hidrogênio e ciano são hidrofóbicos. No arquivo molde da superfície do VEGFR, podemos ver várias regiões capazes de formar diferentes interações intermoleculares. Esse resultado mostra que para haver reconhecimento molecular na região é necessária alta seletividade. Dessa forma, utilizando os arquivos das estruturas PDB do GB1 e do molde mostrados na Figura 5, podemos executar o script shell do APÊNDICE H - Figura 34 para realizar a mutagênese no PyMOL em cada uma das posições dos resíduos 2, 4, 6, 8, 13, 15, 17, 19, 42, 44, 46, 49, 51, 53 e 55 do GB1.

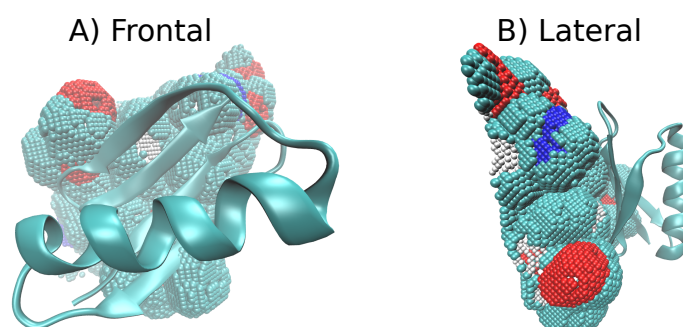


Figura 5 – Ilustração do complexo formado pela estrutura do GB1 e o molde VEGFR utilizado pela função de pontuação.

Fonte: elaborado pelo Autor (2024).

A mutagênese é realizada testando os rotâmeros dos aminoácidos essenciais, que estão presentes no banco de dados do PyMOL. É importante destacar a importância da função mutagênese do PyMOL, pois além de permitir a escolha do rotâmero desejado, ela também pode escolher automaticamente, com base em critérios energéticos, evitando problema de impedimento estérico. Além de permitir que todas essas funções sejam executadas milhares de vezes por meio de shell script.

Embora todos os 20 aminoácidos essenciais possam participar do reconheci-

mento molecular (Ramaraj *et al.*, 2012), optamos por excluir a possibilidade de haver fenilalanina na região variável do AM, para evitar problemas de enovelamento, uma vez que ela participa do núcleo hidrofóbico de GB1. Todos os átomos de cada rotâmero têm sua distância calculada em relação a todos os átomos do molde. Os átomos que estão dentro do raio de corte (0,5 Å) recebem uma pontuação. As funções de pontuação empíricas estimam a afinidade de ligação com base em um conjunto de termos de energia dados pela equação 3:

$$\Delta V = \sum_i \omega_i \Delta V_i \quad (3)$$

Onde ω_i representa o peso, dado pelo número de átomos rotulados dentro do raio de corte, e ΔV_i representa o módulo do termo de energia potencial de interação. Com a função de pontuação, podemos ordenar os resíduos de aminoácidos em todas as posições (APÊNDICE A -Tabela 2), informação usada para gerar uma população inicial de mutantes (Espírito Santo; Feliciano, 2021b). Neste caso, o primeiro mutante é gerado pelos aminoácidos com melhor pontuação em cada uma das posições, o segundo mutante é gerado pelos aminoácidos com a segunda melhor pontuação e assim sucessivamente. O mutante de melhor pontuação na função de pontuação para o VEGFR é mostrado na Figura 6:

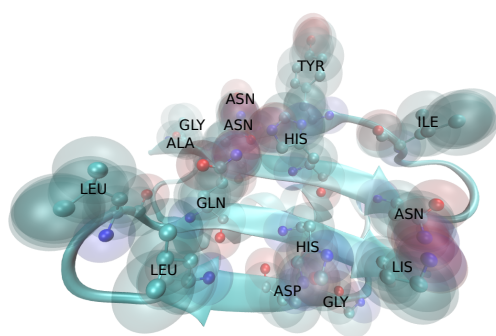


Figura 6 – Ilustração da estrutura do mutante mGB1-1 melhor classificado na função de pontuação do rotâmeros para VEGFR. Os resíduos são projetados com seus rotâmeros na direção de maior pontuação.

Fonte: elaborado pelo Autor (2024).

A Figura 6 mostra os resíduos de aminoácidos do mutante mGB1-1 com os seus rotâmeros melhor classificados em cada posição pela função de pontuação. As esferas translúcidas são átomos do molde VEGFR, as esferas opacas são átomos dos rotâmeros. Como pode ser observado, os resíduos se ajustam conforme as propriedades desejadas nos grupos químicos das cadeias laterais. Grupos apolares são favorecidos na região ciano, grupos doadores e aceptores de ligações de hidrogênio nas regiões vermelha e azul.

3.6 Dinâmica molecular

Em todos os modelos para AM propostos, o sistema base com espaçamento, gerado anteriormente, é modificado nas posições dos resíduos 2, 4, 6, 8, 13, 15, 17, 19, 42, 44, 46, 49, 51, 53 e 55 de GB1 para formar os arquivos da população inicial dos mutantes. A conformação inicial dos rotâmeros é sempre escolhida automaticamente pelo PyMOL, a fim de minimizar a energia. Como o sistema contém apenas resíduos de aminoácidos, usamos o campo de força ff14SB, com passo de tempo de 2fs.

Todas as DMs foram realizadas em GPU usando pmemd.cuda implementado em Amber 18. Cada sistema foi solvatado com modelo de água TIP3P em um caixa cúbica de simulação, com 1 nm de distância entre o sistema e a borda da caixa. O método de malha de partículas de Ewald foi usado para tratar interações não ligadas de longo alcance. Em interações de curto alcance, foi utilizado um raio de corte padrão do Amber de 8 Å. O campo de força clássico considera um estado de protonação fixo da proteína referente ao pH 7. Cada um dos sistemas foi submetido ao processo de minimização de energia com o algoritmo de descida mais íngreme. A temperatura do sistema foi elevada para 300 K usando termostato dinâmico Langevin. A pressão foi controlada a 1 bar usando o barostato de Berendsen.

Todas as DMs foram realizadas em duas etapas. Na primeira etapa, realizamos uma DM por um período de 90 ns, onde foram salvos os parâmetros de temperatura, pressão e posição do primeiro e último *frame*. A segunda etapa da DM foi então continuada de 90 ns para 100 ns onde 5000 *frames* foram salvos. A EL foi calculada usando o script (MMPBSA.py) do AmberTools16, usando concentração de sal de 0.1 e constante dielétrica da água. Neste projeto, utilizamos uma máquina equipada com processador Intel Xeon Gold 6130 e placa de vídeo NVIDIA Geforce GTX 1080.

3.7 Algoritmo Genético

O AG é iniciado com a execução das DMs da população inicial gerada pela função de pontuação. As condições iniciais dos mutantes são sempre as mesmas, determinadas pela distância do sistema base com um espaçamento uniforme gerado anteriormente. Em seguida, a população é submetida à seleção natural com base na EL, calculada pelo MM-GBSA. Os mutantes com a energia desfavorável, que sofrem desnaturação ou se ligam em outras regiões, são eliminados no processo. Em seguida, o tamanho da população é restaurado pelo código de permutação mostrado no APÊNDICE H, cruzando os mutantes sobreviventes de maneira semelhante ao crossing over, onde os resíduos de aminoácidos seriam equivalentes aos genes.

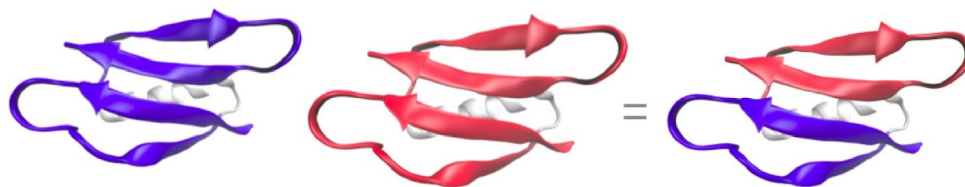


Figura 7 – Ilustração do cruzamento inicial realizado pelo AG. Os dois pares de fitas- β de dois mutantes são trocados resultando em dois novos mutantes filhos.

Fonte: elaborado pelo Autor (2024).

Como visto na Figura 7, a primeira permutação é realizada cruzando dois pares de fitas- β . O código então pesquisa no banco de dados para ver se a energia desse mutante já foi calculada. Em caso de repetição, a segunda permutação é realizada entre as quatro fitas- β . Em um novo caso de repetição, da terceira permutação em diante, são realizadas trocas em resíduos individuais, da direção N-terminal para C-terminal. Quando esgotada a possibilidade de geração de novos mutantes por cruzamento, o recurso de mutação é utilizado para aumentar a variabilidade populacional. A escolha do resíduo mutado não deve ser aleatória para evitar custos computacionais desnecessários. Assim, escolhemos entre os aminoácidos que apresentam a melhor pontuação na posição em que a mutação é realizada. Este processo de seleção e replicação é repetido até que toda uma população de novas moléculas, com características desejáveis, seja obtida no processo de convergência evolutiva do AG.

3.8 Síntese e expressão

Devido ao tamanho de 56 resíduos, a forma mais apropriada de expressar os AM é usar o vetor para clonagem. A linhagem de expressão *E. Coli* BL21 Rosetta foi adquirida pela Thermo Fisher Scientific. Os genes sintéticos foram adquiridos da empresa GenOne Brasil (Rio de Janeiro, RJ), já devidamente clonados no vetor de expressão pET-28a (+), contendo os genes mostrados nas páginas posteriores, localizados entre os sítios de restrição NdeI e XhoI. Além disso, o plasmídeo contém o promotor de RNA polimerase T7 e o operon lac que reprime a expressão não induzida por isopropil-b-D-galactosídeo (IPTG). Essa etapa da pesquisa foi realizada em parceria com a equipe de pesquisadores da Faculdade de Ciências Farmacêuticas da UNESP (FCFAR): Aline Reis, Anderson Amendola Pinheiro e o Prof. Dr. Paulo Inácio da Costa. Protocolos experimentais utilizados em testes de expressão, purificação e afinidade antigênica podem ser encontrados no APÊNDICE C.

3.9 Dicroísmo circular

O dicroísmo circular (CD) é uma forma de espectroscopia amplamente utilizada no estudo de estruturas secundárias de proteínas (Fasman, 1996). Baseia-se na diferente absorção de luz polarizada no sentido horário e anti-horário (Woody, 1995). A elipticidade molar (θ_m) é obtida a partir das seguintes equações:

$$\begin{aligned} A &= \ln \frac{I_0}{I} = \epsilon cl \\ \Delta\epsilon &= \epsilon_{esquerda} - \epsilon_{Direita} = \frac{A_{esquerda} - A_{Direita}}{cl} \\ \theta_m &= 100 \frac{\theta}{cl} = 2398 \Delta\epsilon \end{aligned} \quad (4)$$

Onde θ_m é a diferença entre o coeficiente de extinção molar da luz polarizada da esquerda para a direita $\Delta\epsilon$, que depende da absorção da luz A , a concentração da espécie c , a distância percorrida l e a intensidade da luz I .

3.10 Detecção de anticorpos

O kit de detecção de anticorpos neutralizantes (cPass) é baseado em ensaio imunoenzimático (ELISA). Destina-se à avaliação qualitativa de anticorpos neutralizantes (Green; Sambrook, 2012). Devido a esta função, pode ser utilizado como validação experimental da capacidade neutralizante de anticorpos miméticos. O valor de corte da percentagem de inibição $\%_{Sinal}$ para anticorpos neutralizantes é de 30%, para valores abaixo deste considera-se que não há neutralização, para valores superiores considera-se que o anticorpo é capaz de neutralizar o antígeno. A percentagem de inibição é obtida usando a equação abaixo:

$$\%_{Sinal} = \left[1 - \frac{OD_{Amostra}}{OD_{Controle}} \right] \times 100\% \quad (4)$$

Onde $OD_{Amostra}$ é a densidade óptica da amostra e $OD_{Controle}$ é a densidade óptica do controle negativo.

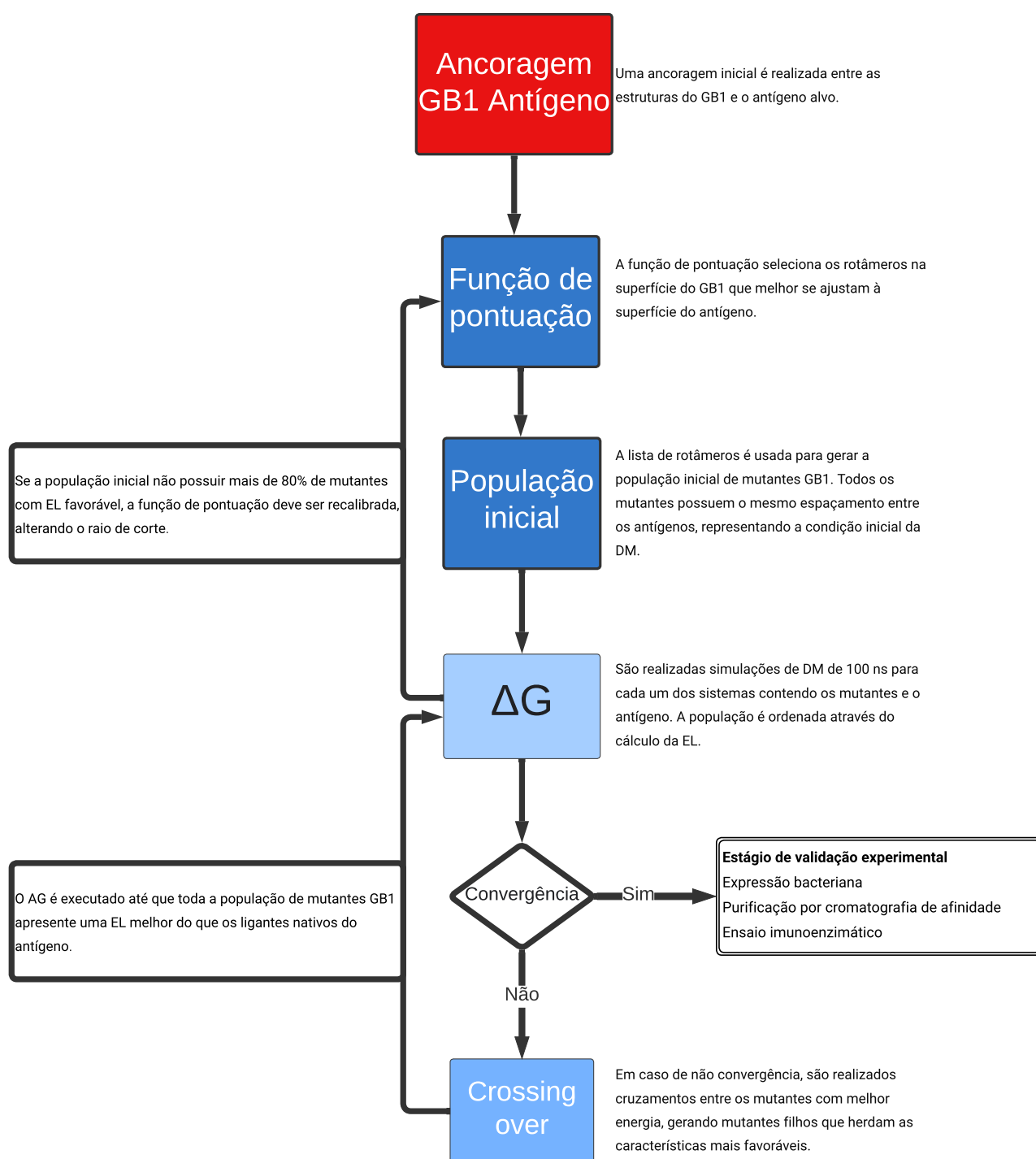


Figura 8 – Fluxograma representando a metodologia completa utilizada no AG.

Fonte: elaborado pelo Autor (2024).

4 Resultados

4.1 Função de pontuação

A função de pontuação previamente definida foi aplicada aos antígenos VEGFR, HC-E2 e RBD, como mencionado na introdução. É importante ressaltar que não se espera que a população inicial, gerada por meio de uma função de pontuação tão simples, apresente afinidade pela região de reconhecimento molecular. Seu objetivo é apenas evitar problemas que surgiriam ao gerar uma população inicial de forma aleatória, como a adição de grupos hidrofílicos em regiões hidrofóbicas.

Para facilitar a compreensão, os resultados da função de pontuação para os três sistemas estudados foram ilustrados graficamente. Em cada posição selecionada, um gráfico de barras mostra a classificação dos resíduos nessa posição. Quanto maior o valor da pontuação, melhor será a capacidade do resíduo em ocupar aquela posição. Nos gráficos a seguir, as colunas representam os aminoácidos mais bem classificados em cada uma das posições do GB1 para os três sistemas.

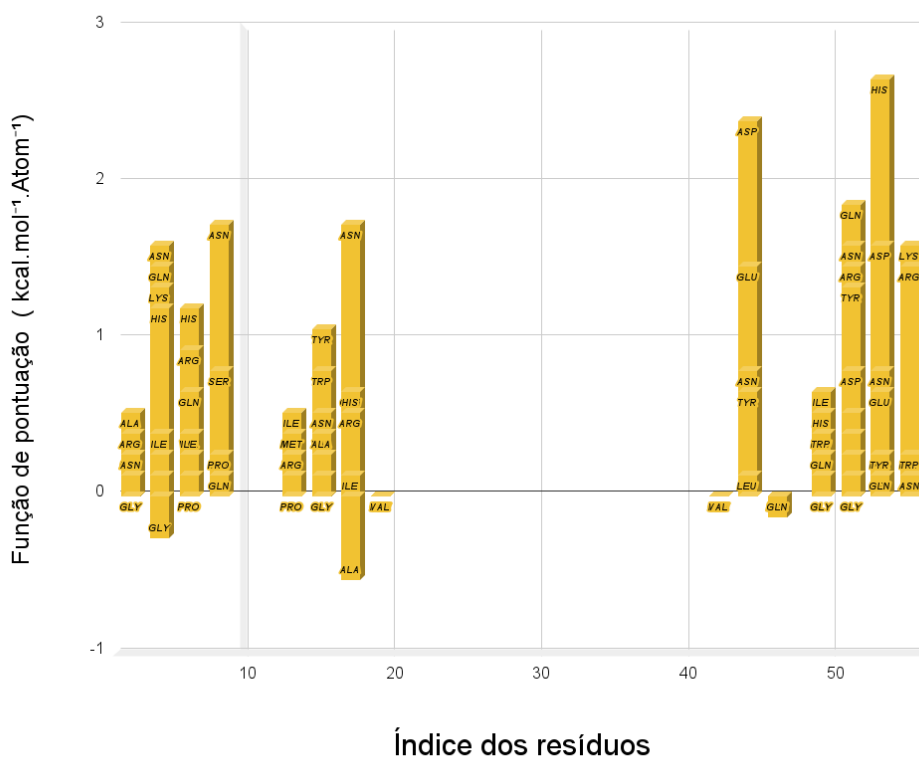


Figura 9 – Representação da ordenação dos resíduos conforme a função de pontuação para o antígeno VEGFR.

Fonte: elaborado pelo Autor (2024).

Na Figura 9, no resultado da função de pontuação do antígeno VEGFR, podemos observar um domínio de valores de energia positivos em quase todas as posições, indicando que existem muitos resíduos com interações favoráveis. Observa-se que, nas posições 19 e 42, nenhum dos rotâmeros consegue pontuar devido à grande distância do molde VEGFR. Nas regiões polares, os resíduos com maior pontuação geralmente são aqueles que podem atuar como doadores ou aceptores de ligações de hidrogênio. Nas regiões apolares, o que determina a pontuação é a capacidade do resíduo de ocupar melhor o espaço. Além disso, existem resíduos com pontuações negativas, indicando uma penalidade elevada para interações desfavoráveis.

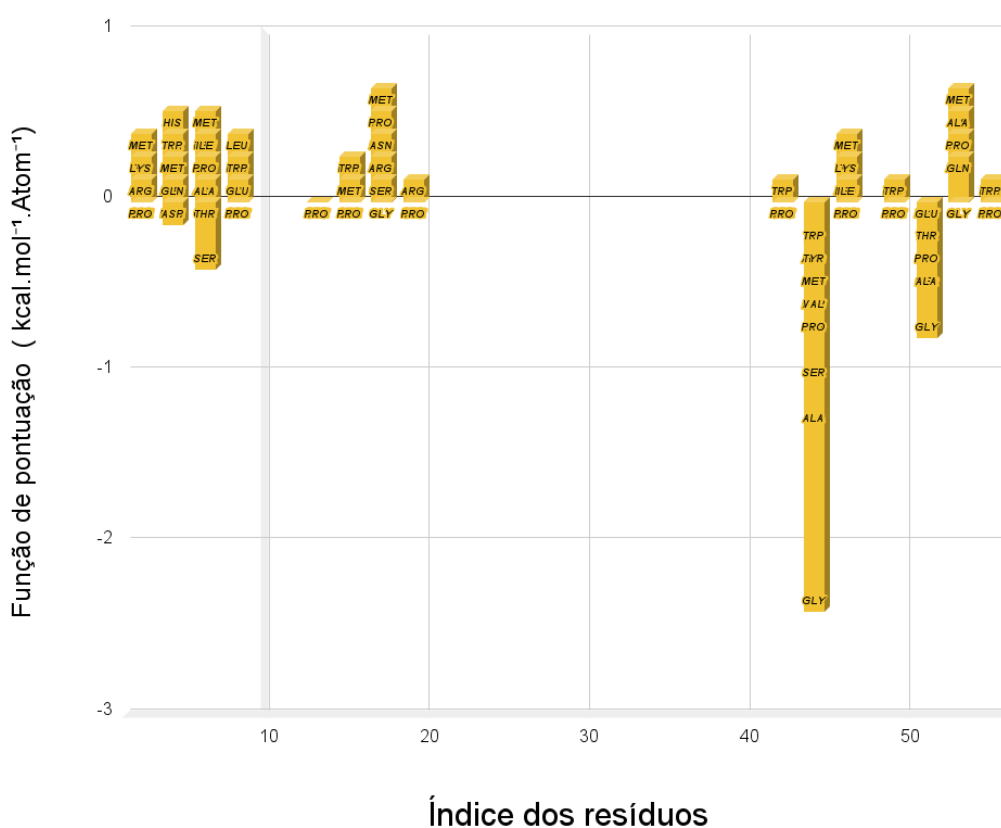


Figura 10 – Representação da ordenação dos resíduos conforme a função de pontuação para o antígeno HC-E2.

Fonte: elaborado pelo Autor (2024).

Na Figura 10, no resultado da função de pontuação para antígeno HC-E2, notamos que não há muita separação de energia entre os resíduos. Este resultado pode ser explicado pelo pequeno tamanho do antígeno, que é capaz de gerar um número limitado de interações com o AM. A maior separação de energia ocorre na posição 44, onde os valores negativos de energia indicam uma alta penalidade devido a interações desfavoráveis.

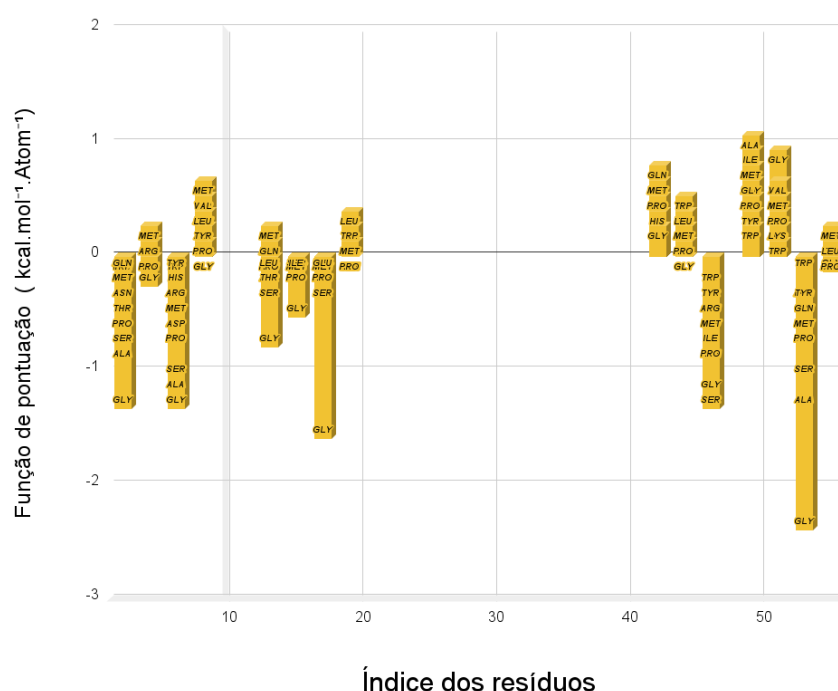


Figura 11 – Representação da ordenação dos resíduos conforme a função de pontuação para o antígeno RBD.

Fonte: elaborado pelo Autor (2024).

Na Figura 11, o resultado da função de pontuação para o antígeno RBD mostra um domínio de valores negativos, indicando uma predominância de penalidades em quase todas as posições. No entanto, a região superficial do antígeno possibilita diversos tipos de interações, favorecendo uma boa separação entre os resíduos.

Esses resultados mostram que a função de pontuação, definida por parâmetros empíricos, é um método plausível para gerar uma população inicial. Sua aplicação pode ser interpretada como um mecanismo de direcionamento da população inicial de AM para as imediações de um mínimo local, colaborando com o processo de otimização do AG.

A EL das populações iniciais é calculada com o MM-GBSA e utilizada como principal critério de calibração para a função de pontuação. Uma grande porcentagem de mutantes de primeira geração com EL favorável é o objetivo desejado. Caso contrário, a função de pontuação pode ser recalibrada variando o raio de corte. O AG inicia sua tarefa atuando sobre uma população inicial de AM que já possui alguma afinidade pelo antígeno, evitando a necessidade de escolha entre 20^{15} possibilidades de permutação.

4.2 Algoritmo genético do sistema VEGF

A etapa de execução do AG, conforme explicitado anteriormente, utiliza um sistema de base inicial, como pode ser visto na Figura 12B, formado pela estrutura PDB original do antígeno de interesse VEGFR, ancorado no peptídeo GB1, servindo como condição inicial onde todos os mutantes são gerados. Além disso, essa abordagem nos permite utilizar a estrutura cristalográfica do antígeno VEGFR com seu ligante natural VEGF, como pode ser visto na Figura 12A, para calcular a EL que serve como critério de convergência para o AG.

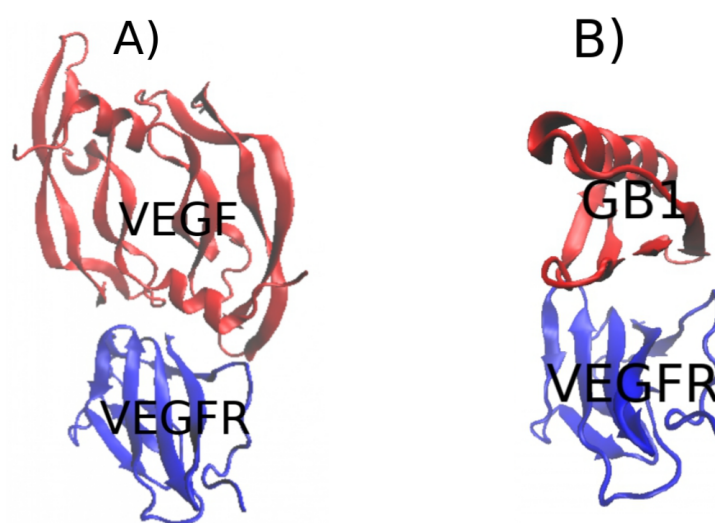


Figura 12 – Ilustração mostrando o sistema base para a DM no projeto de AM. A) Sistema formado pelo VEGFR e o VEGF. B) Sistema base formado pelo GB1 e o VEGFR.

Fonte: elaborado pelo Autor (2024).

A partir dessas estruturas iniciais, testamos o protocolo de AG previamente definido. As DMs de 100 ns foram executadas na população inicial, gerada pela função de pontuação, formada por 10 mutantes (mGB1) (APÊNDICE A - Tabela 2). O método MM-GBSA foi aplicado aos 10 ns finais para estimar a EL de cada mutante. Esse intervalo se justifica porque, diferentemente de moléculas pequenas, que podem ter uma grande número de conformações no sítio de ligação, os anticorpos são estruturas maiores e estáveis, que reconhecem o antígeno em poucas conformações. Além disso, as sequências de resíduos favoráveis à ligação são transmitidas para gerações futuras de AM, sendo amostradas dezenas de vezes no AG, o que equivale a uma simulação de DM de milhares de nanossegundos.

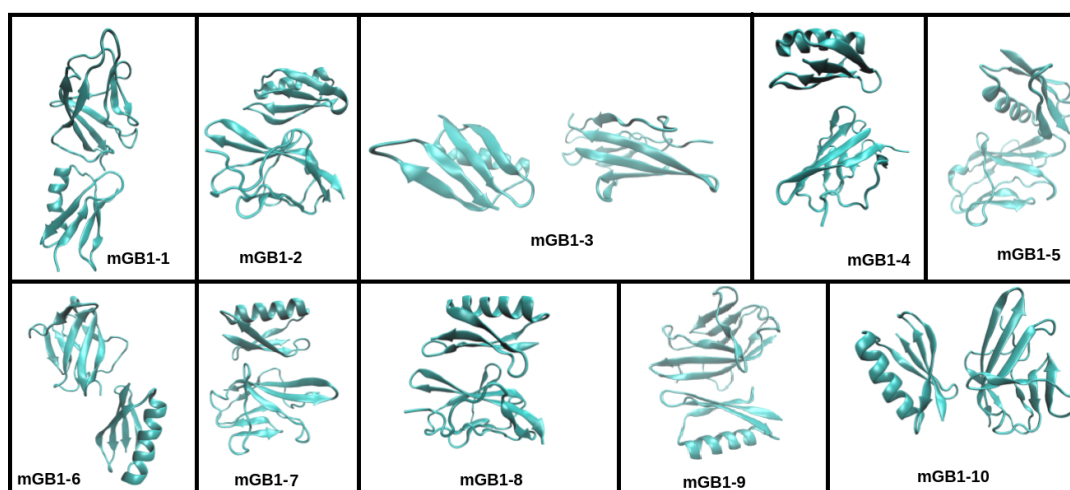


Figura 13 – Ilustração da conformação final dos mutantes mGB1 da população inicial do AG após 100ns em cada DM.

Fonte: elaborado pelo Autor (2024).

Na Figura 13, podemos ver o frame final dos mutantes da população inicial após 100 ns de DM. Nota-se que as estruturas mGB1-2, mGB1-7, mGB1-8 e mGB1-9, que possuem afinidade maior pela região, permanecem ligadas ao antígeno durante todo o tempo de simulação. De maneira oposta, nas estruturas mGB1-1, mGB1-3, mGB1-4, mGB1-5 e mGB1-6, ocorre a separação dos complexos, mesmo com valores negativos para a EL.

Esse resultado é explicado porque, para aumentar a eficiência, o método MM-GBSA ignora alguns termos entrópicos, que seriam cancelados ao fazermos a comparação de $\Delta\Delta G$. Este resultado revela que o tempo de simulação é suficiente para que os AM se afastem da região de interesse caso não tenham afinidade. Nesse caso, as energias menos favoráveis podem ser descartadas pelo AG, sem a necessidade de calcular termos entrópicos computacionalmente caros. Assim, constatamos que apenas as estruturas mais promissoras são capazes de permanecer ligadas ao antígeno durante todo o tempo de simulação.

Como podemos ver na validação do AG (APÊNDICE B - Tabela 3), a comparação entre a EL calculada pelo MM-GBSA e valores calculados de métodos experimentais apresenta algumas limitações relacionadas à parametrização do campo de força e à ausência de alguns termos entrópicos. No entanto, a comparação de $\Delta\Delta G$ entre diferentes moléculas é bastante confiável (Wu *et al.*, 2020; Hou *et al.*, 2010). Nesse caso, calculamos o MM-GBSA da estrutura cristalográfica do complexo formado pelo receptor VEGFR e seu ligante nativo VEGF (Figura 12A), obtendo o valor de $-39,4 \text{ kcal.mol}^{-1}$. Esse valor fornece um parâmetro para ser usado como critério de convergência do AG, como vemos na Figura 14:

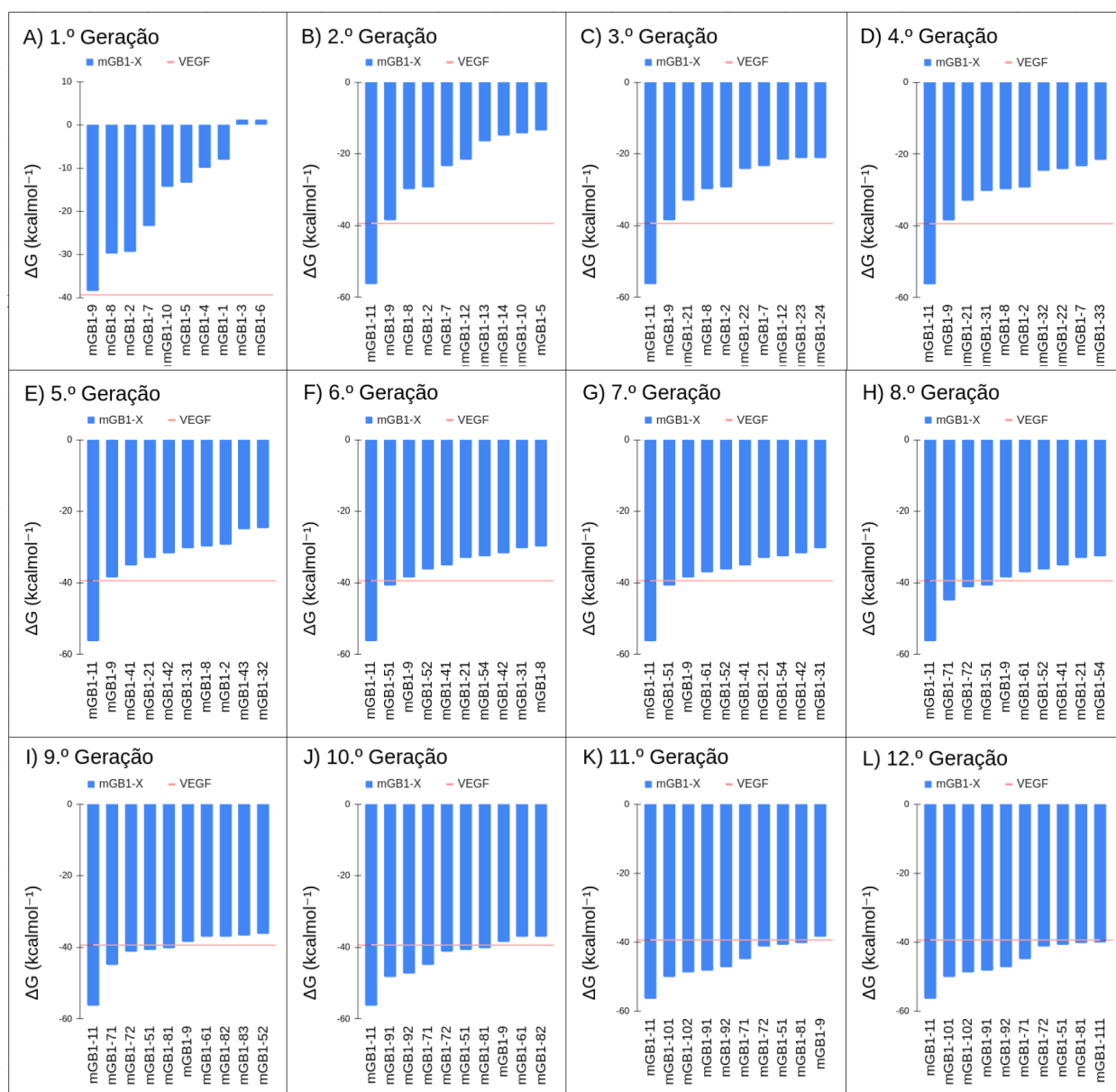


Figura 14 – Energias calculadas em todas as gerações do AG para o sistema VEGFR. O critério de convergência é a EL entre o VEGF e o receptor VEGFR.

Fonte: elaborado pelo Autor (2024).

Na Figura 14A podemos ver o resultado da comparação da EL da primeira geração de mutantes mGB1 (APÊNDICE D - Tabela 4). A função de pontuação resulta em 80% dos mutantes com energia favorável, mostrando que o reconhecimento molecular depende do conjunto de resíduos de aminoácidos envolvidos. Vale ressaltar que, dado que a função de pontuação é bastante simples para determinar qual microdomínio possui maior afinidade, cabe ao AG selecionar as melhores combinações de resíduos. O crossing over, descrito na metodologia, é realizado aos pares, considerando as melhores energias.

Na primeira geração, mGB1-9 com mGB1-8, mGB1-2 com mGB1-7 e assim por diante, originam uma nova população de mutantes filhos. A etapa de DM com cálculo

de MM-GBSA é aplicada novamente. O melhor resultado é alcançado ao cruzar as duas primeiras fitas- β de mGB1-9 com as duas últimas fitas- β de mGB1-8, resultando no mutante mGB1-11, com uma energia de $-56.4 \text{ kcal.mol}^{-1}$. Este valor de energia está abaixo do critério de convergência, com um aumento significativo na afinidade na primeira etapa de otimização (APÊNDICE D - Tabela 5). As melhores energias da população inicial, juntamente com os mutantes filhos, são reordenadas, eliminando mutantes com energias menos favoráveis, formando a segunda geração (APÊNDICE D - Tabela 6).

Na segunda geração (Figura 14 B), obtemos uma redução significativa de energia da população, evidenciando a capacidade de otimização do AG. A população de segunda geração é cruzada novamente, para gerar uma nova população de mutantes (APÊNDICE D - Tabela 7). O AG muitas vezes dá origem a mutantes com energias desfavoráveis, mas como a seleção natural de AG elimina esses mutantes, espera-se que a energia da população seja otimizada a cada geração. De todos os mutantes filhos da segunda geração, apenas 4 obtêm energia abaixo de $-20 \text{ kcal.mol}^{-1}$. As melhores energias ocorrem em mutantes com grupos de aminoácidos mais heterogêneos, que não competem pelo mesmo local de ligação.

Na terceira geração (Figura 14 C), a convergência de energias começa a selecionar as melhores combinações de aminoácidos, gerando microdomínios mais favoráveis ao reconhecimento molecular. Embora a população de terceira geração tenha sido otimizada para energias abaixo de $-20 \text{ kcal.mol}^{-1}$, não obtemos mais mutantes com energias abaixo do critério de convergência, mostrando que o AG pode exigir mais etapas de otimização. Além disso, as colunas da Tabela 8 começam a mostrar as propriedades que as cadeias laterais dos resíduos precisam para otimizar o reconhecimento molecular. Nos mutantes filhos da terceira geração novamente não temos energias com valores abaixo do critério de convergência, mas novamente há uma atualização da população (APÊNDICE D - Tabela 9).

Na quarta geração (Figura 14 D), observa-se uma perda de variabilidade populacional à medida que os microdomínios favoráveis são mantidos (APÊNDICE D - Tabela 10). Isso cria um problema para o AG, que não pode gerar novos mutantes apenas com crossing over. Uma solução para este problema é inserir mutações para aumentar a variabilidade.

No AG padrão, essa mutação seria aleatória, mas quando olhamos a Tabela 10, o resíduo de cisteína na posição 8, que aparece em metade da população, torna-se uma boa opção. Uma solução racional seria realizar mutações de aminoácidos com tamanhos e propriedades semelhantes. No APÊNDICE A - Tabela 2, a função de pontuação prevê que o resíduo de serina pode ocupar esse local. Uma serina tem o mesmo número de átomos e ocupa praticamente o mesmo espaço da cisteína, com

a vantagem de ser doador de ligação de hidrogênio. Assim, por meio de informações preexistentes da função de pontuação, definimos uma possibilidade de mutação, para que o algoritmo possa convergir mais rapidamente, sem necessidade de escolhas aleatórias.

Nos mutantes filhos da quarta geração (APÊNDICE D - Tabela 11), observa-se que o mutante mGB1-46, gerado pela mutação do resíduo de cisteína para serina em mGB1-11, obtém a energia de $-15.5 \text{ kcal.mol}^{-1}$, resultando em uma enorme perda de afinidade. Isto se deve à interação entre o grupo álcool do resíduo serina e o grupo ácido do aspartato na posição 55. Este resultado mostra como os vírus conseguem escapar do sistema imunológico ao revelar que a mutação de um único resíduo pode desestabilizar toda a capacidade de reconhecimento molecular dos anticorpos. Além disso, podemos compreender por que modelos matemáticos muito simples, baseados apenas em interações intermoleculares, são insuficientes para prever a atividade de moléculas bioativas, uma vez que desconsideram a interação entre os grupos químicos da própria molécula.

Na quinta geração (Figura 14 E), toda a população apresenta energia abaixo de $-24 \text{ kcal.mol}^{-1}$, na qual todos os mutantes permanecem ligados ao epítipo durante todo o tempo de simulação. Isso significa que, teoricamente, essas moléculas já possuem potencial para serem AM, com EL favorável durante o tempo de simulação. Um dos mutantes filhos da quinta geração (APÊNDICE D - Tabela 13) possui energia inferior a $-40 \text{ kcal.mol}^{-1}$, mostrando que ainda há muito espaço para otimizar a energia da população.

Na sexta geração (Figura 14 F), a energia da população se aproxima de $-30 \text{ kcal.mol}^{-1}$, uma melhoria significativa em relação à primeira geração. O AG demonstra a capacidade de gerar diversos AM com afinidade, mesmo que não houvesse critérios de convergência. Nos mutantes filhos da sexta geração, obtemos novamente energias entre as 10 primeiras (APÊNDICE D - Tabela 15).

Na sétima geração (Figura 14 G), a população apresenta EL média abaixo de $-30 \text{ kcal.mol}^{-1}$. Observa-se que as mutações de serina na posição 8 permanecem, embora estejam nas últimas posições (APÊNDICE D - Tabela 16). Nos mutantes filhos da sétima geração, obtemos duas energias abaixo do critério de convergência, indicando que o AG pode convergir mais rapidamente (APÊNDICE D - Tabela 17).

Na oitava geração (Figura 14 H), quase metade da população possui energia abaixo do critério de convergência, o último mutante remanescente da primeira geração está agora no meio da Tabela 18. Um dos mutantes filhos da oitava geração obtém energia abaixo do critério de convergência, porém sua sequência é um pouco diferente das demais, o que é interessante do ponto de vista da variabilidade (APÊNDICE D -

Tabela 19).

Na nona geração (Figura 14 I), existem 5 mutantes com energia abaixo de $-40 \text{ kcal.mol}^{-1}$, sendo observada grande similaridade entre mGB1-11 e mGB1-51, mGB1-71 e mGB1-71 (APÊNDICE D - Tabela 20). Nos mutantes filhos da nona geração, obtemos novamente dois mutantes com energias abaixo do critério de convergência, mas observa-se uma saturação de possibilidades de cruzamento (APÊNDICE D - Tabela 21).

Na décima geração (Figura 14 J), a variabilidade populacional é bastante reduzida, o número de permutações possíveis para gerar novas populações é limitado (APÊNDICE D - Tabela 22). Nos mutantes filhos da décima geração obtivemos duas energias abaixo do critério de convergência, com destaque para $-51 \text{ kcal.mol}^{-1}$ do mutante mGB1-101, segunda melhor EL obtida pelo AG (APÊNDICE D - Tabela 23).

Na décima primeira geração (Figura 14 K), 90% da população possui energia abaixo do critério de convergência (APÊNDICE D - Tabela 24). Entre os mutantes filhos da décima primeira geração (APÊNDICE D - Tabela 25), obtivemos o mutante mGB1-111 com energia de $-40,1 \text{ kcal.mol}^{-1}$, completando a convergência do AG na décima segunda geração (Figura 14 L).

Uma observação importante a ser feita é que, embora o VEGFR seja semelhante em tamanho ao GB1, sua estrutura bem definida permite a identificação da região responsável pelo reconhecimento molecular. Com esta informação podemos evitar o problema criado por ignorar outras porções da proteína principal onde o epítipo está localizado. Este é um fator muito importante para a confiabilidade nas etapas experimentais subsequentes.

Após a convergência do AG, notamos que embora a região de ligação ao VEGFR seja predominantemente hidrofóbica, a maioria das cadeias laterais dos resíduos que convergem não são completamente hidrofóbicas. Isto pode ser explicado porque o AG seleciona os resíduos que melhor preenchem as diferentes regiões de ligação do VEGFR com diferentes tipos de interações, o que aumenta a seletividade.

O AG foi capaz de otimizar a EL de uma população de 10 mutantes para valores melhores do que o ligante natural VEGF em apenas 12 gerações. Este resultado demonstra a capacidade do AG para projetar AM com elevada afinidade para o antígeno. Alguns microdomínios, como I49N51D53R55, são extremamente importantes para essa afinidade e, como pode ser observado no APÊNDICE D - Tabela 26, acabam sendo preservados.

4.2.1 Análise do desvio padrão mGB1-11

Um ponto importante a destacar é que a EL não é uma constante, os anticorpos não estão ligados covalentemente ao antígeno. Portanto, se pudéssemos executar uma simulação por um tempo suficientemente longo, veríamos periodicamente uma separação entre os dois. Sendo assim, o erro padrão neste tipo de cálculo de energia não diminui com o maior número de *frames* e tempo de simulação. O que fazemos neste sentido é escolher um intervalo de simulação, que represente um mínimo local dentro do tempo de simulação.

Ainda assim, o desvio padrão é consideravelmente elevado, devido à maior liberdade de movimento de um anticorpo, em comparação com pequenas moléculas confinadas em cavidades proteicas, onde normalmente é utilizado o MM-GBSA. Neste caso, dentro do protocolo AG faz mais sentido comparar apenas as médias da EL dentro do mínimo local. Contudo, é oportuno mostrar como o desvio padrão acompanha a otimização da energia dos AM, como podemos ver na Figura 15:

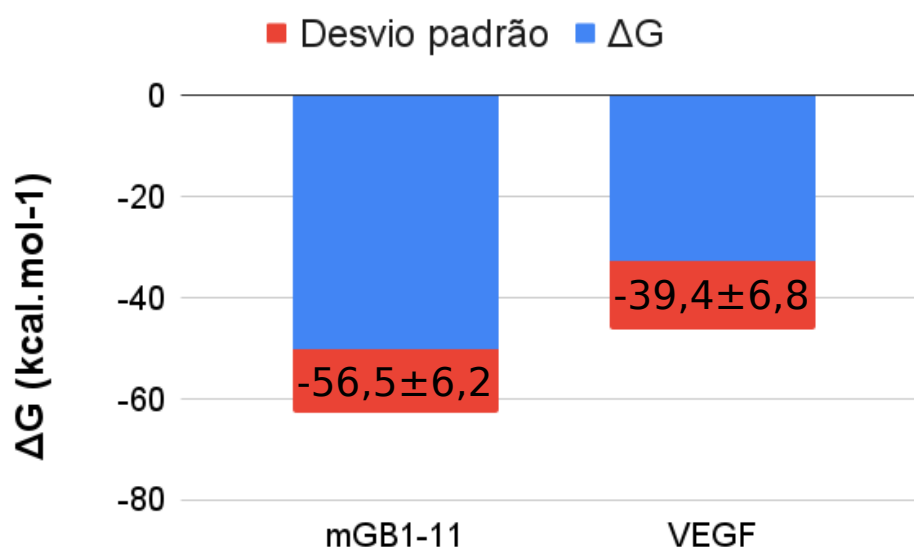


Figura 15 – Desvio padrão da energia calculada para o VEGFR, utilizado como critério de convergência, e a energia do mutante de melhor energia mGB1-11.

Fonte: elaborado pelo Autor (2024).

Na Figura 15 podemos observar que, mesmo com o desvio padrão de aproximadamente 10%, devido à liberdade conformacional do sistema, existe uma diferença de energia considerável entre os -56,5 ± 6,2 kcal.mol⁻¹ do mutante mGB1-11 e - 39,4 ± 6,8 kcal.mol⁻¹ do critério de convergência. Este resultado mostra como o AG pode otimizar a EL de forma confiável e estatisticamente significativa dentro do intervalo de simulação.

4.2.2 Decomposição de energia do mGB1-11

Ao calcular a EL usando o método MM-GBSA, podemos usar a ferramenta de decomposição de energia do Amber para obter informações sobre a contribuição energética de resíduos individuais para a EL. Nesse caso, aplicamos essa metodologia no mutante melhor classificado mGB1-11, como podemos ver na Figura 16:

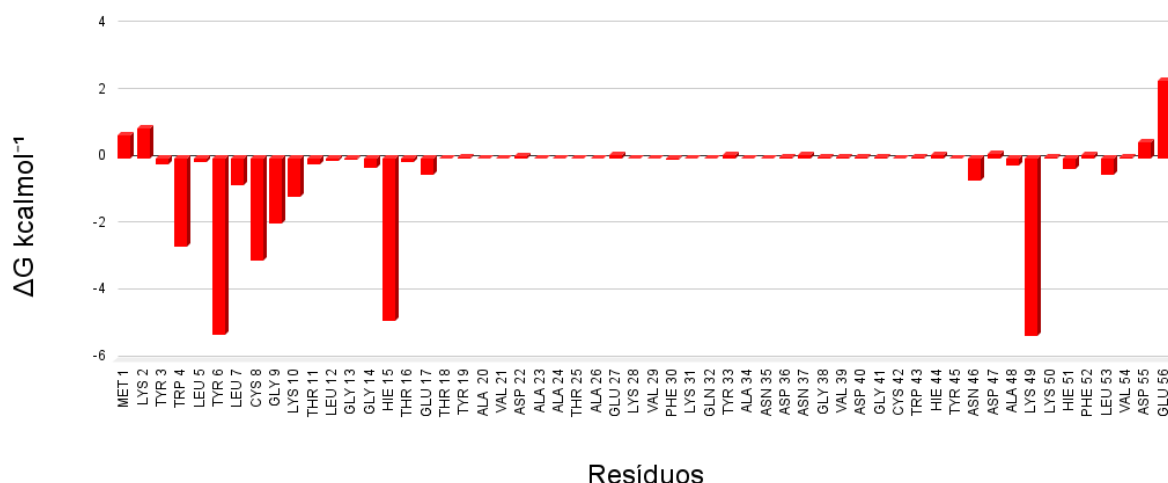


Figura 16 – Contribuições individuais dos resíduos do mutante mGB1-11 na EL total.

Fonte: elaborado pelo Autor (2024).

Como previsto, a ferramenta de decomposição de energia Amber mostra que os resíduos que mais contribuem para a EL são aqueles selecionados pelo AG. A contribuição energética mais favorável é observada nos resíduos de tirosina na posição 6, histidina na posição 15 e lisina na posição 49, todos com aproximadamente -5 kcal/mol⁻¹. Nota-se que os resíduos próximos ao N-terminal têm um peso maior na energia total. Isso pode ser facilmente explicado, uma vez que, quando as possibilidades de cruzamento das fitas- β são esgotadas, o AG realiza trocas individuais do N-terminal para o C-terminal. Este resultado pode ser interpretado como uma medida do nível de otimização do AM pelo AG, indicando que haveria possibilidade de melhoria das energias caso o algoritmo continuasse a ser executado.

O fato desta ferramenta revelar os resíduos mais importantes para a energia total, intuitivamente nos leva a pensar em seu uso para projetar mutantes melhores. Um exemplo seria modificar o resíduo de glutamato na posição 56, que tem uma contribuição desfavorável para a ligação. No entanto, conforme demonstrado pelo AG, a capacidade de reconhecimento molecular é realizada pelo conjunto de aminoácidos e a alteração de um resíduo poderia desestabilizar essa função.

4.3 Algoritmo genético do sistema HC-E2

Ao desenvolver uma nova abordagem para projeto de MA, devemos mostrar a aplicação em outros sistemas, para verificar a confiabilidade da metodologia. Portanto, aplicamos o protocolo do AG no desenvolvimento de AM para o envelope HC-E2 do vírus da hepatite C. Esse epítipo é formado somente por duas fitas- β cuja estrutura é mantida por ligações de dissulfeto. O sistema base usado pelo AG pode ser visto na Figura 17.

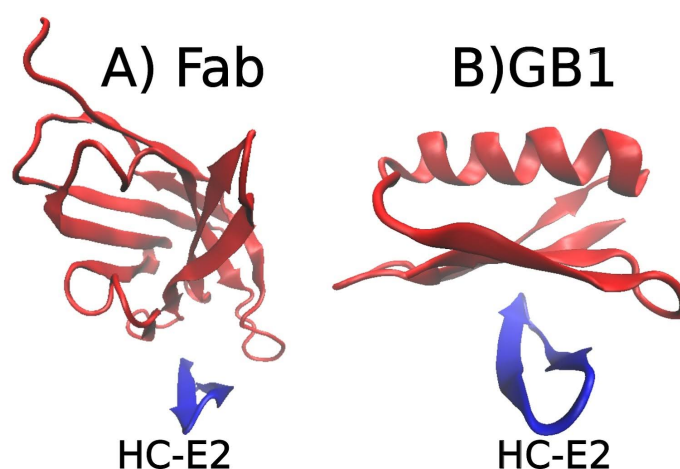


Figura 17 – Ilustração mostrando o sistema base para a DM no projeto de AM. A) Sistema formado pelo fragmento Fab do anticorpo e o envelope HC-E2. B) Sistema base formado pelo GB1 e o envelope HC-E2.

Fonte: elaborado pelo Autor (2024).

Na Figura 17 A, podemos observar como o fragmento Fab do anticorpo se liga ao envelope HC-E2 na estrutura cristalográfica. Enquanto isso, na Figura 17 B, é apresentada a conformação inicial do sistema base do GB1 antes da execução do AG. Em ambas as ilustrações podem-se observar a grande diferença de tamanho em relação ao epítipo HC-E2. É importante destacar que um anticorpo completo contém dois fragmentos Fab que podem se ligar ao antígeno de forma independente. Porém, para critérios de comparação com AM, faz mais sentido utilizar apenas um dos fragmentos, pois queremos otimizar o reconhecimento molecular para uma única região.

Portanto, aplicamos o MM-GBSA ao complexo formado pelo fragmento Fab e pelo envelope HC-E2, com base na estrutura cristalográfica resolvida por Pierce *et al.* (2017). A partir deste cálculo obtemos a EL de $-27,5 \text{ kcal.mol}^{-1}$, que foi utilizada como critério de convergência. O mesmo protocolo do AG utilizado anteriormente foi aplicado ao sistema base, com a vantagem de utilizar um antígeno de tamanho extremamente reduzido.

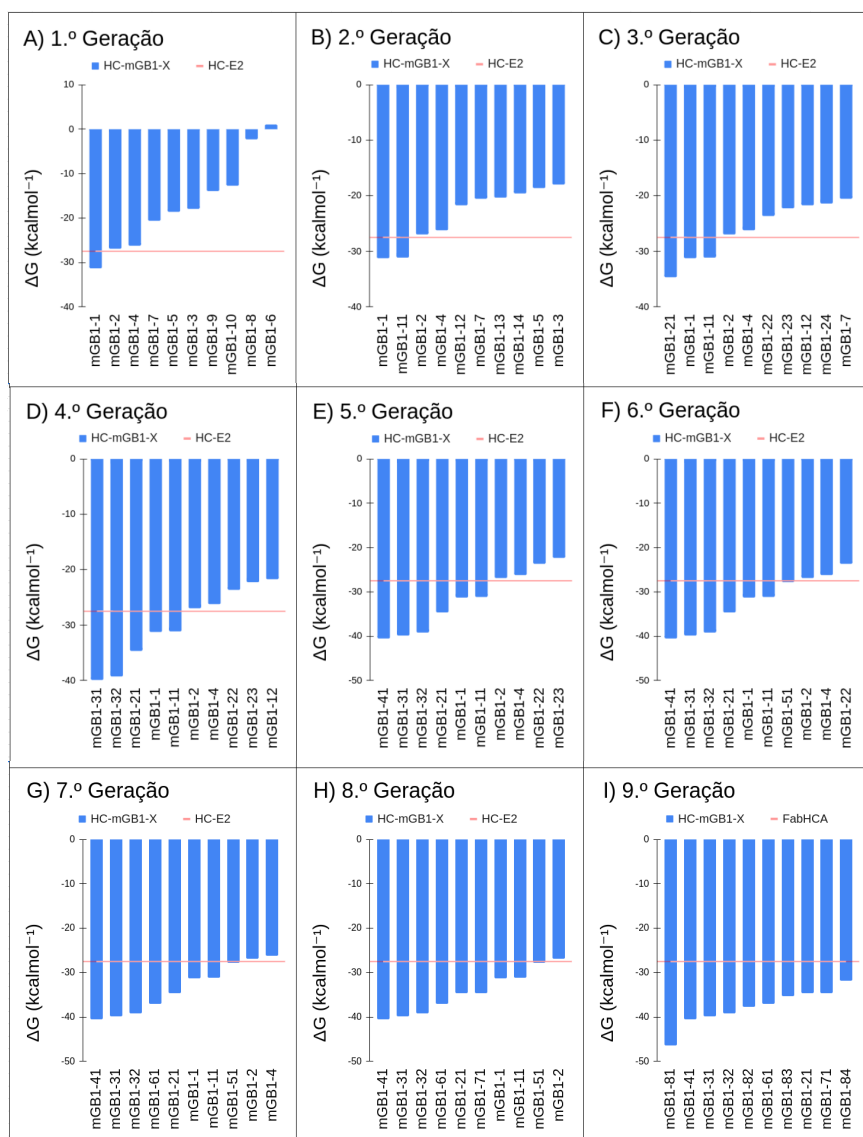


Figura 18 – Energias calculadas em todas as gerações do AG para o sistema HC-E2. O critério de convergência é a EL entre o Grupo Fab do anticorpo e o peptídeo HC-E2.

Fonte: elaborado pelo Autor (2024).

A Figura 18 mostra o funcionamento do AG no segundo sistema, otimizando a afinidade dos AM para o envelope HC-E2 do vírus da hepatite. Na Figura 18 A, a população inicial, gerada pela função de pontuação, possui 90% de mutantes com EL favorável à ligação. Além disso, o mutante de primeira geração HC-mGB1 possui uma EL menor que o critério de convergência (APÊNDICE E - Tabela 27). Este resultado mostra uma boa qualidade da calibração da função de pontuação, que é capaz de prever a ordem dos dois mutantes com melhor afinidade na primeira geração.

Devido ao pequeno tamanho do antígeno, que contém apenas 15 resíduos, a convergência ocorre muito mais rapidamente. A população mutante é atualizada com energias abaixo do critério de convergência em todas as gerações do AG.

É interessante notar que, à medida que o AG converge, os microdomínios mais

favoráveis passam a garantir a EL favorável, aumentando o número de mutantes viáveis por geração. Este efeito pode ser observado nos mutantes filhos da oitava geração (APÊNDICE E - Tabela 42), onde existem quatro mutantes com energia abaixo do critério de convergência.

A velocidade de otimização permite que o AG convirja em apenas nove gerações, mostrando novamente sua aplicabilidade no projeto AM. A análise da nona geração (APÊNDICE E - Tabela 43) mostra que o AG converge para o microdomínio L13D15K17L19 que está presente em diversos mutantes com energias favoráveis. Este resultado destaca a possibilidade de redução do tamanho do AM em abordagens futuras.

Devemos ressaltar que o pequeno tamanho do HC-E2, além de sua função puramente estrutural, pode gerar problemas na sua utilização como antígeno para AM. Isto se deve à falta de informação sobre quais estruturas virais estão localizadas nas proximidades deste epítipo. Portanto, pode-se afirmar que o HC-E2 seria um antígeno pouco confiável em casos experimentais, sendo interessante apenas como prova do conceito teórico do AG.

4.3.1 Análise do desvio padrão HC-mGB1-81

Novamente, é importante avaliar se o resultado do AG no sistema HC-E2 é estatisticamente confiável. Portanto, foi realizada uma análise do desvio padrão das energias calculadas entre o HC-E2, utilizado como critério de convergência, e o mutante com melhor energia HCmGB1-81, como podemos observar na Figura 19:

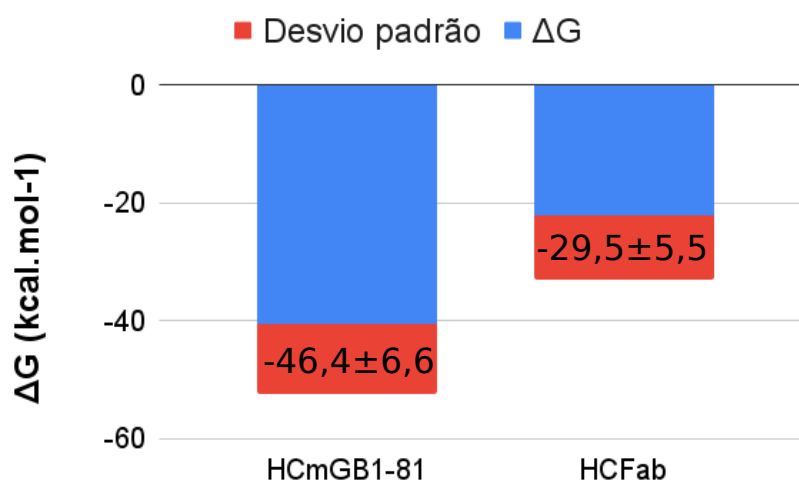


Figura 19 – Desvio padrão da energia calculada para o HC-E2, utilizado como critério de convergência, e a energia do mutante de melhor energia HCmGB1-81.

Fonte: elaborado pelo Autor (2024).

Na Figura 19 vemos novamente como o desvio padrão acompanha a otimização energética dentro do AG. Embora o desvio padrão tenha um valor consideravelmente elevado, existe uma separação estatisticamente significativa entre as energias de $-46,4 \pm 6,6 \text{ kcal.mol}^{-1}$ do HCmGB1-81 e $-29,5 \pm 5,5 \text{ kcal.mol}^{-1}$ do critério de convergência, revelando que a população foi de fato otimizada.

4.3.2 Decomposição de energia do mutante HCmGB1-81

A ferramenta de decomposição de energia Amber foi aplicada ao mutante HCmGB1-81 e os resultados podem ser vistos no gráfico da Figura 20:

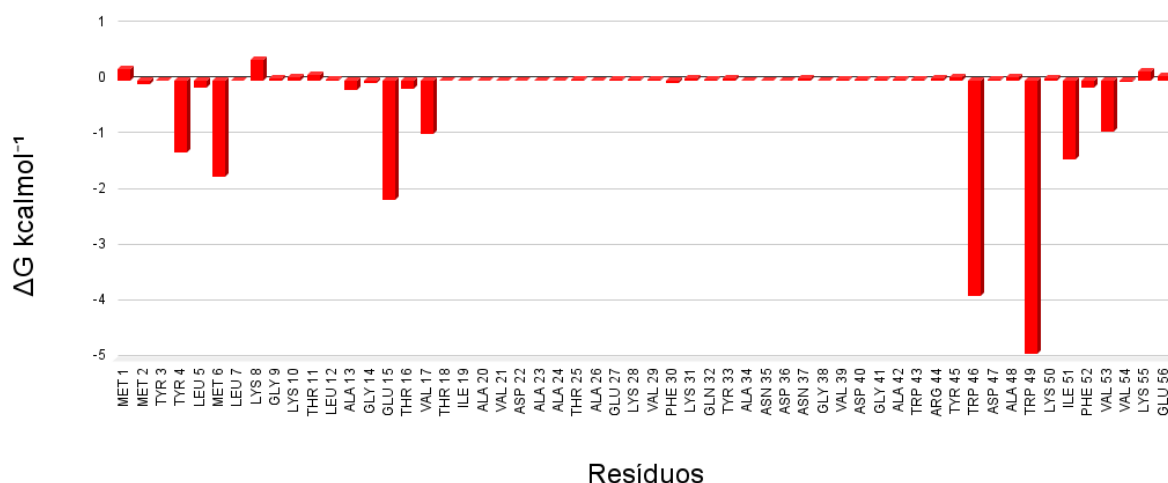


Figura 20 – Contribuições individuais dos resíduos do mutante HCmGB1-81 na EL total.

Fonte: elaborado pelo Autor (2024).

Na Figura 20, a análise da energia para o mutante HC-mGB1-81 mostra que os resíduos de triptofano, nas posições 46 e 49, têm a maior contribuição para a energia total. Podemos observar que os resíduos mais próximos do C-terminal possuem maior peso na EL total, o que pode ser interpretado como uma saturação na capacidade de otimização do AG. Essa saturação pode ser atribuída ao pequeno tamanho do antígeno, fazendo com que o AG convirja mais rapidamente. Deste resultado podemos concluir que o tamanho do antígeno alvo, com dimensões mais compatíveis com a área variável do AM, oferece melhores resultados devido ao maior número de possibilidades de interação.

4.4 Algoritmo genético do sistema RBD

Nos sistemas anteriores, o protocolo de AG foi aplicado em modelos puramente teóricos, como uma prova de conceito, para demonstrar a sua viabilidade. No terceiro sistema, composto pelo domínio RBD da proteína spike, a intenção principal era o desenvolvimento de AM que pudessem ser usados posteriormente na fase experimental. Esse epítipo, além de muito estável, não contém modificações pós-traducionais, podendo ser expresso facilmente por meio de plasmídeos bacterianos.

Nas modificações do GB1, com base na experiência adquirida nos sistemas anteriores, optamos por remover a presença de resíduos de cisteína na população inicial (Liao; Rabideau; Pentelute, 2014), que poderiam formar ligações de dissulfeto, indesejadas durante a expressão em laboratório. Assim sendo, executamos novamente o protocolo de AG no projeto de AM para o domínio RBD, como podemos observar na Figura 21:

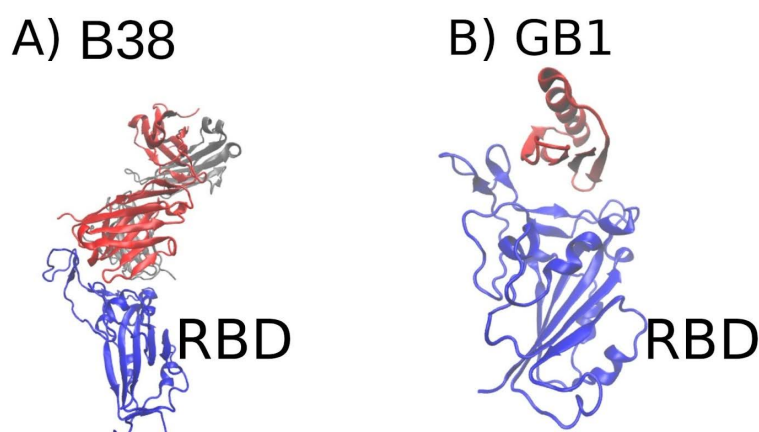


Figura 21 – Ilustração da região do domínio RBD (Azul) neutralizada por anticorpos. A) Estrutura do anticorpo B38 (vermelho) neutralizando a região de reconhecimento ACE2. B) Estrutura do domínio GB1 (vermelho) na região de otimização do sistema base.

Fonte: elaborado pelo Autor (2024).

Na Figura 21 A podemos ver como o fragmento Fab do anticorpo B38 se liga à região de reconhecimento ACE2, bloqueando o mecanismo de entrada do vírus na célula. Esta região é a mais importante do RBD porque, além de bloquear a função de reconhecimento molecular, permite a comparação da EL do anticorpo B38. Portanto, optamos por otimizar a capacidade de reconhecimento molecular do GB1 nesta região. Devido ao grande tamanho dos anticorpos, há muitos microdomínios de ligação, que são capazes de manter a capacidade de reconhecimento molecular. No entanto, essa flexibilidade é limitada pelo pequeno tamanho do GB1. Portanto, optamos por utilizar a EL média entre os fragmentos Fab do anticorpo B38, que são mais próximos em

tamanho do GB1. Nesse caso, obtemos a EL de $-50,23 \text{ kcal.mol}^{-1}$ usada como critério de convergência.

As estruturas do antígeno RBD e do peptídeo GB1 mostradas na Figura 20B foram utilizadas como condição inicial do AG. Neste sistema optamos por gerar uma população inicial de 12 mutantes. A população inicial foi submetida ao AG conforme pode ser observado na Figura 22:

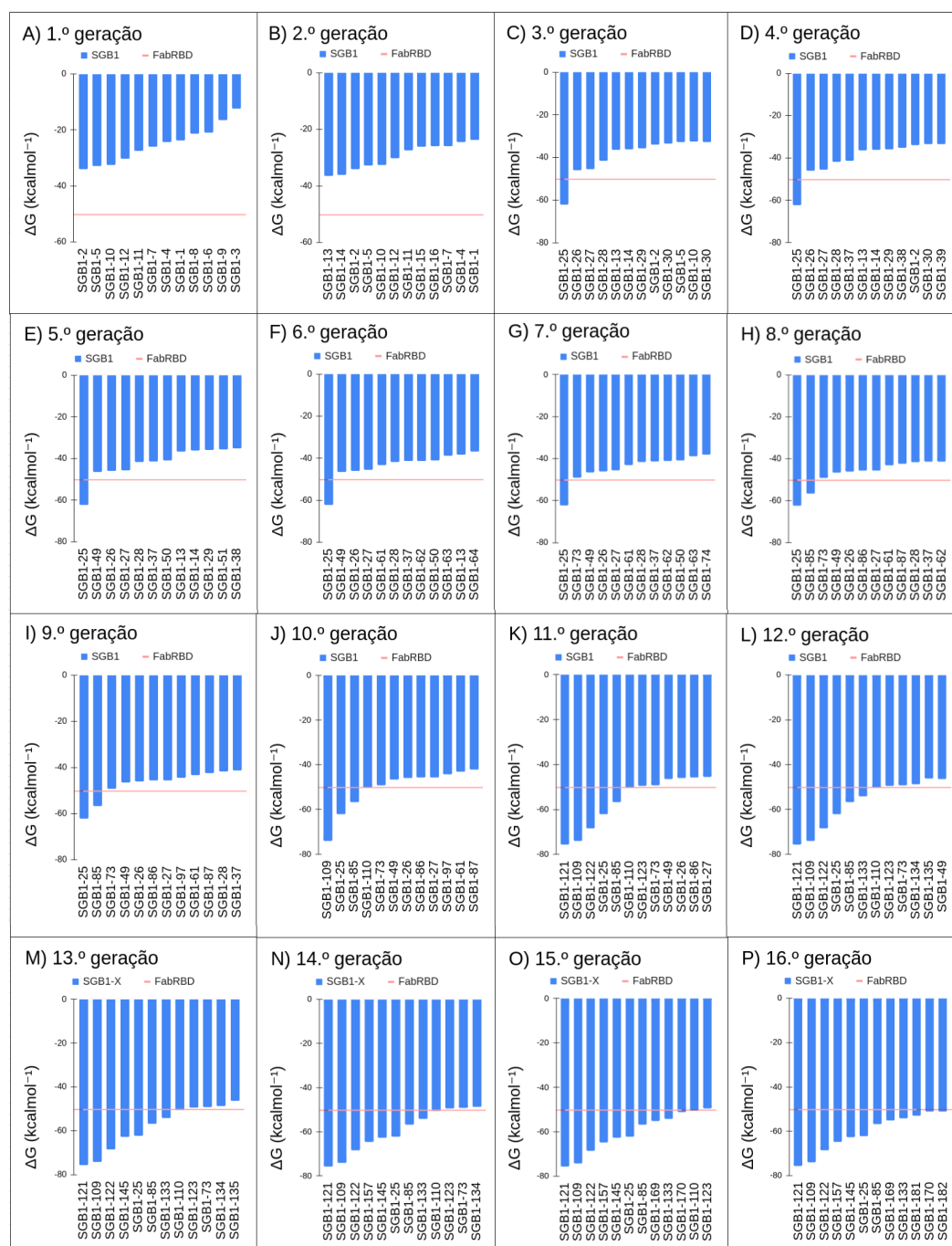


Figura 22 – Energias calculadas em todas as gerações do AG para o sistema RBD. O critério de convergência é a EL entre o fragmento Fab do anticorpo B38 e o Domínio RBD da proteína spike.

Fonte: elaborado pelo Autor (2024).

Na Figura 22 A, podemos ver que todos os mutantes da população inicial tem uma EL negativa, com alguma afinidade pelo antígeno. Mais uma vez, a função de pontuação alcança o objetivo de gerar uma população inicial com energia favorável ao reconhecimento molecular. Contudo, estas energias continuam bem acima do critério de convergência de $-50,23 \text{ kcal.mol}^{-1}$.

O primeiro mutante com energia abaixo do critério de convergência aparece na 3ª geração (APÊNDICE F - Tabela 48) e o segundo mutante apenas na 8ª geração (APÊNDICE F - Tabela 58). Devido ao critério de convergência muito mais exigente, a convergência ocorre de forma muito mais lenta quando comparada com os sistemas anteriores, no entanto, é contínua se olharmos para a média da população ao longo das gerações.

A partir da décima geração, mutantes com energias abaixo do critério de convergência começam a surgir com maior frequência, com o AG convergindo na 16ª geração. Ao longo dessas 16 gerações o AG favorece a formação do microdomínio Q4E6R8E10, como pode ser visto no APÊNDICE F - Tabela 74.

É interessante notar na Tabela 74 que o GB1 original obtém uma EL negativa, mesmo não tendo correlação conhecida com o domínio RBD. Porém, esse resultado é enganoso, pois como vimos nas mais de 400 DMs realizadas no AG, essa energia desconsidera os termos entrópicos, sendo apenas um artefato resultante da compatibilidade estérica entre RBD e GB1.

O mutante SGB1-121, surgido na décima primeira geração, obteve uma energia de $-75,6 \text{ kcal. mol}^{-1}$, sendo a melhor energia dentro da população final. Este resultado mostra novamente que não é necessário esperar pela convergência total do AG para obter mutantes confiáveis.

Devido ao tamanho consideravelmente maior do RBD em comparação com o GB1, temos muito mais confiança sobre a região de reconhecimento molecular e a função que ela desempenha. Além disso, sabemos exatamente quais estruturas proteicas estão localizadas nas proximidades, tornando este antígeno ideal para prova de conceito experimental de AM.

É importante destacar que o mutante SGB1-121, melhor classificado durante o AG, obtém uma energia melhor do que a calculada para os peptídeos mutantes do domínio ACE2, desenvolvidos por Sitthiyotha e Chunsrivirod (2021), usando a mesma ferramenta do MM-GBSA do Amber. Este resultado mostra o potencial teórico dos AM baseados em GB1, capazes de se ligar ao antígeno com melhor afinidade do que os peptídeos que são ligantes naturais do domínio RBD.

4.4.1 Análise do desvio padrão SGB1-121

O modelo do AG para RBD foi projetado tendo em mente as etapas experimentais subsequentes. Portanto, mais duas DMs, cujas trajetórias podem ser visualizadas no APÊNDICE G, foram realizadas com este sistema, seguindo a mesma metodologia, como forma de avaliar a confiabilidade do modelo mutante SGB1-121.

No primeiro, foi realizada uma nova DM de 100 ns com a mesma estrutura SGB1-121 e RBD nas condições iniciais do AG (SGB1-121_{MD}). Então, após adicionar um grupo his-tag de 6 histidinas a região C-terminal do mutante SGB1-121, foi realizada uma nova DM de 100 ns (SGB1-121_{histag}).

Esta simulação é necessária porque a his-tag é utilizada na etapa de purificação por cromatografia de afinidade em coluna de Níquel. Esta comparação é crucial para avaliar se a capacidade de reconhecimento molecular é mantida mesmo com a presença do his-tag na região C-terminal. A análise do desvio padrão das energias calculadas entre RBD, utilizado como critério de convergência, e a DM de SGB1-121 é mostrada na Figura 23:

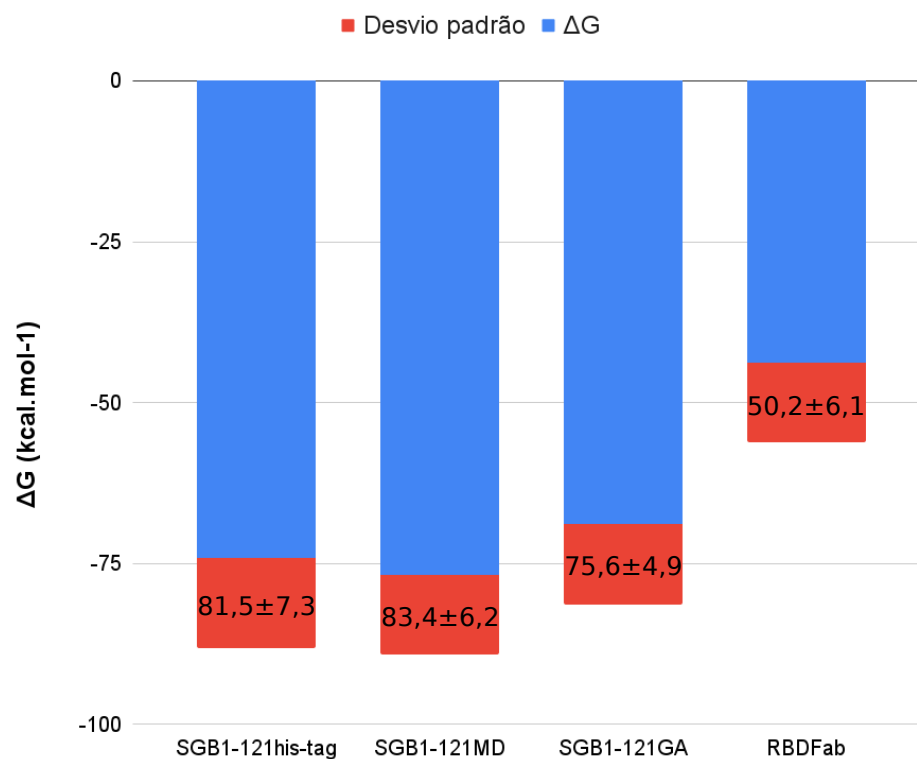


Figura 23 – Desvio padrão da energia calculada para o RBD, utilizado como critério de convergência, e as energias do mutante de melhor energia SGB1-121.

Fonte: elaborado pelo Autor (2024).

A Figura 23 mostra as energias de $-83,0 \pm 6,2$ kcal.mol⁻¹ para o mutante SGB1-121_{MD} e $-81,5 \pm 7,3$ kcal. mol⁻¹ para o mutante SGB1-121_{his-tag}, resultados melhores que o valor de $-75,6 \pm 4,9$ kcal. mol⁻¹ obtido por SGB1-121 no AG. Os valores de energia estão muito próximos, dentro do desvio padrão, mesmo em simulações diferentes. Este resultado confirma que o tempo de amostragem da DM é suficiente para obter resultados confiáveis.

Esses valores de energia estão muito distantes do valor do critério de convergência, mesmo considerando o desvio padrão, mostrando que as energias foram otimizadas pelo AG de forma estatisticamente significativa. Além disso, observamos que mesmo com a adição da his-tag no mutante SGB1-121, não parece ter ocorrido uma perda na capacidade de reconhecimento molecular, um fator importante a ser considerado antes de prosseguir para a etapa experimental.

4.4.2 Análise de termos entrópicos

Uma vez convergido o AG, podemos analisar a EL de um ponto de vista mais completo, inserindo os termos entrópicos de origem vibracional (Genheden *et al.*, 2012). Neste caso, utilizamos as trajetórias de DM dos mutantes SGB1-121_{MD} e SGB1-121_{his-tag} e recalculamos o MM-GBSA, juntamente com o termo entrópico, através dos modos normais de vibração, como pode ser visto na Figura 24.

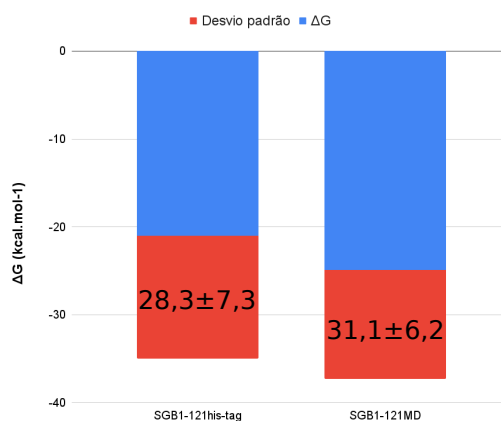


Figura 24 – Valores da EL dos mutantes SGB1-121_{MD} e SGB1-121_{his-tag} com adição do termo entrópico dos modos normais de vibração.

Fonte: elaborado pelo Autor (2024).

Ao somar os termos entrópicos da EL dos mutantes obtemos $-31,1 \pm 6,2$ kcal.mol⁻¹ para o mutante SGB1-121_{MD} e $-28,3 \pm 7,3$ kcal.mol⁻¹ para o mutante SGB1-121_{his-tag}. Este resultado mostra uma EL negativa, favorável ao reconhecimento molecular, mesmo com a contribuição positiva do termo entrópico e do desvio padrão. Além disso, demonstra que a his-tag tem pouco efeito nas propriedades esperadas para o mutante

SGB1-121, sem diferença estatisticamente significativa. Este resultado mostra novamente como o $\Delta\Delta G$ consegue ordenar as energias independentemente do cálculo da componente entrópica vibracional, computacionalmente muito cara para ser usada no AG.

4.4.3 Decomposição de energia do mutante SGB1-121

A ferramenta de decomposição de energia Amber foi aplicada ao mutante SGB1-121 e os resultados podem ser vistos no gráfico da Figura 25:

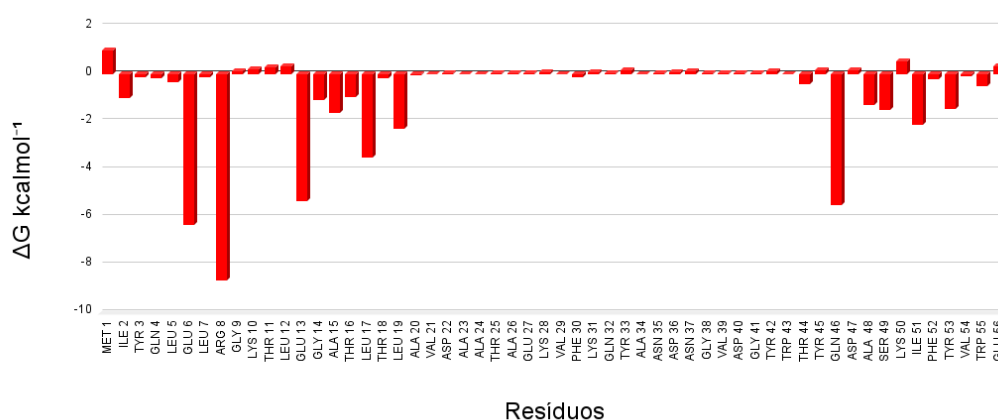


Figura 25 – Contribuições individuais dos resíduos do mutante SGB1-121 na EL total.

Fonte: elaborado pelo Autor (2024).

Na Figura 25, a análise da decomposição de energia do mutante SGB1-121 mostra vários resíduos importantes ao reconhecimento molecular, como o glutamato nas posições 6 e 13, a arginina na posição 8 e glutamina na posição 46, todas com valores abaixo de $-5 \text{ kcal}\cdot\text{mol}^{-1}$. Além disso, há uma contribuição maior na energia para resíduos próximos ao N-terminal, indicando que haveria possibilidade de otimizar ainda mais a energia total na região C-terminal em caso de continuidade do AG.

Devemos ressaltar que o epítipo do domínio RBD é muito maior que os antígenos anteriormente utilizados, oferecendo muito mais possibilidades de interação, o que parece favorecer a capacidade de otimização do AG.

Uma otimização ideal dos AM seria alcançada quando os resíduos modificados apresentassem uma EL muito mais favorável do que os resíduos originais do GB1. Dessa forma, resultados próximos a esse padrão podem indicar o estágio de otimização dos AM pelo AG. Nesse caso, a ferramenta de decomposição de energia poderia ser utilizada como critério de convergência nos casos em que os ligantes naturais são desconhecidos.

4.4.4 Modelagem por homologia SGB1-121

A análise de estabilidade dos mutantes pode ser realizada usando o servidor online de predição estrutural por homologia (SWISS-MODEL). A modelagem por homologia (MH) é um dos métodos mais precisos para gerar modelos tridimensionais da estrutura de proteínas (Waterhouse *et al.*, 2018).

Esta ferramenta usa informações de estruturas experimentais, com sequências homólogas, para prever o enovelamento de uma proteína. As sequências de aminoácidos do mutante SGB1-121: MIYQLELRGKTLEGATLTAVDAATAEKVFKQYANDNGVDG-YWTYQDASKIFYVWE foram analisadas pelo SWISS-MODEL e os resultados podem ser vistos na Figura 26:

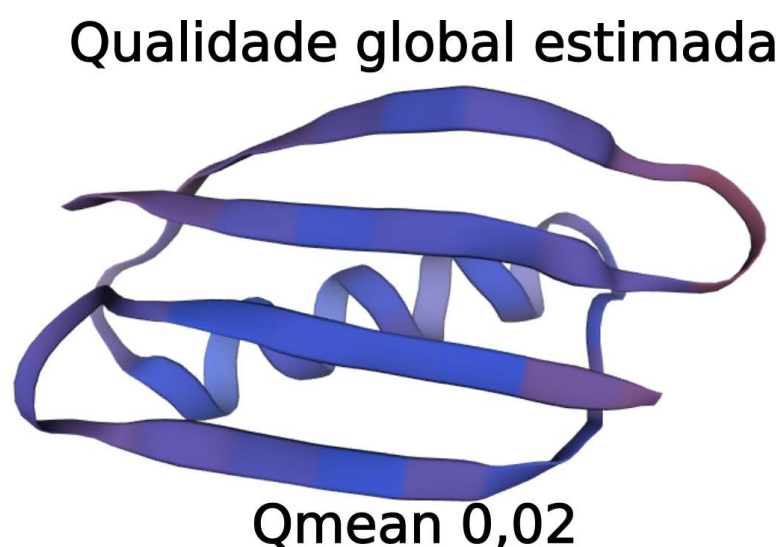


Figura 26 – Ilustração da previsão estrutural da ferramenta de MH. Valores de QMEAN próximos a zero indicam grande similaridade com estruturas experimentais.

Fonte: elaborado pelo Autor (2024).

Na Figura 26, a predição estrutural realizada por MH mostra que o mutante SGB1-121 é formado por uma α -hélice e quatro fitas- β . O modelo gerado não parece ter sofrido nenhuma modificação estrutural que descaracterize a estrutura original do GB1, que é o resultado desejado ao modelar um AM. Isto é confirmado pelos valores de QMEAN próximos de zero, que indicam alta similaridade com estruturas experimentais homólogas.

4.4.5 Teste de ancoragem molecular

Devido à novidade do modelo de AG proposto, devemos garantir a confiabilidade dos mutantes, evitando que o modelo seja viesado apenas para esta abordagem. Para isso, podemos aplicar uma abordagem mais tradicional, usando MH em conjunto com a ancoragem molecular. Dessa forma, optamos por usar a estrutura do mutante SGB1-121, modelada anteriormente pelo SWISS-MODEL, para realizar a predição de ancoragem molecular na estrutura do domínio RBD, usando o servidor online HDock (Yan *et al.*, 2020). A ancoragem molecular realizada de forma independente pode fornecer uma forte evidência da confiabilidade da metodologia aplicada no AG, como vemos na Figura 27:

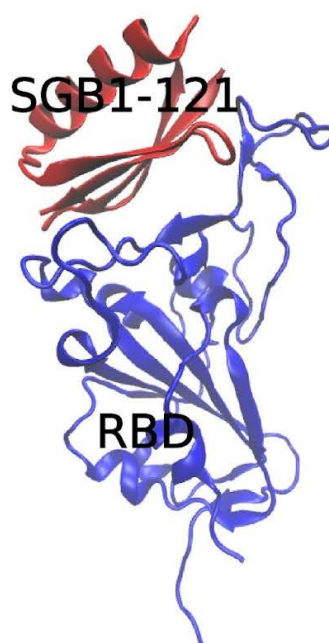


Figura 27 – Ilustração do modelo de ancoragem molecular realizado pelo servidor online HDock.

Fonte: elaborado pelo Autor (2024).

Na Figura 27, podemos observar o resultado da melhor predição do servidor HDock. O resultado prevê que o mutante SGB1-121 se liga ao domínio RBD na mesma região e conformação do sistema base usado na condição inicial do AG. Isto significa que se não conhecermos a região de ligação e apenas ancorarmos as estruturas do modelo, como tradicionalmente é feito para estimar a EL, obteremos a mesma conclusão. Os resultados mostram que o AG consegue otimizar a capacidade de reconhecimento molecular o suficiente para ser detectado por software independente, obtendo resultados que só seriam esperados para bancos de dados de proteínas com atividade biológica conhecida.

4.4.6 Teste de His-tag

Nosso modelo de AM é uma representação simplificada de um sistema mais complexo e, conseqüentemente, existem muitas condições no ambiente biológico desconsideradas durante a simulação. A capacidade de prever computacionalmente as condições reais do sistema ajuda a evitar possíveis problemas nas fases experimentais. Um exemplo disso é a hist-tag de 6 histidinas, utilizada na etapa de purificação por cromatografia de afinidade. Devido ao pequeno tamanho dos AM, a his-tag poderia gerar problemas, como impedimento estérico, que afetaria o reconhecimento molecular nas fases experimentais. A sobreposição dos *frames* da trajetória de DM dos mutantes SGB1-121_{MD} e SGB1-121_{his-tag} pode ser vista na Figura 28:

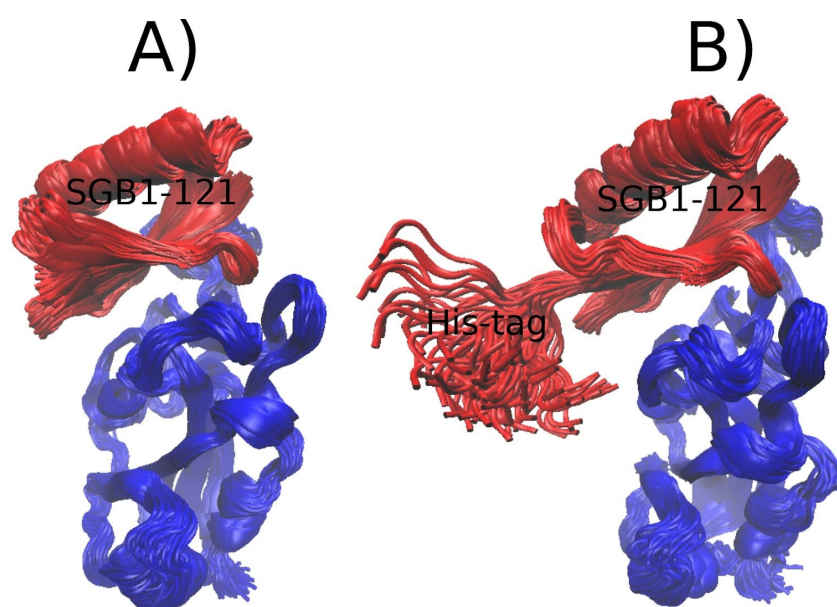


Figura 28 – Ilustração da sobreposição dos *frames* durante 10 ns finais das DMs do mutante SGB1-121. A) Sobreposição dos *frames* da trajetória do mutante SGB1-121_{MD}. B) Sobreposição dos *frames* da trajetória do mutante SGB1-121_{his-tag}.

Fonte: elaborado pelo Autor (2024).

Na Figura 28, podemos observar a sobreposição dos *frames* durante 10 ns finais da DM dos mutantes SGB1-121_{MD} e SGB1-121_{his-tag}. Nota-se que os mutantes permanecem firmemente ligados ao antígeno RBD mesmo com a presença da his-tag, que tem um grau de liberdade bem maior. Desse modo, observa-se que a presença de his-tag não deve gerar problemas que atrapalhem as etapas experimentais, não havendo necessidade de métodos de modificação pós-traducional para extração his-tag após a expressão proteica.

4.4.7 Análise da estabilidade estrutural

A escolha do domínio GB1 deve-se a sua formidável estabilidade frente a mudança de aminoácidos. Deste modo, é necessário analisar a estabilidade estrutural dos AM. Neste caso, optamos por realizar DMs de 100 ns dos mutantes SGB1-121 e SGB1-121_{his-tag} sem a presença do antígeno RBD. Sua estabilidade foi avaliada por meio da raiz quadrada do desvio quadrático médio (RMSD) e da raiz quadrada da flutuação quadrática média (RMSF), conforme pode ser observado na Figura 29:

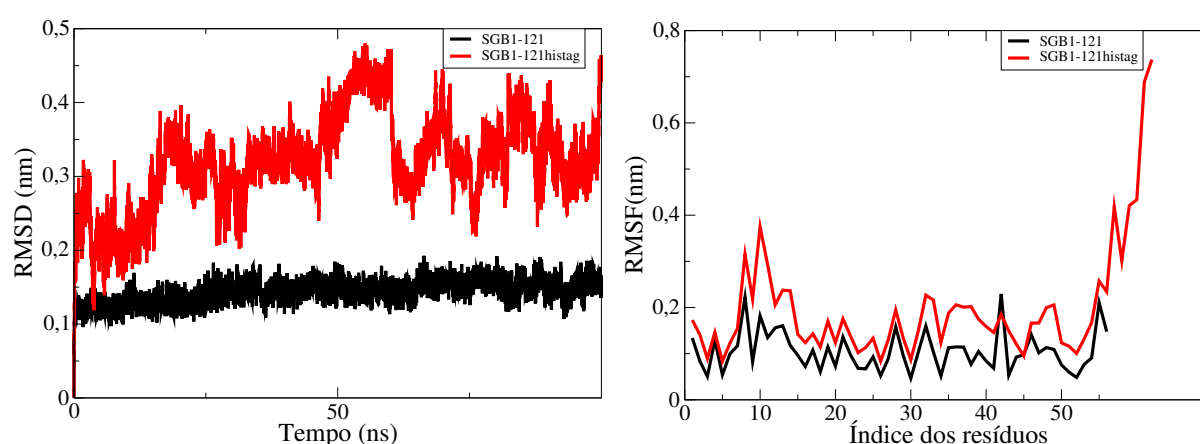


Figura 29 – Análise estrutural de todos os mutantes SGB1-121. A) Gráfico de RMSD dos mutantes com e sem his-tag. B) Gráfico de RMSF dos mutantes com e sem his-tag.

Fonte: elaborado pelo Autor (2024).

A análise gráfica do RMSD (Figura 29A) mostra que, durante as DMs de 100 ns, as estruturas variam em menos de 0,5 nm quando comparadas ao frame inicial, mostrando alta estabilidade. Há um RMSD muito maior no mutante SGB1-121_{his-tag}, o que pode ser explicado pela maior liberdade de movimento da his-tag. O RMSD não é suficiente para definir a estabilidade da região de interesse devido ao apêndice muito flexível formado pela his-tag. A análise do RMSF (Figura 29B) mostra que os átomos dos mutantes têm uma flutuação muito baixa, dentro dos graus de liberdade das cadeias laterais dos resíduos de aminoácidos, com exceção da região do apêndice formado pela his-tag. O RMSF consegue revelar melhor como a estrutura do mutante SGB1-121_{his-tag} tem uma alta estabilidade.

4.4.8 Análise de estrutura secundária

A verificação da estabilidade estrutural do mutante SGB1-121 também pode ser feita através da ferramenta de análise de estrutura secundária (DSSP) do software VMD, como podemos ver na Figura 30:

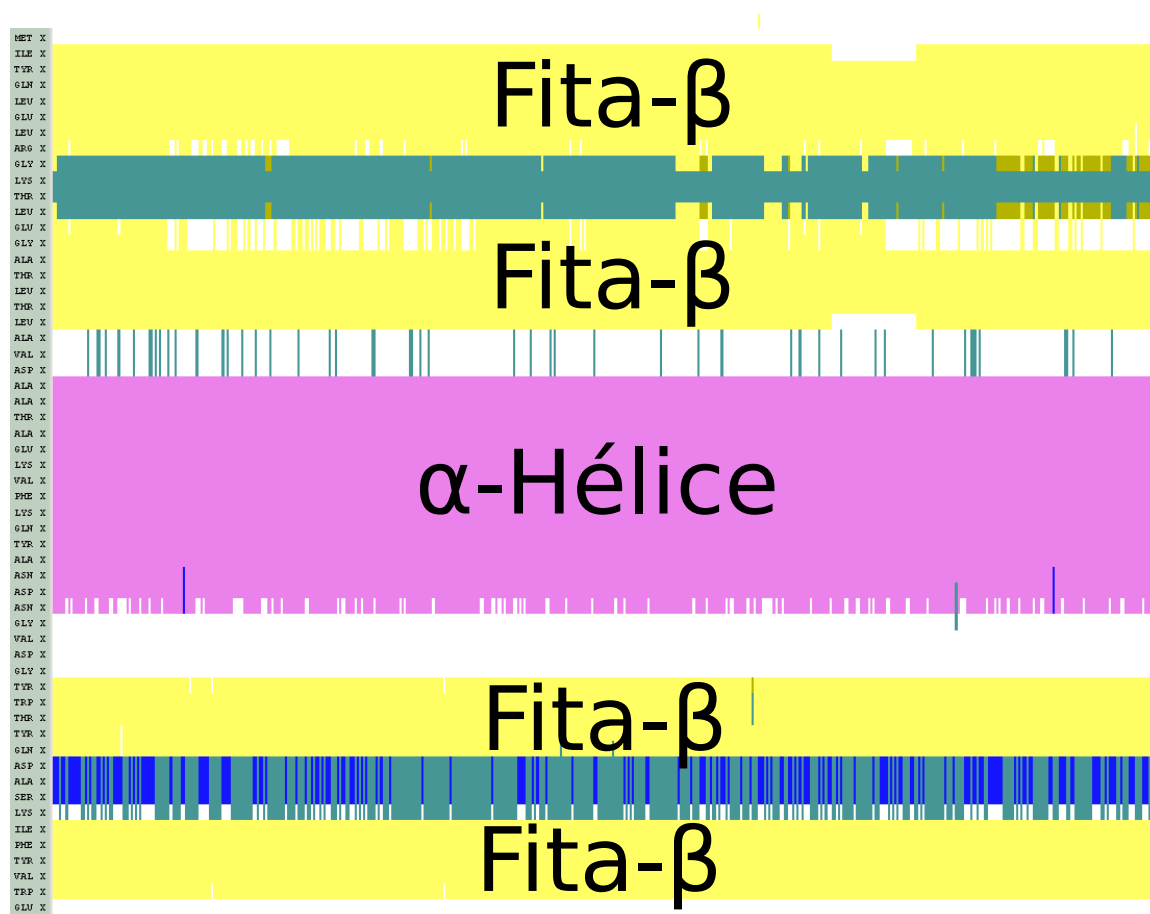


Figura 30 – Análise estrutural da estrutura secundária do mutante SGB1-121 durante a DM de 100 ns.

Fonte: elaborado pelo Autor (2024).

A Figura 30 mostra como a estrutura formada por quatro fitas- β (Amarelo) e uma α -hélice (Lilás) permanece estável durante a DM de 100 ns. Esse resultado demonstra mais uma vez a grande estabilidade do domínio GB1 frente a modificação de aminoácidos. Os resultados da análise estrutural por meio de MH, RMSD, RMSF e DSSP demonstram a alta estabilidade do mutante SGB1-121 gerado a partir da estrutura GB1. Este resultado indica que o AM têm boas possibilidades de manter a integridade estrutural da região otimizada para reconhecimento molecular.

4.4.9 Moléculas estruturais de água

Sempre que buscamos o desenvolvimento de novas moléculas bioativas, devemos levar em consideração a importância das moléculas estruturais de água como mediadoras da capacidade de reconhecimento molecular. A possibilidade de incluir moléculas de água explícitas na simulação do reconhecimento molecular é uma das principais vantagens da DM. Neste caso, analisamos a trajetória de DM do mutante SGB1-121_{MD} para localizar este tipo de água estrutural, como podemos ver na Figura 31:



Figura 31 – Ilustração da sobreposição dos *frames* da trajetória de DM do mutante SGB1-121_{MD}. Moléculas de água são coloridas em vermelho no interior do arcaabouço proteico.

Fonte: elaborado pelo Autor (2024).

Na Ilustração da sobreposição dos *frames* da trajetória de DM do mutante SGB1-121 podemos localizar duas moléculas de água (Vermelho) que permanecem nas cavidades formadas entre o mutante SGB1-121 e o domínio RBD. Este resultado mostra claramente um papel mediador na ligação entre moléculas maiores. Após completar esta análise, podemos concluir que o modelo teórico de AM para o RBD é suficientemente confiável para prosseguir para a fase experimental deste estudo.

5 Síntese e validação experimental

Após a convergência do AG, encerramos a parte computacional deste projeto e iniciamos o planejamento da parte experimental. Um dos objetivos do projeto é a testagem experimental dos AM. Conforme explicado na metodologia, o GB1 tem um tamanho de 56 resíduos de aminoácidos, o que o torna muito grande para ser sintetizado quimicamente.

No entanto, o GB1 pode ser facilmente sintetizado pela expressão bacteriana por meio de plasmídeos, como pode ser visto em Bozkurt *et. al* (2018). Portanto, para a parte experimental, foram encomendados plasmídeos já clonados com genes para o mutante SGB1-121_{his-Tag} e o peptídeo original GB1 (GB1-Ref), utilizado como controle de qualidade. Abaixo estão as sequências em formato fasta das duas moléculas utilizadas na parte experimental:

- SGB1-121: MIYQLELRGKTLEGATLTAVDAATAEKVFKQYANDNGVDGYWTYQ-DASKIFYVWEHHHHHH
- GB1-Ref: MTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYD-DATKTFTVTEHHHHHH

As duas sequências são muito semelhantes, diferindo apenas nas regiões dos 15 resíduos modificados pelo AG. Além disso, a sequência mostra o marcador his-tag na região C-terminal. Como observado por Liao, Rabideau e Pentelute (2014), a possibilidade de conjugar outros domínios funcionais nas extremidades C-terminais é outra das vantagens do GB1, o que permite a liberação intracelular de anticorpos miméticos.

Detalhes dos protocolos utilizados na parte experimental, incluindo preparação e transformação celular, expressão e extração de biomassa, purificação cromatográfica e avaliação de afinidade antigênica, podem ser encontrados detalhadamente no APÊNDICE C - Protocolos Experimentais.

5.1 Análise de dicroísmo circular

A análise do dicroísmo circular pode fornecer informações importantes sobre a estrutura secundária, permitindo a investigação do estado de enovelamento dos peptídeos sintetizados. Portanto, amostras purificadas de solução contendo GB1-Ref e SGB1-121_{his-Tag} foram analisadas no aparelho de dicroísmo circular, com três varreduras de 185 nm a 260 nm e largura de banda de 0,5 nm (Bozkurt *et. al*, 2018). Além disso, foi realizada uma previsão teórica do espectro de dicroísmo circular da GB1-Ref fornecida pelo servidor online PDBMD2CD (Drew; Janes, 2020). Os resultados podem ser visualizados na Figura 32:

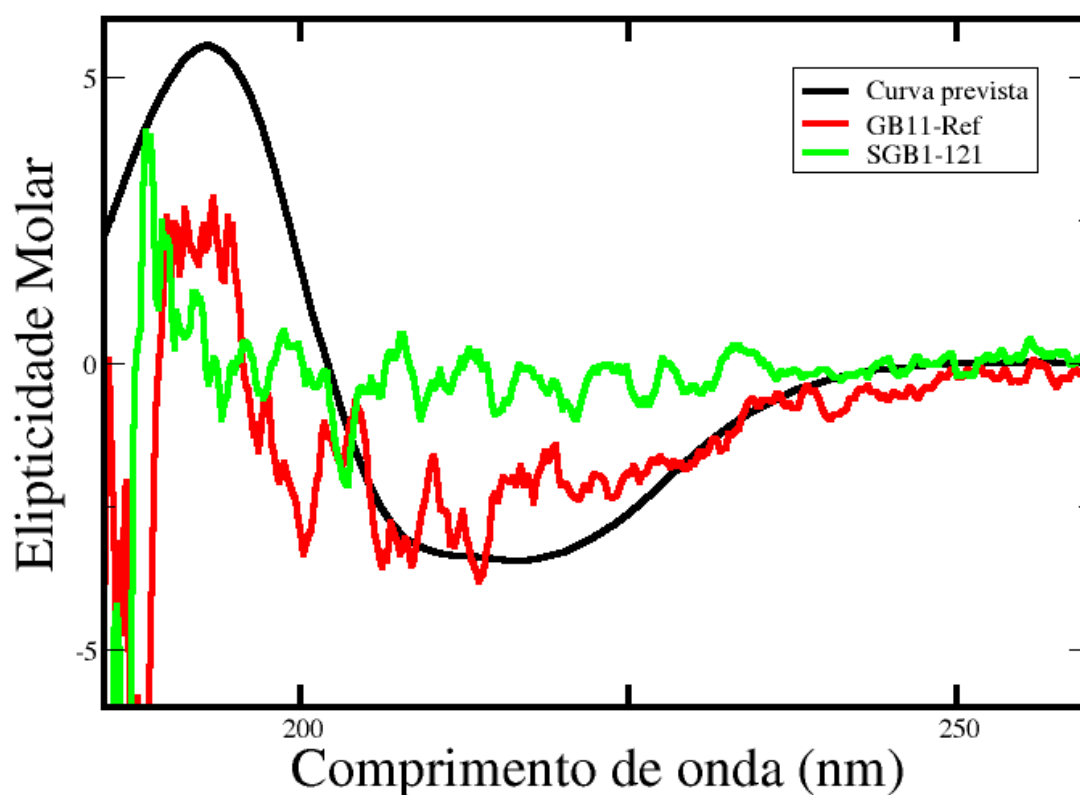


Figura 32 – Gráfico mostrando o espectro de dicroísmo circular das soluções contendo GB1-Ref e SGB1-121 (diluídas 150 vezes), com a previsão teórica fornecida pelo servidor online PDBMD2CD.

Fonte: elaborado pelo Autor (2024).

Inicialmente, as soluções contendo GB1-Ref ($59,5 \text{ ng} \cdot \mu\text{L}^{-1}$) e SGB1-121 ($55,0 \text{ ng} \cdot \mu\text{L}^{-1}$) foram testadas no Dicroísmo Circular. No entanto, devido à concentração muito elevada, não foi possível realizar a medição. Portanto, optamos por diluir as soluções em 150 vezes e 200 vezes, obtendo as concentrações nas quais foram realizadas as medidas.

Na Figura 32, as curvas obtidas para a GB1-Ref e SGB1-121 revelam picos na região de 190-200 nm, seguidos por uma queda na região de 200-225 nm e estabilização em torno de zero após 225 nm. Nenhuma das curvas indica uma região de queda com um mínimo entre 190-210 nm, sugerindo que não há desenovelamento randômico, mas sim um reequilíbrio entre folhas- β e α -hélices. Nesse caso, a GB1-REF parece estar mais próxima de sua curva ideal, uma conclusão corroborada pela análise mais detalhada do gráfico utilizando o servidor online BeStSel (Micsonai et al., 2018).

Segundo o BeStSel, a curva de modelo ideal da GB1 prevista pelo PDBMD2CD indica que para essa estrutura deveria haver 18,8% de α -hélices, 23,4% de fitas- β antiparalelas e 7,3% de fitas- β paralelas. No entanto, ao compararmos as curvas de elipticidade molar das moléculas reais, observamos diferenças significativas. Na curva obtida para a GB1-Ref ($0,39 \text{ ng} \cdot \mu\text{L}^{-1}$), obtemos 17,0% de α -hélices, 20,2% de fitas- β antiparalelas e 9,9% de fitas- β paralelas. Esse resultado é muito próximo do esperado, o que sugere que a GB1-Ref obteve um enovelamento adequado.

Na curva obtida para SGB1-121 ($0,36 \text{ ng} \cdot \mu\text{L}^{-1}$), obtemos 0,0% de α -hélices, 42,2% de fitas- β antiparalelas e 4,7% de fitas- β paralelas. À primeira vista, os resultados indicam que a SGB1-121 não se enovelou na estrutura esperada. No entanto, ao analisarmos novamente a GB1-Ref ($0,29 \text{ ng} \cdot \mu\text{L}^{-1}$), obtemos 0,0% de α -hélices, 37,5% de fitas- β antiparalelas e 0% de fitas- β paralelas.

Esse resultado indica que o dicroísmo circular depende significativamente da descoberta de uma concentração ideal, visto que o aparelho é altamente sensível à diluição. Como obtivemos previsões estruturais muito diferentes ao variar as concentrações, isso levou a resultados duvidosos. Portanto, o experimento de dicroísmo circular demonstra que o GB1-Ref possui uma estrutura muito próxima daquela prevista teoricamente em uma concentração específica. No entanto, não podemos tomar decisões definitivas sobre a estrutura do SGB1-121 devido às limitações no número de concentrações testadas.

5.2 Avaliação de afinidade antigênica

A etapa experimental dos testes de afinidade antigênica dos AM gerados pelo AG foi realizada pela mestrande da FCFAR Aline Reis, que colaborou conosco na parte experimental deste projeto. Para confirmar experimentalmente a afinidade do AM ao RBD, foi realizado o teste cPass adquirido em kit comercial, cujo protocolo é mostrado no APÊNDICE C. O cPass foi aplicado as soluções purificadas de $55.0 \text{ ng} \cdot \mu\text{L}^{-1}$ do mutante SGB1-121_{his-Tag} e $59.5 \text{ ng} \cdot \mu\text{L}^{-1}$ GB1-Ref, utilizado como controle negativo. O resultado preliminar do teste pode ser visto na Figura 33:

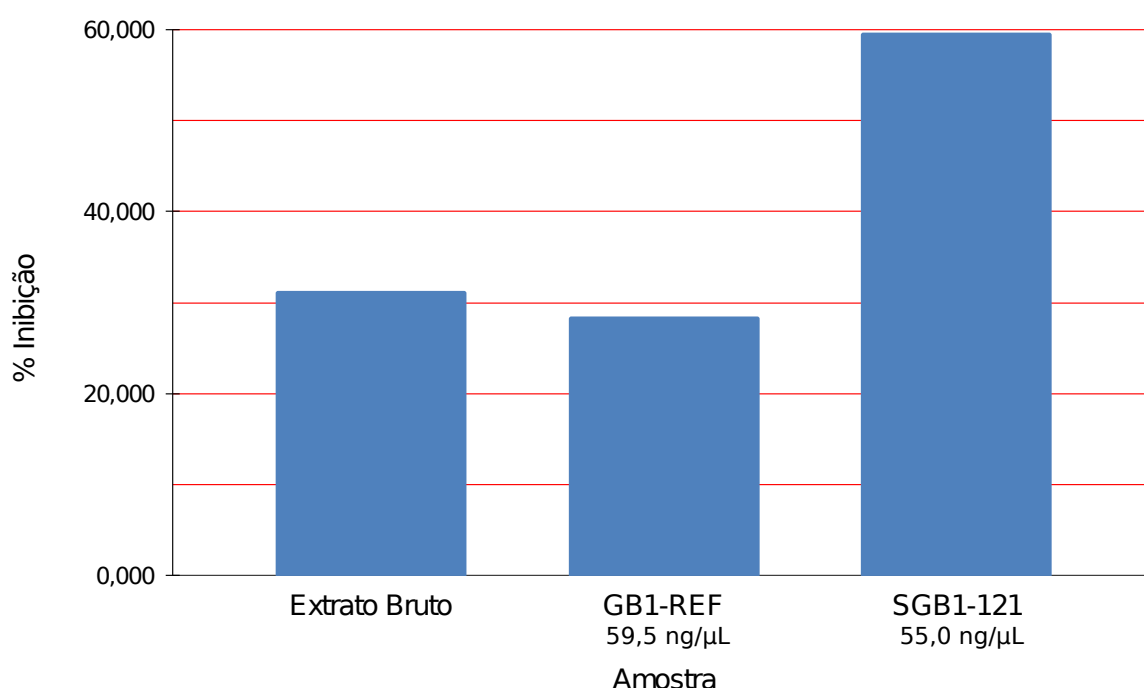


Figura 33 – Gráfico mostrando a porcentagem de inibição do extrato bruto não purificado de *E. Coli* e dos extratos purificados contendo GB1-Ref e SGB1-121. Valores acima de 30% são considerados positivos para inibição.

Fonte: elaborado pelo Autor (2024).

Como pode ser visto na Figura 33, os dados preliminares do cPass indicam porcentagem de inibição de 31.24% para o extrato Bruto de *E. Coli*. Esse resultado mostra como excesso de impurezas na amostra pode gerar resultados enganosos, sendo interpretado como um falso positivo. Para os extratos purificados obtemos porcentagem de inibição de 28.51% para GB1-Ref e 59.64% para o SGB1-121. Conforme detalhado na metodologia, o teste cPass realizado individualmente é apenas qualitativo e valores acima de 30% são considerados positivos para inibição.

Dessa forma, o resultado pode ser considerado positivo para o mutante SGB1-121. No entanto, por se tratar de um teste preliminar, é necessário analisar o resultado

com cautela. Embora ainda não seja possível afirmar com certeza sobre a capacidade neutralizante do mutante SGB1-121, os resultados preliminares são promissores. Neste ponto, devemos destacar a importância de realizar mais testes com diferentes concentrações, aumentando a confiabilidade dos resultados e permitindo, assim, uma conclusão definitiva sobre os experimentos.

6 Observações gerais

A principal vantagem do AG é o seu baixo custo computacional, cada simulação de DM requer uma GPU e pode ser executada em paralelo. Dessa forma, a população pode ser tão grande quanto o número de GPUs disponíveis. Isto permite que o protocolo do AG seja executado rapidamente sob condições favoráveis.

A utilização de um modelo para AM baseado em peptídeo tem a vantagem de não necessitar de parametrização, todos os resíduos já são nativos do campo de força Amber. Isso nos permite aproveitar ao máximo a DM em GPU, que pode rodar centenas de nanossegundos por dia, totalmente controlado por scripts de shell.

O protocolo AG realiza apenas comparação $\Delta\Delta G$, utilizando o método MM-GBSA sob condições ideais de desempenho. Cabe ressaltar que as energias calculadas são uma média de um mínimo local, dentro do intervalo de simulação de 100 ns, porém, os microdomínios são amostrados dezenas de vezes, o que equivale a simulações de milhares de nanossegundos.

A ferramenta de decomposição de energia do MM-GBSA pode fornecer informações sobre os resíduos mais importantes para o reconhecimento molecular, servindo como um marcador para o nível de otimização dos AM dentro do AG. Esta informação pode ser muito útil nos casos onde não existem ligantes naturais para estimar um critério de convergência.

Os mutantes GB1 que não reconhecem o antígeno, acabam se deslocando da região de interesse durante a simulação do DM, sendo eliminados pelo próprio AG devido à energia desfavorável. Da mesma forma, mutantes que sofrem desnaturação, como SGB1-24 (APÊNDICE F - Tabela 45), podem ser eliminados facilmente com análise estrutural. Escolher uma população inicial com energia favorável potencializa o AG, que pode convergir mais rapidamente, como visto no mutante mGB1-11 obtido na segunda geração.

Alterar um único resíduo de aminoácido pode desestabilizar toda a capacidade de reconhecimento molecular, como vemos na mutação CYS8SER. Esta mutação faz com que resíduos do próprio AM interajam entre si, reduzindo a afinidade pelo

antígeno. Isso mostra como os vírus podem escapar do sistema imunológico apenas com mutações pontuais.

As cadeias laterais de aminoácidos vizinhos com propriedades semelhantes, que competem pelo mesmo tipo de sítio, causam perda de seletividade. Um exemplo seria um AM que estabelece apenas ligações de hidrogênio com o antígeno, resultando em alta afinidade e pouca seletividade. Nesse sentido, podemos dizer que o melhor modelo de AM não é aquele que possui a melhor EL, mas sim aquele que realiza o maior número de interações intermoleculares distintas.

A enxertia de motivos estruturais já conhecidos é uma das estratégias mais utilizadas no design de proteínas, mas é limitada por problemas estruturais ao inserir esses motivos no arcabouço de outras proteínas. Na técnica de AG, os motivos estruturais são desenhados no próprio GB1 de forma inovadora, sem depender de motivos estruturais preexistentes.

O AG favorece a formação de motivos estruturais totalmente originais, como K2W4Y6C8, I49N51D53R55, L13D15K17L19 e Q4E6R8E10, fundamentais para o reconhecimento molecular. Esses motivos, formados por vários tipos de resíduos, garantem a alta seletividade dos AM projetados pelo AG.

Os *Softwares* comerciais capazes de realizar aprimoramento de capacidade de reconhecimento molecular dependem de bancos de dados preexistentes para realizar a tarefa. As regiões de grande variabilidade ainda são um desafio para esse tipo de ferramenta computacional. A técnica de AG desenvolvida nesse projeto tem a capacidade de criar seu próprio banco de dados, que pode ser tão grande quanto o número de gerações do AG.

A escolha de um modelo de 56 aminoácidos traz inúmeras vantagens, como facilidade de síntese e maior penetração celular dos peptídeos. Algo impensável para moléculas de tamanho semelhante a anticorpos reais, contendo centenas de aminoácidos e modificações pós-traducionais. A estabilidade dos microdomínios projetados pelo AG poderá ser uma oportunidade para, no futuro, minimizar a estrutura dos AM.

A escolha do antígeno para AM deve levar em consideração o tamanho do epítopo, o conhecimento exato da região de reconhecimento molecular e as estruturas proteicas localizadas em regiões próximas. Isto evita problemas ao criar AM que se ligam fortemente ao epítopo, mas são incapazes de se ligar à proteína completa devido ao impedimento estérico.

O protocolo AG demonstra que ao combinar condições ótimas de operação das técnicas de DM e MM-GBSA, associadas à modularidade e portabilidade, com estruturas experimentalmente confiáveis, é capaz de gerar resultados robustos de forma rápida e abundante.

Neste ponto devemos destacar que o protocolo AG ainda não é totalmente automático, revisado a cada geração do algoritmo genético. Essa prática garante que apenas mutantes com as qualidades desejadas avancem no protocolo.

A metodologia utilizada neste estudo foi desenvolvida durante a execução do AG e obviamente pode ser melhorada. Uma das melhorias mais óbvias seria a possibilidade de utilizar simulações de DM com tempos maiores, ou diversas simulações com o mesmo mutante, para melhor amostrar as possibilidades de conformações, aumentando a confiabilidade do resultado. A análise estrutural e conformacional atualmente realizada visualmente poderia ser automatizada, permitindo a execução contínua do AG.

A ideia geral deste estudo é bastante intuitiva e de fácil compreensão. Na verdade, é apenas uma nova abordagem para usar protocolos básicos de DM de proteínas. O grande diferencial está na estrutura modular em que são realizados os protocolos, na forma de um algoritmo de busca e otimização capaz de reprogramar a estrutura e função do GB1.

Após uma revisão da literatura, podemos afirmar que o artigo publicado por Cao *et al.* (2020) é o mais próximo do estudo realizado neste projeto. Nesse artigo, utilizaram o *software* Rifdock e uma estratégia baseada na escolha dos rotâmeros que melhor se ligam à superfície de um antígeno para criar motivos estruturais. Posteriormente, foi feita uma seleção a partir de um banco de dados de peptídeos cuja estrutura melhor se adaptava ao enxerto desses motivos estruturais. Esse estudo foi validado experimentalmente por Hunt *et al.* (2022).

Esses estudos são muito recentes, publicados após o início desta pesquisa e conforme afirmam Chauhan e Pantazes (2022), o Rifdock ainda não teve sua metodologia publicada e revisada por pares, tendo sido idealizado como um pacote para ser utilizado no Rosetta. A maior inovação proposta neste estudo em relação ao artigo citado é a criação de motivos estruturais diretamente na estrutura dos AM, eliminando a necessidade de utilizar bancos de dados externos.

Deve-se notar que, como ocorre com qualquer outra abordagem computacional, a validação experimental é sempre necessária para comprovar a atividade biológica. Nosso resultado experimental mais importante, obtido no teste de afinidade antigênica do mutante SGB1-121, indica um resultado positivo para neutralização do RBD. No entanto, ainda é necessário agir com cautela devido à necessidade de realizar mais experimentos.

O protocolo desenvolvido neste estudo permite uma formidável portabilidade, dependendo apenas de *softwares* extensivamente usados em DM, além de ser facilmente auditável em todas as etapas, garantindo a transparência necessária à metodologia científica.

7 Conclusão

Neste estudo, apresentamos uma abordagem demonstrando que a modelagem de AM pode ser efetuada utilizando uma combinação de *softwares* amplamente empregados em simulação molecular. Esses *softwares*, quando estruturados sob forma de um AG, podem redefinir o papel do domínio GB1. Tal reconfiguração, alcançada através da reprogramação da sua estrutura e função, possui um potencial significativo para orientar a pesquisa experimental por novas moléculas bioativas. A execução do AG é viável computacionalmente devido à escolha de um arcabouço para AM e epítomos antigênicos com dimensões extremamente pequenos, mas capazes de manter a estabilidade estrutural.

Mostramos que, a partir de bases de dados confiáveis da superfície de um antígeno, usando uma função de pontuação muito simples, o AG pode otimizar a capacidade de reconhecimento molecular de populações de AM em poucas gerações. Nos três casos de estudo apresentados, o AG foi capaz de otimizar a EL das populações de AM para valores melhores do que os ligantes nativos. Essa otimização acontece muito rapidamente, comprovando que a escolha racional das populações iniciais tem potencial de resolver um dos grandes problemas do projeto de moléculas bioativas.

Os resultados computacionais indicam que, utilizando a metodologia empregada neste estudo, é possível desenvolver motivos estruturais capazes de reconhecer antígenos. Esses motivos estruturais podem ser projetados a partir da própria estrutura dos AM, de forma original e inovadora, sem necessidade de *softwares* de predição estrutural, ancoragem molecular ou bancos de dados preexistentes. Os mutantes gerados conciliam a alta estabilidade do GB1 com as propriedades otimizadas dos AM. Portanto, podemos afirmar que o objetivo de desenvolver uma base peptídica estável para conversão em AM foi atingido.

Os testes experimentais de afinidade antigênica mostraram um resultado positivo, confirmando que o anticorpo SGB1-121 foi capaz de neutralizar o antígeno RBD para o qual foi projetado. Esses resultados ainda não são definitivos, demandando novos experimentos com diferentes concentrações. Contudo, a viabilidade da metodologia utilizada neste estudo é evidente. Todos os procedimentos podem ser concluídos em períodos significativamente mais curtos do que a busca convencional em extensas bases de dados de moléculas bioativas. Consequentemente, podemos afirmar que a fase de validação experimental produziu resultados satisfatórios, demonstrando que os objetivos deste estudo foram alcançados.

Referências

ADOLF-BRYFOGLE, Jared; KALYUZHNIY, Oleks; KUBITZ, Michael; WEITZNER, Brian D.; HU, Xiaozhen; ADACHI, Yumiko; SCHIEF, William R.; DUNBRACK, Roland L. RosettaAntibodyDesign (RAbD): a general framework for computational antibody design. **Plos Computational Biology**, v. 14, n. 4, p. 1006112, 27 abr. 2018.

ARYA, Namrata; KAUR, Arminder. Molecular Docking: a review paper. **International Journal of Innovative Research in Engineering Management**, p. 140-146, 1 fev. 2022.

BAROZZI, Annalisa; LAVOIE, Ashton; DAY, Kevin N.; PRODROMOU, Raphael; MENEGATTI, Stefano. Affibody-Binding Ligands. **International Journal of Molecular Sciences**, v. 21, n. 11, p. 3769, 27 maio 2020.

BOZKURT, Esra; ASHARI, Negar; BROWNING, Nicholas; BRUNK, Elizabeth; CAMPOMANESA, Pablo; PEREZ, Marta A. S.; ROTH LISBERGER, Ursula. Lessons from Nature: computational design of biomimetic compounds and processes. **Chimia**, v. 68, n. 9, p. 642, 24 set. 2014.

BOZKURT, Esra; PEREZ, Marta A. S.; HOVIUS, Ruud; BROWNING, Nicholas J.; ROTH LISBERGER, Ursula. Genetic Algorithm Based Design and Experimental Characterization of a Highly Thermostable Metalloprotein. **Journal of the American Chemical Society**, v. 140, n. 13, p. 4517-4521, 16 jan. 2018.

BRUCE, Virginia J.; TA, Angeline N.; MCNAUGHTON, Brian R. Minimalist Antibodies and Mimetics: an update and recent applications. **Chembiochem**, v. 17, n. 20, p. 1892-1899, 4 set. 2016.

CAO, Longxing; GORESHNIK, Inna; COVENTRY, Brian; CASE, James Brett; MILLER, Lauren; KOZODOY, Lisa; CHEN, Rita E.; CARTER, Lauren; WALLS, Alexandra C.; PARK, Young-Jun. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. **Science**, v. 370, n. 6515, p. 426-431, 23 out. 2020.

CASE, David A.; BROZELL, Scott R.; BEN-SHALOM, Ido Y.; CERUTTI, David; CHEATHAM, Thomas E.; CRUZEIRO, Vinícius W. D.; DARDEN, TOM; DUKE, Robert E.; GHOREISHI, Delaram; GILSON, MIKE; GOHLKE, Holger; Goetz, Andreas W.; GREENE, D'Artagnan; HARRIS, Robert; HOMEYER, Nadine; Huang, Yandong; IZADI, Saeed; KOVALENKO, Andriy; KURTZMAN, Tom; LEE, Taisung; LEGRAND, Scott; LI, Pengfei; LIN, Charles; LIU, Jian; LUCHKO, Tyler; LUO, Ray; MERMELSTEIN, Daniel J.; MERZ, Kenneth M.; MIAO, Yinglong; MONARD, Gérald; NGUYEN, Hai; OMELIAN, IGOR; ONUFRIEV, Alexey; PAN, Feng;

SALOMON-FERRER, Romelia; SWAILS, Jason; WALKER, Ross C.; WANG, Junmei; WEI, Haixin; WOLF, Romain M.; WU, Xiongwu; XIAO, Li; YORK, Darrin; KOLLMAN, Peter. **Amber 2018**. versão 2018. University of California: San Francisco, 2018.

CHAUHAN, Varun M.; PANTAZES, Robert J. MutDock: a computational docking approach for fixed-backbone protein scaffold design. **Frontiers in Molecular Biosciences**, v. 9, 29 ago. 2022.

CHEN, Yu; LI, Lanjuan. SARS-CoV-2: virus dynamics and host response. **The Lancet Infectious Diseases**, v. 20, n. 5, p. 515-516, maio 2020.

CHEN, Yu; ZHENG, Yongxiang; FONG, Pedro; MAO, Shengjun; WANG, Qiantao. The application of the MM/GBSA method in the binding pose prediction of FGFR inhibitors. **Physical Chemistry Chemical Physics**, v. 22, n. 17, p. 9656-9663, 2020.

CONWAY, Patrick; TYKA, Michael D.; DIMAIO, Frank; KONERDING, David E.; BAKER, David. Relaxation of backbone bond geometry improves protein energy landscape modeling. **Protein Science**, v. 23, n. 1, p. 47-55, 22 nov. 2013.

DELANO, W. L. *The PyMOL Molecular Graphics System*. Versão 2.5. DeLano Scientific: Palo Alto, 2002.

HUMPHREY, William; DALKE, Andrew; SCHULTEN, Klaus. VMD: visual molecular dynamics. **Journal Of Molecular Graphics**, v. 14, n. 1, p. 33-38, fev. 1996.

DREW, Elliot D; JANES, Robert W. PDBMD2CD: providing predicted protein circular dichroism spectra from multiple molecular dynamics-generated protein structures. **Nucleic Acids Research**, v. 48, n. 1, p. 17-24, 28 abr. 2020.

ELLIS, Lee M.; HICKLIN, Daniel J. VEGF-targeted therapy: mechanisms of anti-tumour activity. **Nature Reviews Cancer**, v. 8, n. 8, p. 579-591, 3 jul. 2008.

ESPIRITO SANTO, Anderson A.; FELICIANO, Gustavo T. Thermodynamic and structural aspects of molecular recognition in mannose-binding protein complexes: a theoretical study over hrp-artinm association. **Journal of Molecular Modeling**, v. 27, n. 4, 15 mar. 2021.

ESPIRITO SANTO, Anderson A.; FELICIANO, Gustavo T. Genetic Algorithms Applied to Thermodynamic Rational Design of Mimetic Antibodies Based on the GB1 Domain of Streptococcal Protein G: an atomistic simulation study. **The Journal of Physical Chemistry B**, v. 125, n. 29, p. 7985-7996, 15 jul. 2021.

FASMAN, G. D. **Circular Dichroism and the Conformational Analysis of Biomolecules**. New York: Kluwer Academic/Plenum, 1996. 738 p.

FREJD, Fredrik y; KIM, Kyu-Tae. Affibody molecules as engineered protein drugs. **Experimental Molecular Medicine**, v. 49, n. 3, p. 306, mar. 2017.

- GELLER, Ron; ESTADA, Úrsula; PERIS, Joan B.; ANDREU, Iván; BOU, Juan-Vicente; GARIJO, Raquel; CUEVAS, José M.; SABARIEGOS, Rosario; MAS, Antonio; SANJUÁN, Rafael. Highly heterogeneous mutation rates in the hepatitis C virus genome. **Nature Microbiology**, v. 1, n. 7, n.p , 18 abr. 2016.
- GENHEDEN, Samuel; KUHN, Oliver; MIKULSKIS, Paulius; HOFFMANN, Daniel; RYDE, Ulf. The Normal-Mode Entropy in the MM/GBSA Method: effect of system truncation, buffer region, and dielectric constant. **Journal of Chemical Information and Modeling**, v. 52, n. 8, p. 2079-2088, 2 ago. 2012.
- GENHEDEN, Samuel; RYDE, Ulf. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. **Expert Opinion on Drug Discovery**, v. 10, n. 5, p. 449-461, 2 abr. 2015.
- GREEN, Michael R., SAMBROOK, Joseph. **Molecular Cloning: A Laboratory Manual**. 4th Ed., Vol. II, New York : Cold Spring Harbor Laboratory Press. 2012. 2028 p.
- GUPTA, Ashish; CHAUDHARY, Neha; APAROY, Polamarasetty. MM-PBSA and per-residue decomposition energy studies on 7-Phenyl-imidazoquinolin-4(5H)-one derivatives: identification of crucial site points at microsomal prostaglandin e synthase-1 (mpges-1) active site. **International Journal of Biological Macromolecules**, v. 119, p. 352-359, nov. 2018.
- GELLMAN, Samuel H . Introduction: molecular recognition. **Chemical Reviews**, v. 97, n. 5, p. 1231-1232, 1 ago. 1997.
- HOLLINGSWORTH, Scott A.; DROR, Ron O . Molecular Dynamics Simulation for All. **Neuron**, v. 99, n. 6, p. 1129-1143, set. 2018.
- HOU, Tingjun; WANG, Junmei; LI, Youyong; WANG, Wei. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. **Journal of Chemical Information and Modeling**, v. 51, n. 1, p. 69-82, 30 nov. 2010.
- HUANG, Sheng-You; GRINTER, Sam Z.; ZOU, Xiaoqin. Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. **Physical Chemistry Chemical Physics**, v. 12, n. 40, p. 12899, 2010.
- HUNT, Andrew C.; CASE, James Brett; PARK, Young-Jun; CAO, Longxing; WU, Kejia; WALLS, Alexandra C.; LIU, Zhuoming; BOWEN, John E.; YEH, Hsien-Wei; SAINI, Shally. Multivalent designed proteins neutralize SARS-CoV-2 variants of concern and confer protection against infection in mice. **Science Translational Medicine**, v. 14, n. 646, 25 maio 2022.
- KADONOSONO, Tetsuya; YIMCHUEN, Wanaporn; OTA, Yumi; SEE, Kyra; FURUTA,

- Tadaomi; SHIOZAWA, Tadashi; KITAZAWA, Maika; GOTO, Yu; PATIL, Akash; KUCHIMARU, Takahiro. Design Strategy to Create Antibody Mimetics Harboring Immobilised Complementarity Determining Region Peptides for Practical Use. **Scientific Reports**, v. 10, n. 1, 21 jan. 2020.
- KUHN, Bernd; GERBER, Paul; SCHULZ-GASCH, Tanja; STAHL, Martin. Validation and Use of the MM-PBSA Approach for Drug Discovery. **Journal of Medicinal Chemistry**, v. 48, n. 12, p. 4040-4048, 10 maio 2005.
- LADNER, Robert C . Mapping the Epitopes of Antibodies. **Biotechnology and Genetic Engineering Reviews**, v. 24, n. 1, p. 1-30, jan. 2007.
- LIAO, Xiaoli; RABIDEAU, Amy E.; PENTELUTE, Bradley L. Delivery of Antibody Mimics into Mammalian Cells via Anthrax Toxin Protective Antigen. **Chembiochem**, v. 15, n. 16, p. 2458-2466, 22 set. 2014.
- LU, Ruei-Min; HWANG, Yu-Chyi; LIU, I-Ju; LEE, Chi-Chiu; TSAI, Han-Zen; LI, Hsin-Jung; WU, Han-Chung. Development of therapeutic antibodies for the treatment of diseases. **Journal of Biomedical Science**, v. 27, n. 1, 2 jan. 2020.
- MAIER, James A.; MARTINEZ, Carmenza; KASAVAJHALA, Koushik; WICKSTROM, Lauren; HAUSER, Kevin E.; SIMMERLING, Carlos. Ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99sb. **Journal of Chemical Theory and Computation**, v. 11, n. 8, p. 3696-3713, 23 jul. 2015.
- MICSONAI, András; WIEN, Frank; BULYÁKI, Éva; KUN, Judit; MOUSSONG, Éva; LEE, Young-Ho; GOTO, Yuji; RÉFRÉGIERS, Matthieu; KARDOS, József. BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. **Nucleic Acids Research**, v. 46, n. 1, p. 315-322, 11 jun. 2018.
- NELSON, David L.; COX, Michael M. **Princípios de bioquímica de Lehninger**. Porto Alegre. 6. ed. Porto Alegre: Artmed, 2014. 1300 p.
- NERATTINI, Francesca; TUBIANA, Luca; CARDELLI, Chiara; BIANCO, Valentino; DELLAGO, Christoph; COLUZZA, Ivan. Design of Protein–Protein Binding Sites Suggests a Rationale for Naturally Occurring Contact Areas. **Journal of Chemical Theory and Computation**, v. 15, n. 2, p. 1383-1392, 11 dez. 2018.
- OLUKITIBI, Titus A.; AO, Zhujun; WARNER, Bryce; UNAT, Rodrigo; KOBASA, Darwyn; YAO, Xiaojian. Significance of Conserved Regions in Coronavirus Spike Protein for Developing a Novel Vaccine against SARS-CoV-2 Infection. **Vaccines**, v. 11, n. 3, p. 545, 24 fev. 2023.
- PETTERSEN, Eric F.; GODDARD, Thomas D.; HUANG, Conrad C.; COUCH, Gregory S.; GREENBLATT, Daniel M.; MENG, Elaine C.; FERRIN, Thomas E . UCSF

- Chimera A visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1605-1612, jul. 2004.
- PIERCE, Brian G.; BOUCHER, Elisabeth N.; PIEPENBRINK, Kurt H.; EJEMEL, Monir; RAPP, Chelsea A.; THOMAS, William D.; SUNDBERG, Eric J.; WENG, Zhiping; WANG, Yang. Structure-Based Design of Hepatitis C Virus Vaccines That Elicit Neutralizing Antibody Responses to a Conserved Epitope. **Journal of Virology**, v. 91, n. 20, 15 out. 2017.
- RAMARAJ, Thiruvarangan; ANGEL, Thomas; DRATZ, Edward A.; JESAITIS, Algirdas J.; MUMEY, Brendan. Antigen–antibody interface properties: composition, residue interactions, and features of 53 non-redundant structures. **Biochimica Et Biophysica Acta (Bba) - Proteins And Proteomics**, v. 1824, n. 3, p. 520-532, mar. 2012.
- SCOTT, Andrew M.; WOLCHOK, Jedd D.; OLD, Lloyd J. Antibody therapy of cancer. **Nature Reviews Cancer**, v. 12, n. 4, p. 278-287, 22 mar. 2012.
- SHEU, Sheh-Yi; YANG, Dah-Yen; SELZLE, H. L.; SCHLAG, E. W. Energetics of hydrogen bonds in peptides. **Proceedings of the National Academy Of Sciences**, v. 100, n. 22, p. 12683-12687, 14 out. 2003.
- SHUAIB, Wan Muhammad Azfar Wan; BADARUDDIN, Izzatul Aliaa; MANSOR, Munirah; SALLEH, Sharifah Azura; HASSAN, Mohd Rohaizat; LINDONG, Steward; SAMAD, Shahril Nizam; OTHMAN, Hanita. SARS-CoV-2 S-RBD IgG Neutralizing antibodies among different categories of health care workers post third dose BNT162b2 mRNA COVID-19 vaccine. **Human Vaccines Immunotherapeutics**, v. 19, n. 3, p. 1, 13 out. 2023.
- SIMEON, Rudo; CHEN, Zhilei. In vitro-engineered non-antibody protein therapeutics. **Protein Cell**, v. 9, n. 1, p. 3-14, 7 mar. 2017.
- SITTHIYOTHA, Thassanai; CHUNSRIVIROT, Surasak. Computational design of SARS-CoV-2 peptide binders with better predicted binding affinities than human ACE2 receptor. **Scientific Reports**, v. 11, n. 1, 2 ago. 2021.
- SPIEGEL, Jacob O.; DURRANT, Jacob D. AutoGrow4: an open-source genetic algorithm for de novo drug design and lead optimization. **Journal of Cheminformatics**, v. 12, n. 1, 17 abr. 2020.
- SU, Chinh Tran-To; NGUYEN, Thuy-Diem; ZHENG, Jie; KWOH, Chee-Keong. IFACEwat: the interfacial water-implemented re-ranking algorithm to improve the discrimination of near native structures for protein rigid docking. **Bmc Bioinformatics**, v. 15, n. 16, dez. 2014.
- VALIENTE, Pedro A.; WEN, Han; NIM, Satra; LEE, Jinah; KIM, Hyeon Ju; KIM,

- Jinhee; PEREZ-RIBA, Albert; PAUDEL, Yagya Prasad; HWANG, Insu; KIM, Kyun-Do. Computational Design of Potent D-Peptide Inhibitors of SARS-CoV-2. **Journal of Medicinal Chemistry**, v. 64, n. 20, p. 14955-14967, 8 out. 2021.
- VOET, Donald; VOET, Judith G. **Bioquímica**. 4 ed. Porto Alegre: Artmed. 2013. 1481 p.
- WANG, Ercheng; SUN, Huiyong; WANG, Junmei; WANG, Zhe; LIU, Hui; ZHANG, John Z. H.; HOU, Tingjun. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. **Chemical Reviews**, v. 119, n. 16, p. 9478-9508, 24 jun. 2019.
- WATERHOUSE, Andrew; BERTONI, Martino; BIENERT, Stefan; STUDER, Gabriel; TAURIELLO, Gerardo; GUMIENNY, Rafal; HEER, Florian T; DEBEER, Tjaart A P; REMPFER, Christine; BORDOLI, Lorenza. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research**, v. 46, n. 1, p. 296-303, 21 maio 2018.
- WHITTEN, Kenneth W.; DAVIS, Raymond E.; PECK, M. Larry; STANLEY, G. G. **Chemistry**. 10. ed. Belmont: Brooks/Cole Cengage Learning, 2014. 1184 p.
- WIESMANN, Christian; FUH, Germaine; CHRISTINGER, Hans W.; EIGENBROT, Charles; WELLS, James A.; VOS, Abraham M. de. Crystal Structure at 1.7 Å Resolution of VEGF in Complex with Domain 2 of the Flt-1 Receptor. **Cell**, v. 91, n. 5, p. 695-704, nov. 1997.
- WOODY, Robert W. Circular dichroism. **Methods in Enzymology**. v. 246, p. 34-71, 1995.
- WU, Yifei; LOU, Lei; XIE, Zhong-Ru. A Pilot Study of All-Computational Drug Design Protocol—From Structure Prediction to Interaction Analysis. **Frontiers in Chemistry**, v. 8, 12 fev. 2020.
- YAN, Yumeng; TAO, Huanyu; HE, Jiahua; HUANG, Sheng-You. The HDock server for integrated protein–protein docking. **Nature Protocols**, v. 15, n. 5, p. 1829-1852, 8 abr. 2020.
- ZAHND, Christian; SPINELLI, Silvia; LUGINBÜHL, Béatrice; AMSTUTZ, Patrick; CAMBILLAU, Christian; PLÜCKTHUN, Andreas. Directed in Vitro Evolution and Crystallographic Analysis of a Peptide-binding Single Chain Antibody Fragment (scFv) with Low Picomolar Affinity. **Journal of Biological Chemistry**, v. 279, n. 18, p. 18870-18877, abr. 2004.

APÊNDICES

APÊNDICE A – Parâmetros da Função de pontuação

Tabela 1 – Parâmetros usados na calibração da função de pontuação de rotâmeros.

Parâmetros da função de pontuação		
Tipo de interação	kcal.mol ⁻¹	Referência
Ligação de hidrogênio	1.5	Sheu <i>et al.</i> (2003)
Dipolo-dipolo	0.95	Whitten (2014, p. 450)
Lennard-Jones	0.17	Su <i>et al.</i> (2014)

Tabela 2 – Ordem dos resíduos de aminoácidos mais bem classificados em cada posição na função de pontuação do rotâmeros

ID\Resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55
mGB1-1	A	N	H	N	I	Y	N	G	G	D	L	L	Q	H	K
mGB1-2	R	Q	R	S	M	W	Q	A	A	E	I	I	N	D	R
mGB1-3	C	K	Q	A	K	N	H	L	L	N	V	H	R	N	L
mGB1-4	Y	H	V	W	W	V	R	I	I	Y	S	W	Y	E	I
mGB1-5	N	I	I	P	R	A	I	V	V	L	T	Q	D	I	W
mGB1-6	Q	R	W	Q	A	L	V	S	S	I	C	D	W	K	N
mGB1-7	E	P	M	R	L	D	W	T	T	K	Y	E	L	Y	Q
mGB1-8	L	L	E	H	H	K	D	C	C	H	N	K	H	L	D
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V
mGB1-10	H	M	K	M	N	I	Y	N	N	A	D	Y	M	W	T

APÊNDICE B – Validação do AG

Para validar a técnica de AG, precisamos demonstrar como o MM-GBSA é capaz de classificar corretamente as energias através do $\Delta\Delta G$. Assim, optamos por analisar um estudo experimental sobre a evolução dirigida de anticorpos (Zahnd *et al.*, 2004). Neste estudo, Zahnd *et al.* (2004) inseriram mutações em um fragmento de anticorpo de cadeia simples (scFv) para melhorar a constante de afinidade. Usando a estrutura cristalográfica (PDB ID: 1P4B), que contém um scFv complexado com o peptídeo GCN4, aplicamos o método MM-GBSA. Neste caso, por termos uma estrutura cristalográfica, para maior precisão do MM-GBSA, utilizamos apenas a estrutura relaxada da etapa de minimização de energia da DM (Chen *et al.*, 2020; Hou *et al.*, 2010). O resultado pode ser visto na Tabela 3.

Tabela 3 – Validação Experimental do AG. Todos os valores de ΔG estão em quilocaloria por mol.

Comparação entre valor teórico e experimental de ΔG		
Tipo de interação	ΔG_{K_a} (Zahnd <i>et al.</i> , 2004)	$\Delta G_{MM-GBSA}$
L107	-14.1 ± 0.1	-55.6163
L135	-14.6 ± 0.2	-57.2596
H6	-14.7 ± 0.1	-57.4534
L107L135	-14.2 ± 0.1	-56.8989
L135H6	-14.8 ± 0.8	-57.7258

Como podemos ver na Tabela 3, há uma grande diferença na magnitude das energias calculadas entre os métodos experimental e teórico. Além dos termos entrópicos desconsiderados pelo MM-GBSA, a energia pode variar muito dependendo do método de parametrização do campo de força usado na DM (Kuhn *et al.*, 2005; Wang *et al.*, 2019). No entanto, como demonstrado por Wu *et al.* (2020), podemos fazer uma comparação muito precisa das diferenças de EL calculadas pelo MM-GBSA (Wu *et al.*, 2020; Hou *et al.*, 2010). O resultado na tabela nos mostra como os mutantes L107L135 e L135H6, resultantes do crossing over dos mutantes L107, L135 e H6, permanecem na ordem correta de classificação energética. Usando apenas uma estrutura relaxada, o MM-GBSA é capaz de ordenar corretamente as energias dos mutantes. Este resultado demonstra que o MM-GBSA tem uma precisão necessária para ser utilizado em AG. Vale ressaltar que o MM-GBSA já é amplamente utilizado para filtrar grupos de moléculas promissoras para o planejamento de fármacos, o que gera muitas críticas quanto ao método de parametrização dessas moléculas. Este problema não existe neste estudo, que usa apenas resíduos de aminoácidos padrão do campo de força Amber, que já foram extensivamente validados pelos desenvolvedores.

APÊNDICE C – Protocolos experimentais

Preparo das células e Transformação

As células da cepa *E. coli* BL21 foram inoculadas em 5 mL de meio Luria-Bertani (LB) 24 horas antes da transformação. O cultivo foi incubado em um agitador orbital a 37°C e 240 rpm durante a noite. Em seguida, 50 μ L do cultivo foram adicionados a 5 mL de meio LB e deixados crescer por 2 horas, até atingir uma densidade óptica de 0,4. Após esse período, as células foram centrifugadas a 4500 rpm por 3 minutos, e o sobrenadante foi descartado. O pellet foi lavado duas vezes com CaCl₂ 0,1 M gelado, ressuspenso na mesma solução e mantido sobre gelo. Após 30 minutos, a suspensão foi centrifugada e o sobrenadante descartado. O pellet foi ressuspenso em 300 μ L de CaCl₂ 0,1 M para transformação. Esse processo foi realizado para duas alíquotas referentes a GB1-Ref e SGB1-121.

As alíquotas das células receberam 50 ng dos plasmídeos (pET-28a/GB1ref ou pET-28a/GB1-121) e foram mantidas sobre gelo por 25 minutos. Em seguida, foram submetidas a um choque térmico a 42 °C por 90 segundos, seguido de retorno ao gelo por 5 minutos.

Após esse processo, cada alíquota recebeu 1 mL de meio LB fresco e foi deixada em um agitador orbital por 1 hora a 240 rpm.

Após o período de recuperação, foi realizado o plaqueamento em meio LB-Ágar contendo 25 mg/mL de Canamicina. As placas foram incubadas a 37°C para formação de colônias.

Expressão e extração de biomassa

Para a expressão do mutante SGB1-121 e da referência GB1, foram utilizados dois inóculos contendo a cepa de *E. coli* BL21 Rosetta transformada com os vetores pET-28a/GB1ref e pET-28a/SGB1-121. Cada inóculo foi cultivado durante a noite em 5 mL de meio LB líquido com canamicina. De cada cultura, foi retirada uma alíquota de 1 mL, adicionada a 50 mL de meio LB líquido com a correspondente concentração de

canamicina. Cada cultura foi incubada em um agitador a 240 rpm e 37°C até atingir uma densidade óptica (O.D.) de 0,4 a 0,6. Uma solução de IPTG foi adicionada a cada cultura até alcançar uma concentração final de 1 mM. Os meios de cultura foram mantidos em indução de expressão por 5 horas a 37°C.

Após a indução da expressão, cada meio foi centrifugado a 1600 x g durante 10 minutos a 10°C, e o sobrenadante foi descartado. A massa celular foi lavada com tampão Tris-EDTA pH 7,4. Em seguida, foi adicionado um tampão de lise contendo inibidor de protease ao pellet. O pellet foi lisado utilizando ultrassom de ponteira por 4 minutos, em intervalos de 30 segundos, em um banho de gelo. O lisado foi centrifugado a 2200 x g durante 10 minutos a 4°C. O sobrenadante resultante foi separado em um novo tubo e armazenado a -20°C até o uso.

Purificação por cromatografia de afinidade

A purificação do AM foi realizada em coluna de níquel utilizando colunas cromatográficas HisTrap™ (Sephacrose TM6 Fast Flow-Ni²⁺), que possuem afinidade por histidinas his-tag 6 fundidas às extremidades C-terminais dos AM. As colunas foram previamente lavadas com água ultrapura para remover a solução de estocagem. Em seguida, para preparar o leito cromatográfico, foram introduzidos 25 mL de solução de ligação. A amostra contendo o sobrenadante foi acrescida de 20 mM de imidazol. Esta fração de sobrenadante foi filtrada por uma membrana de 0,22 µm e aplicada na coluna. As proteínas não ligadas foram removidas por meio de uma lavagem com 50 mL de solução de ligação. As proteínas ligadas foram eluídas com a adição de 50 mL de solução de eluição (500 mM de imidazol). Após a eluição, a coluna foi lavada com água ultrapura e armazenada em etanol 20%. O perfil cromatográfico foi construído por meio de amostras de 5 mL eluídas da coluna. A concentração de proteínas foi medida pelo kit comercial Micro BCATM Protein Assay (Thermo Scientific). Assim, 100 µL de cada amostra foram adicionados aos poços de microplaca, com 100 µL do reagente de trabalho. Após a agitação por 30 segundos, a placa foi vedada e incubada a 37°C por 2 horas. Após a incubação, a absorbância foi medida a 562 nm utilizando o leitor de placas GO accuSkan (Fisher Scientific).

Avaliação de afinidade antigênica

Para testes experimentais de afinidade antigênica, foi utilizado o kit comercial cPass™ SARS-CoV-2 Neutralization Antibody Detection, uma ferramenta destinada a

detecção de anticorpos do tipo ELISA por competição. Este método utiliza o domínio de ligação RBD e o receptor ACE2. Este teste foi projetado para simular o mecanismo de interação entre o vírus e o receptor ACE2 na placa. Esta interação pode ser bloqueada na presença de anticorpos inibidores de RBD. No caso deste projeto, foi avaliada a capacidade de inibição do mutante SGB1-121.

O protocolo começa com a pré-incubação da proteína quimérica HRP-RBD para permitir a ligação de anticorpos inibitórios. Esta mistura é adicionada a uma placa de captura revestida com o domínio da proteína ACE2. Desta forma, os conjugados HRP-RBD que não são neutralizados pelos anticorpos são ligados à placa, enquanto os conjugados HRP-RBD que foram neutralizados são removidos durante o processo de lavagem.

Após o processo de lavagem, uma solução de substrato cromogênico (TMB) é adicionada à placa, tornando-a azul na presença da HRP. Quando a solução de parada é adicionada, a reação é interrompida, tornando-se amarela. A solução final é analisada a 450 nm no leitor de placas de microtitulação. A absorbância medida é inversamente dependente do título de anticorpos neutralizantes RBD.

APÊNDICE D – Energias do AG para os mutantes mGB1 anti-FLT-1

Tabela 4 – Primeira geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-8	L	L	E	H	H	K	D	C	C	H	N	K	H	L	D	-29.8
mGB1-2	R	Q	R	S	M	W	Q	A	A	E	I	I	N	D	R	-29.4
mGB1-7	E	P	M	R	L	D	W	T	T	K	Y	E	L	Y	Q	-23.4
mGB1-10	H	M	K	M	N	I	Y	N	N	A	D	Y	M	W	T	-14.3
mGB1-5	N	I	I	P	R	A	I	V	V	L	T	Q	D	I	W	-13.4
mGB1-4	Y	H	V	W	W	V	R	I	I	Y	S	W	Y	E	I	-10.0
mGB1-1	A	N	H	N	I	Y	N	G	G	D	L	L	Q	H	K	-8.1
mGB1-3	C	K	Q	A	K	N	H	L	L	N	V	H	R	N	L	1.2
mGB1-6	Q	R	W	Q	A	L	V	S	S	I	C	D	W	K	N	1.3

Tabela 5 – Mutantes filhos da primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
mGB1-12	E	P	M	R	L	D	W	T	A	E	I	I	N	D	R	-21.6
mGB1-13	N	I	I	P	R	A	I	V	N	A	D	Y	M	W	T	-16.6
mGB1-14	L	L	E	H	H	K	D	C	Y	M	Q	V	I	Q	V	-15.0
mGB1-15	R	Q	R	S	M	W	Q	A	T	K	Y	E	L	Y	Q	-12.0
mGB1-16	Q	R	W	Q	A	L	V	S	L	N	V	H	R	N	L	-11.9
mGB1-17	A	N	H	N	I	Y	N	G	I	Y	S	W	Y	E	I	-11.0
mGB1-18	C	K	Q	A	K	N	H	L	S	I	C	D	W	K	N	-7.91
mGB1-19	Y	H	V	W	W	V	R	I	G	D	L	L	Q	H	K	-6.2
mGB1-20	H	M	K	M	N	I	Y	N	V	L	T	Q	D	I	W	1.8

Tabela 6 – Segunda geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-8	L	L	E	H	H	K	D	C	C	H	N	K	H	L	D	-29.8
mGB1-2	R	Q	R	S	M	W	Q	A	A	E	I	I	N	D	R	-29.4
mGB1-7	E	P	M	R	L	D	W	T	T	K	Y	E	L	Y	Q	-23.4
mGB1-12	E	P	M	R	L	D	W	T	A	E	I	I	N	D	R	-21.6
mGB1-13	N	I	I	P	R	A	I	V	N	A	D	Y	M	W	T	-16.6
mGB1-14	L	L	E	H	H	K	D	C	Y	M	Q	V	I	Q	V	-15.0
mGB1-10	H	M	K	M	N	I	Y	N	N	A	D	Y	M	W	T	-14.3
mGB1-5	N	I	I	P	R	A	I	V	V	L	T	Q	D	I	W	-13.4

Tabela 7 – Mutantes filhos da segunda geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-21	E	P	M	R	L	D	W	T	T	K	Y	I	N	D	R	-33.1
mGB1-22	K	W	Y	C	G	H	E	Y	C	H	N	V	I	Q	V	-24.2
mGB1-23	E	P	M	R	L	D	W	T	A	E	I	E	L	Y	Q	-21.2
mGB1-24	K	W	Y	C	G	H	E	Y	Y	M	Q	K	H	L	D	-21.1
mGB1-25	N	I	I	P	N	I	Y	N	V	L	T	Y	M	W	T	-17.5
mGB1-26	N	I	I	P	R	A	I	V	Y	M	Q	V	I	Q	V	-16.3
mGB1-27	L	L	E	H	H	K	D	C	A	E	I	I	N	D	R	-14.5
mGB1-28	R	Q	R	S	M	W	Q	A	C	H	N	K	H	L	D	-6.8
mGB1-29	H	M	K	M	R	A	I	V	N	A	D	Q	D	I	W	-6.3
mGB1-30	L	L	E	H	H	K	D	C	N	A	D	Y	M	W	T	-3.7

Tabela 8 – Terceira geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-21	E	P	M	R	L	D	W	T	T	K	Y	I	N	D	R	-33.1
mGB1-8	L	L	E	H	H	K	D	C	C	H	N	K	H	L	D	-29.8
mGB1-2	R	Q	R	S	M	W	Q	A	A	E	I	I	N	D	R	-29.4
mGB1-22	K	W	Y	C	G	H	E	Y	C	H	N	V	I	Q	V	-24.2
mGB1-7	E	P	M	R	L	D	W	T	T	K	Y	E	L	Y	Q	-23.4
mGB1-12	E	P	M	R	L	D	W	T	A	E	I	I	N	D	R	-21.6
mGB1-23	E	P	M	R	L	D	W	T	A	E	I	E	L	Y	Q	-21.2
mGB1-24	K	W	Y	C	G	H	E	Y	Y	M	Q	K	H	L	D	-21.1

Tabela 9 – Mutantes filhos da terceira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-31	R	Q	R	S	M	W	Q	A	C	H	N	V	I	Q	V	-30.3
mGB1-32	K	W	Y	C	G	H	E	Y	C	H	N	K	I	L	V	-24.7
mGB1-33	K	W	Y	C	G	H	E	Y	A	E	I	I	N	D	R	-21.7
mGB1-34	E	P	M	R	L	D	W	T	A	E	I	I	L	D	Q	-18.2
mGB1-35	E	P	M	R	L	D	W	T	T	K	Y	E	N	Y	R	-17.9
mGB1-36	K	W	Y	C	G	H	E	Y	A	E	I	E	L	Y	Q	-17.1
mGB1-37	K	W	Y	C	G	H	E	Y	Y	M	Q	V	H	Q	D	-11.6
mGB1-38	L	L	E	H	H	K	D	C	T	K	Y	I	N	D	R	-3.2
mGB1-39	E	P	M	R	L	D	W	T	C	H	N	K	H	L	D	-1.8
mGB1-40	E	P	M	R	L	D	W	T	Y	M	Q	K	H	L	D	0.8

Tabela 10 – Quarta geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔH
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-21	E	P	M	R	L	D	W	T	T	K	Y	I	N	D	R	-33.1
mGB1-31	R	Q	R	S	M	W	Q	A	C	H	N	V	I	Q	V	-30.3
mGB1-8	L	L	E	H	H	K	D	C	C	H	N	K	H	L	D	-29.8
mGB1-2	R	Q	R	S	M	W	Q	A	A	E	I	I	N	D	R	-29.4
mGB1-32	K	W	Y	C	G	H	E	Y	C	H	N	K	I	L	V	-24.7
mGB1-22	K	W	Y	C	G	H	E	Y	C	H	N	V	I	Q	V	-24.2
mGB1-7	E	P	M	R	L	D	W	T	T	K	Y	E	L	Y	Q	-23.4
mGB1-33	K	W	Y	C	G	H	E	Y	A	E	I	I	N	D	R	-21.7

Tabela 11 – Mutantes filhos da quarta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-41	L	L	E	H	M	W	Q	A	C	H	N	I	N	D	R	-35.2
mGB1-42	K	W	Y	S	G	H	E	Y	Y	M	Q	V	I	Q	V	-31.8
mGB1-43	E	P	M	R	L	D	W	T	C	H	N	V	I	Q	V	-25.1
mGB1-44	K	W	Y	S	G	H	E	Y	C	H	N	V	I	Q	V	-19.9
mGB1-45	K	W	Y	S	G	H	E	Y	C	H	N	K	I	L	V	-19.7
mGB1-46	K	W	Y	S	G	H	E	Y	C	H	N	K	H	L	D	-15.5
mGB1-47	E	P	M	R	G	H	E	Y	T	K	Y	I	N	D	R	-14.9
mGB1-48	R	Q	R	S	M	W	Q	A	T	K	Y	I	N	D	R	-12.0
mGB1-49	K	W	Y	C	L	D	W	T	A	E	I	E	L	Y	Q	-10.6
mGB1-50	R	Q	R	S	H	K	D	C	A	E	I	K	H	L	D	1.2

Tabela 12 – Quinta geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-41	L	L	E	H	M	W	Q	A	C	H	N	I	N	D	R	-35.1
mGB1-21	E	P	M	R	L	D	W	T	T	K	Y	I	N	D	R	-33.1
mGB1-42	K	W	Y	S	G	H	E	Y	Y	M	Q	V	I	Q	V	-31.8
mGB1-31	R	Q	R	S	M	W	Q	A	C	H	N	V	I	Q	V	-30.3
mGB1-8	L	L	E	H	H	K	D	C	C	H	N	K	H	L	D	-29.8
mGB1-2	R	Q	R	S	M	W	Q	A	A	E	I	I	N	D	R	-29.4
mGB1-43	E	P	M	R	L	D	W	T	C	H	N	V	I	Q	V	-25.1
mGB1-32	K	W	Y	C	G	H	E	Y	C	H	N	K	I	L	V	-24.7

Tabela 13 – Mutantes filhos da quinta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-51	K	W	Y	V	G	H	E	Y	Y	M	Q	V	I	Q	V	-40.8
mGB1-52	L	L	E	H	M	W	Q	A	T	K	Y	I	N	D	R	-36.2
mGB1-53	K	W	Y	S	M	W	Q	A	Y	M	Q	V	I	Q	V	-32.5
mGB1-54	R	Q	R	S	H	W	D	A	A	E	I	I	H	D	D	-28.2
mGB1-55	K	W	Y	V	G	H	E	Y	C	H	N	K	H	L	D	-20.4
mGB1-56	R	Q	R	S	G	H	E	Y	C	H	N	V	I	Q	V	-20.2
mGB1-57	L	L	E	H	M	K	Q	C	C	H	N	K	N	L	R	-16.6
mGB1-58	E	P	M	R	L	D	W	T	C	H	N	I	N	D	R	-14.4
mGB1-59	E	P	M	R	G	H	E	Y	C	H	N	K	I	L	V	-8.0
mGB1-60	K	W	Y	C	L	D	W	T	C	H	N	V	I	Q	V	-6.1

Tabela 14 – Sexta geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
mGB1-51	K	W	Y	V	G	H	E	Y	Y	M	Q	V	I	Q	V	-40.8
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-52	L	L	E	H	M	W	Q	A	T	K	Y	I	N	D	R	-36.2
mGB1-41	L	L	E	H	M	W	Q	A	C	H	N	I	N	D	R	-35.1
mGB1-21	E	P	M	R	L	D	W	T	T	K	Y	I	N	D	R	-33.1
mGB1-54	K	W	Y	S	M	W	Q	A	Y	M	Q	V	I	Q	V	-32.5
mGB1-42	K	W	Y	S	G	H	E	Y	Y	M	Q	V	I	Q	V	-31.8
mGB1-31	R	Q	R	S	M	W	Q	A	C	H	N	V	I	Q	V	-30.3
mGB1-8	L	L	E	H	H	K	D	C	C	H	N	K	H	L	D	-29.8

Tabela 15 – Mutantes filhos da sexta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-61	K	W	Y	C	G	H	E	Y	T	K	Y	I	N	D	R	-37.1
mGB1-62	R	Q	R	S	H	K	D	C	C	H	N	K	H	L	D	-29.0
mGB1-63	L	L	E	H	L	D	W	T	C	H	N	I	N	D	R	-26.3
mGB1-64	K	W	Y	C	G	H	E	Y	Y	H	Q	V	H	Q	D	-26.1
mGB1-65	L	L	E	H	M	W	Q	A	Y	M	Q	V	I	Q	V	-15.6
mGB1-66	L	L	E	H	M	W	Q	A	C	H	N	V	I	Q	V	-14.8
mGB1-67	K	W	Y	S	M	H	Q	Y	Y	M	Q	V	I	Q	V	-13.6
mGB1-68	K	W	Y	S	G	W	E	A	Y	M	Q	V	I	Q	V	-13.2
mGB1-69	E	P	M	R	M	W	Q	A	T	K	Y	I	N	D	R	-8.7
mGB1-70	K	W	Y	V	G	H	E	Y	C	M	N	K	I	L	V	-3.2

Tabela 16 – Sétima geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
mGB1-51	K	W	Y	V	G	H	E	Y	Y	M	Q	V	I	Q	V	-40.8
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-61	K	W	Y	C	G	H	E	Y	T	K	Y	I	N	D	R	-37.1
mGB1-52	L	L	E	H	M	W	Q	A	T	K	Y	I	N	D	R	-36.2
mGB1-41	L	L	E	H	M	W	Q	A	C	H	N	I	N	D	R	-35.1
mGB1-21	E	P	M	R	L	D	W	T	T	K	Y	I	N	D	R	-33.1
mGB1-54	K	W	Y	S	M	W	Q	A	Y	M	Q	V	I	Q	V	-32.5
mGB1-42	K	W	Y	S	G	H	E	Y	Y	M	Q	V	I	Q	V	-31.8
mGB1-31	R	Q	R	S	M	W	Q	A	C	H	N	V	I	Q	V	-30.3

Tabela 17 – Mutantes filhos da sétima geração de AM. A todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-71	L	L	E	H	M	W	Q	A	C	K	N	I	N	D	R	-45.0
mGB1-72	L	L	E	H	M	W	Q	A	T	H	Y	I	N	D	R	-41.3
mGB1-73	E	P	M	R	L	D	W	T	Y	M	Q	V	I	Q	V	-32.0
mGB1-74	R	W	R	S	G	W	E	A	C	M	N	V	I	Q	V	-29.0
mGB1-75	K	Q	Y	S	M	H	Q	Y	Y	H	Q	V	I	Q	V	-22.7
mGB1-76	K	W	Y	S	M	W	Q	A	T	K	Y	I	N	D	R	-20.4
mGB1-77	K	W	Y	C	G	H	E	Y	Y	M	Q	I	N	D	R	-14.8
mGB1-78	K	W	Y	C	G	H	E	Y	C	H	Q	V	H	Q	D	-14.5
mGB1-79	K	W	Y	C	G	H	E	Y	T	K	Y	V	I	Q	V	-11.5
mGB1-80	K	W	Y	V	G	H	E	Y	Y	M	N	K	I	L	V	-7.0

Tabela 18 – Oitava geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
mGB1-71	L	L	E	H	M	W	Q	A	C	K	N	I	N	D	R	-45.0
mGB1-72	L	L	E	H	M	W	Q	A	T	H	Y	I	N	D	R	-41.3
mGB1-51	K	W	Y	V	G	H	E	Y	Y	M	Q	V	I	Q	V	-40.8
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-61	K	W	Y	C	G	H	E	Y	T	K	Y	I	N	D	R	-37.1
mGB1-52	L	L	E	H	M	W	Q	A	T	K	Y	I	N	D	R	-36.2
mGB1-41	L	L	E	H	M	W	Q	A	C	H	N	I	N	D	R	-35.1
mGB1-21	E	P	M	R	L	D	W	T	T	K	Y	I	N	D	R	-33.1
mGB1-54	K	W	Y	S	M	W	Q	A	Y	M	Q	V	I	Q	V	-32.5

Tabela 19 – Mutantes filhos da oitava geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-81	E	P	M	R	M	W	Q	A	T	K	Y	V	I	Q	V	-40.3
mGB1-82	K	W	Y	S	L	D	W	T	Y	M	Q	I	N	D	R	-37.0
mGB1-83	L	L	E	H	M	W	Q	A	C	H	Y	I	N	D	R	-36.7
mGB1-84	K	W	Y	V	M	W	Q	A	Y	M	Q	I	N	D	R	-35.9
mGB1-85	K	W	Y	C	G	H	E	Y	C	K	N	I	N	D	R	-28.8
mGB1-86	K	W	Y	C	G	H	E	Y	T	M	Y	I	I	D	V	-21.4
mGB1-87	K	W	Y	C	G	H	E	Y	Y	K	Q	V	N	Q	R	-17.8
mGB1-88	L	L	E	H	M	W	Q	A	C	H	N	K	H	L	D	-12.2
mGB1-89	L	L	E	H	G	H	E	Y	T	H	Y	V	I	Q	V	-11.8
mGB1-90	L	L	E	H	M	W	Q	A	T	K	N	I	N	D	R	-9.2

Tabela 20 – Nona geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
mGB1-71	L	L	E	H	M	W	Q	A	C	K	N	I	N	D	R	-45.0
mGB1-72	L	L	E	H	M	W	Q	A	T	H	Y	I	N	D	R	-41.3
mGB1-51	K	W	Y	V	G	H	E	Y	Y	M	Q	V	I	Q	V	-40.8
mGB1-81	E	P	M	R	M	W	Q	A	T	K	Y	V	I	Q	V	-40.3
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-61	K	W	Y	C	G	H	E	Y	T	K	Y	I	N	D	R	-37.1
mGB1-82	K	W	Y	S	L	D	W	T	Y	M	Q	I	N	D	R	-37.0
mGB1-83	L	L	E	H	M	W	Q	A	C	H	Y	I	N	D	R	-36.7
mGB1-52	L	L	E	H	M	W	Q	A	T	K	Y	I	N	D	R	-36.2

Tabela 21 – Mutantes filhos da nona geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-91	K	W	Y	C	L	D	W	T	T	K	Y	I	N	D	R	-48.3
mGB1-92	K	W	Y	C	M	W	Q	A	C	H	N	I	N	D	R	-47.3
mGB1-93	L	L	E	H	M	W	Q	A	S	H	Y	I	N	D	R	-31.1
mGB1-94	E	P	M	R	G	H	E	Y	T	K	Y	V	I	Q	V	-19.5
mGB1-95	K	W	Y	C	M	W	Q	A	Y	M	Q	V	I	Q	V	-17.8
mGB1-96	L	L	E	H	M	W	Q	A	C	K	Y	I	N	D	R	-16.6
mGB1-97	K	W	Y	S	G	H	E	Y	Y	M	Q	I	N	D	R	-16.2
mGB1-98	L	W	E	V	G	W	E	A	T	M	Y	I	I	D	V	-14.6
mGB1-99	L	L	E	H	G	H	E	Y	C	K	N	K	H	L	D	-11.6
mGB1-100	K	L	Y	H	M	H	Q	Y	Y	H	Q	V	N	Q	R	-10.6

Tabela 22 – Décima geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
mGB1-91	K	W	Y	C	L	D	W	T	T	K	Y	I	N	D	R	-48.3
mGB1-92	K	W	Y	C	M	W	Q	A	C	H	N	I	N	D	R	-47.3
mGB1-71	L	L	E	H	M	W	Q	A	C	K	N	I	N	D	R	-45.0
mGB1-72	L	L	E	H	M	W	Q	A	T	H	Y	I	N	D	R	-41.3
mGB1-51	K	W	Y	V	G	H	E	Y	Y	M	Q	V	I	Q	V	-40.8
mGB1-81	E	P	M	R	M	W	Q	A	T	K	Y	V	I	Q	V	-40.3
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5
mGB1-61	K	W	Y	C	G	H	E	Y	T	K	Y	I	N	D	R	-37.1
mGB1-82	K	W	Y	S	L	D	W	T	Y	M	Q	I	N	D	R	-37.0

Tabela 23 – Mutantes filhos da décima geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-101	L	W	E	C	M	W	Q	A	C	K	N	I	N	D	R	-50.1
mGB1-102	K	W	Y	C	G	D	E	T	T	K	Q	I	N	D	R	-48.7
mGB1-103	K	W	Y	C	G	D	E	T	T	H	Y	I	H	D	D	-36.9
mGB1-104	K	W	Y	C	L	H	W	Y	C	K	N	K	N	L	R	-36.6
mGB1-105	E	W	M	C	G	W	E	A	Y	M	Y	V	I	Q	V	-29.3
mGB1-106	K	L	Y	H	M	W	Q	A	C	H	N	I	N	D	R	-28.3
mGB1-107	L	W	E	V	G	W	E	A	Y	H	Y	I	I	D	V	-25.4
mGB1-108	K	P	Y	R	M	H	Q	Y	T	K	Q	V	I	Q	V	-9.6
mGB1-109	K	W	Y	S	L	H	W	Y	Y	M	Y	I	N	D	R	-7.1
mGB1-110	K	L	Y	H	M	H	Q	Y	T	M	Q	V	N	Q	R	-5.0

Tabela 24 – Décima primeira geração de AM anti-FLT-1 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-11	K	W	Y	C	G	H	E	Y	C	H	N	K	H	L	D	-56.4
mGB1-101	L	W	E	C	M	W	Q	A	C	K	N	I	N	D	R	-50.1
mGB1-102	K	W	Y	C	G	D	E	T	T	K	Q	I	N	D	R	-48.7
mGB1-91	K	W	Y	C	L	D	W	T	T	K	Y	I	N	D	R	-48.3
mGB1-92	K	W	Y	C	M	W	Q	A	C	H	N	I	N	D	R	-47.3
mGB1-71	L	L	E	H	M	W	Q	A	C	K	N	I	N	D	R	-45.0
mGB1-72	L	L	E	H	M	W	Q	A	T	H	Y	I	N	D	R	-41.3
mGB1-51	K	W	Y	V	G	H	E	Y	Y	M	Q	V	I	Q	V	-40.8
mGB1-81	E	P	M	R	M	W	Q	A	T	K	Y	V	I	Q	V	-40.3
VEGF	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-39.4
mGB1-9	K	W	Y	C	G	H	E	Y	Y	M	Q	V	I	Q	V	-38.5

Tabela 25 – Mutantes filhos da décima primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
mGB1-111	K	W	Y	C	M	H	Q	Y	C	K	N	K	N	L	R	-40.1
mGB1-112	K	W	Y	C	G	D	E	T	T	K	Y	I	N	D	R	-37.2
mGB1-113	K	P	Y	R	M	H	Q	Y	T	M	Q	V	I	Q	V	-34.8
mGB1-114	K	W	E	C	M	W	Q	A	C	K	N	I	N	D	R	-34.7
mGB1-115	L	L	Y	H	M	W	Q	A	C	H	N	I	N	D	R	-26.9
mGB1-116	E	W	M	C	G	W	E	A	Y	K	Y	V	I	Q	V	-24.6
mGB1-117	L	W	E	V	G	W	E	A	Y	H	Q	I	I	D	V	-17.5
mGB1-118	L	W	E	C	G	W	E	A	C	H	N	I	H	D	D	-14.8
mGB1-119	K	L	Y	H	M	H	Q	Y	T	M	Y	V	N	Q	R	-13.2
mGB1-120	K	W	Y	C	L	D	W	T	T	K	Q	I	N	D	R	-11.8

APÊNDICE E – Energias do AG para os mutantes mGB1 anti-HC-E2

Tabela 27 – Primeira geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-1	L	H	V	L	G	K	M	R	W	W	M	Y	L	M	Y	-31.3
HC-E2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-27.5
HCmGB1-2	M	Y	M	K	A	W	V	A	A	R	W	K	I	H	K	-26.9
HCmGB1-4	W	V	K	Y	I	E	I	I	I	Y	V	W	R	V	I	-26.2
HCmGB1-7	Y	E	E	T	T	L	D	T	T	I	A	T	M	N	T	-20.6
HCmGB1-5	K	K	W	E	V	A	P	V	V	K	E	V	H	T	V	-18.6
HCmGB1-3	H	W	L	W	L	D	K	L	L	H	K	H	Y	A	W	-17.9
HCmGB1-9	R	L	H	A	Y	Y	H	Y	Y	L	L	G	K	E	G	-13.9
HCmGB1-10	N	T	N	N	N	N	Y	N	N	N	I	N	Q	K	N	-12.7
HCmGB1-8	E	R	Y	C	C	M	L	C	C	M	H	C	N	D	C	-2.3
HCmGB1-6	I	M	I	S	S	H	Q	S	S	E	R	S	W	C	S	1.0

Tabela 28 – Mutantes filhos da primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-11	H	W	L	W	L	D	K	L	V	K	E	V	H	T	V	-31.1
HCmGB1-12	M	Y	M	K	A	W	V	A	W	W	M	Y	L	M	Y	-21.7
HCmGB1-13	L	H	V	L	G	K	M	R	A	R	W	K	I	H	K	-20.3
HCmGB1-14	E	R	Y	C	C	M	L	C	S	E	R	S	W	C	S	-19.5
HCmGB1-15	I	M	I	S	S	H	Q	S	C	M	H	C	N	D	C	-15.4
HCmGB1-16	K	K	W	E	V	A	P	V	L	H	K	H	Y	A	W	-14.3
HCmGB1-17	W	V	K	Y	I	E	I	I	T	I	A	T	M	N	T	-10.9
HCmGB1-18	N	T	N	N	N	N	Y	N	Y	L	L	G	K	E	G	-7.6
HCmGB1-19	Y	E	E	T	T	L	D	T	I	Y	V	W	R	V	I	-6.1
HCmGB1-20	R	L	H	A	Y	Y	H	Y	N	N	I	N	Q	K	N	-2.1

Tabela 29 – Segunda geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-1	L	H	V	L	G	K	M	R	W	W	M	Y	L	M	Y	-31.3
HCmGB1-11	H	W	L	W	L	D	K	L	V	K	E	V	H	T	V	-31.1
HC-E2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-27.5
HCmGB1-2	M	Y	M	K	A	W	V	A	A	R	W	K	I	H	K	-26.9
HCmGB1-4	W	V	K	Y	I	E	I	I	I	Y	V	W	R	V	I	-26.2
HCmGB1-12	M	Y	M	K	A	W	V	A	W	W	M	Y	L	M	Y	-21.7
HCmGB1-7	Y	E	E	T	T	L	D	T	T	I	A	T	M	N	T	-20.6
HCmGB1-13	L	H	V	L	G	K	M	R	A	R	W	K	I	H	K	-20.3
HCmGB1-14	E	R	Y	C	C	M	L	C	S	E	R	S	W	C	S	-19.5
HCmGB1-5	K	K	W	E	V	A	P	V	V	K	E	V	H	T	V	-18.6
HCmGB1-3	H	W	L	W	L	D	K	L	L	H	K	H	Y	A	W	-17.9

Tabela 30 – Mutantes filhos da segunda geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-21	H	W	L	W	L	D	K	L	W	W	M	Y	L	M	Y	-34.7
HCmGB1-22	M	Y	M	K	A	W	V	A	I	Y	V	W	R	V	I	-23.7
HCmGB1-23	K	K	W	E	L	D	K	L	V	K	E	H	Y	A	W	-22.3
HCmGB1-24	H	W	L	W	V	A	P	V	L	H	K	V	H	T	V	-21.4
HCmGB1-25	E	R	Y	C	C	M	L	C	A	R	W	K	I	H	K	-18.3
HCmGB1-26	L	H	V	L	G	K	M	R	S	E	R	S	W	C	S	-16.4
HCmGB1-27	L	H	V	L	G	K	M	R	V	K	E	V	H	T	V	-16.0
HCmGB1-28	W	V	K	Y	I	E	I	I	A	R	W	K	I	H	K	-13.8
HCmGB1-29	Y	E	E	T	T	L	D	T	W	W	M	Y	L	M	Y	-11.5
HCmGB1-30	M	Y	M	K	A	W	V	A	T	I	A	T	M	N	T	0.7

Tabela 31 – Terceira geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-21	H	W	L	W	L	D	K	L	W	W	M	Y	L	M	Y	-34.7
HCmGB1-1	L	H	V	L	G	K	M	R	W	W	M	Y	L	M	Y	-31.3
HCmGB1-11	H	W	L	W	L	D	K	L	V	K	E	V	H	T	V	-31.1
HC-E2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-27.5
HCmGB1-2	M	Y	M	K	A	W	V	A	A	R	W	K	I	H	K	-26.9
HCmGB1-2	W	V	K	Y	I	E	I	I	I	Y	V	W	R	V	I	-26.2
HCmGB1-22	M	Y	M	K	A	W	V	A	I	Y	V	W	R	V	I	-23.7
HCmGB1-23	K	K	W	E	L	D	K	L	V	K	E	H	Y	A	W	-22.3
HCmGB1-12	M	Y	M	K	A	W	V	A	W	W	M	Y	L	M	Y	-21.7
HCmGB1-24	H	W	L	W	V	A	P	V	L	H	K	V	H	T	V	-21.4
HCmGB1-7	Y	E	E	T	T	L	D	T	T	I	A	T	M	N	T	-20.6

Tabela 32 – Mutantes filhos da terceira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-31	K	K	W	E	L	D	K	L	W	W	M	Y	L	M	Y	-39.9
HCmGB1-32	W	V	K	Y	A	W	V	A	I	Y	V	W	R	V	I	-39.2
HCmGB1-33	H	W	L	W	G	K	M	R	W	W	M	Y	L	M	Y	-18.1
HCmGB1-34	L	H	V	L	L	D	K	L	W	W	M	Y	L	M	Y	-11.3
HCmGB1-35	M	Y	M	K	I	E	I	I	I	Y	V	W	R	V	I	-11.0
HCmGB1-36	H	W	L	W	V	A	P	V	T	I	A	T	M	N	T	-6.0
HCmGB1-37	M	Y	M	K	A	W	V	A	V	K	E	V	H	T	V	-5.9
HCmGB1-38	H	W	L	W	L	D	K	L	A	R	W	K	I	H	K	0.4
HCmGB1-39	M	Y	M	K	A	W	V	A	V	K	E	H	Y	A	W	0.4
HCmGB1-40	Y	E	E	T	T	L	D	T	L	H	K	V	H	T	V	1.1

Tabela 33 – Quarta geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-31	K	K	W	E	L	D	K	L	W	W	M	Y	L	M	Y	-39.9
HCmGB1-32	W	V	K	Y	A	W	V	A	I	Y	V	W	R	V	I	-39.2
HCmGB1-21	H	W	L	W	L	D	K	L	W	W	M	Y	L	M	Y	-34.7
HCmGB1-1	L	H	V	L	G	K	M	R	W	W	M	Y	L	M	Y	-31.3
HCmGB1-11	H	W	L	W	L	D	K	L	V	K	E	V	H	T	V	-31.1
HC-E2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-27.5
HCmGB1-2	M	Y	M	K	A	W	V	A	A	R	W	K	I	H	K	-26.9
HCmGB1-4	W	V	K	Y	I	E	I	I	I	Y	V	W	R	V	I	-26.2
HCmGB1-22	M	Y	M	K	A	W	V	A	I	Y	V	W	R	V	I	-23.7
HCmGB1-23	K	K	W	E	L	D	K	L	V	K	E	H	Y	A	W	-22.3
HCmGB1-12	M	Y	M	K	A	W	V	A	W	W	M	Y	L	M	Y	-21.7

Tabela 34 – Mutantes filhos da quarta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-41	W	V	K	Y	A	W	V	A	W	W	M	Y	L	M	Y	-14.1
HCmGB1-42	K	K	W	E	L	D	K	L	I	Y	V	W	R	V	I	-8.4
HCmGB1-43	H	H	L	L	G	D	M	L	W	W	M	Y	L	M	Y	-12.8
HCmGB1-44	L	W	V	W	L	K	K	R	W	W	M	Y	L	M	Y	-10.8
HCmGB1-45	H	W	L	W	A	W	V	A	V	K	E	K	I	H	K	-16.7
HCmGB1-46	M	Y	M	K	L	D	K	L	A	R	W	V	H	T	V	-40.6
HCmGB1-47	W	Y	K	K	A	E	V	I	I	Y	V	W	R	V	I	-0.4
HCmGB1-48	M	V	M	Y	I	W	I	A	I	Y	V	W	R	V	I	-20.6
HCmGB1-49	K	K	W	E	A	W	V	A	V	K	E	Y	L	M	Y	-1.0
HCmGB1-50	M	Y	M	K	L	D	K	L	W	W	M	H	Y	A	W	-16.9

Tabela 35 – Quinta geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-41	M	Y	M	K	L	D	K	L	A	R	W	V	H	T	V	-40.6
HCmGB1-31	K	K	W	E	L	D	K	L	W	W	M	Y	L	M	Y	-39.9
HCmGB1-32	W	V	K	Y	A	W	V	A	I	Y	V	W	R	V	I	-39.2
HCmGB1-21	H	W	L	W	L	D	K	L	W	W	M	Y	L	M	Y	-34.7
HCmGB1-1	L	H	V	L	G	K	M	R	W	W	M	Y	L	M	Y	-31.3
HCmGB1-11	H	W	L	W	L	D	K	L	V	K	E	V	H	T	V	-31.1
HC-E2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-27.5
HCmGB1-2	M	Y	M	K	A	W	V	A	A	R	W	K	I	H	K	-26.9
HCmGB1-4	W	V	K	Y	I	E	I	I	I	Y	V	W	R	V	I	-26.2
HCmGB1-22	M	Y	M	K	A	W	V	A	I	Y	V	W	R	V	I	-23.7
HCmGB1-23	K	K	W	E	L	D	K	L	V	K	E	H	Y	A	W	-22.3

Tabela 36 – Mutantes filhos da quinta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-51	M	Y	M	K	I	E	I	I	A	R	W	W	R	V	I	-27.8
HCmGB1-52	K	K	W	E	A	W	V	A	V	K	E	W	R	V	I	-26.1
HCmGB1-53	M	Y	M	K	L	D	K	L	W	W	M	Y	L	M	Y	-24.0
HCmGB1-54	M	Y	M	K	L	D	K	L	I	Y	V	H	Y	A	W	-23.0
HCmGB1-55	W	V	K	Y	L	D	K	L	I	Y	V	Y	L	M	Y	-22.8
HCmGB1-56	L	H	V	L	L	D	K	L	W	W	M	V	H	T	V	-21.7
HCmGB1-57	H	W	L	W	G	K	M	R	V	K	E	Y	L	M	Y	-21.2
HCmGB1-58	K	K	W	E	L	D	K	L	A	R	W	V	H	T	V	-12.9
HCmGB1-59	H	W	L	W	A	W	V	A	W	W	M	W	R	V	I	-5.6
HCmGB1-60	W	V	K	Y	A	W	V	A	I	Y	V	K	I	H	K	-0.9

Tabela 37 – Sexta geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-41	M	Y	M	K	L	D	K	L	A	R	W	V	H	T	V	-40.6
HCmGB1-31	K	K	W	E	L	D	K	L	W	W	M	Y	L	M	Y	-39.9
HCmGB1-32	W	V	K	Y	A	W	V	A	I	Y	V	W	R	V	I	-39.2
HCmGB1-21	H	W	L	W	L	D	K	L	W	W	M	Y	L	M	Y	-34.7
HCmGB1-1	L	H	V	L	G	K	M	R	W	W	M	Y	L	M	Y	-31.3
HCmGB1-11	H	W	L	W	L	D	K	L	V	K	E	V	H	T	V	-31.1
HCmGB1-51	M	Y	M	K	I	E	I	I	A	R	W	W	R	V	I	-27.8
HC-E2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-27.5
HCmGB1-2	M	Y	M	K	A	W	V	A	A	R	W	K	I	H	K	-26.9
HCmGB1-4	W	V	K	Y	I	E	I	I	I	Y	V	W	R	V	I	-26.2
HCmGB1-22	K	K	W	E	A	W	V	A	V	K	E	W	R	V	I	-26.1

Tabela 38 – Mutantes filhos da sexta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-61	M	Y	M	K	L	D	K	L	A	R	W	Y	L	M	Y	-37.1
HCmGB1-62	K	K	W	E	L	D	K	L	W	W	M	V	H	T	V	-2.4
HCmGB1-63	W	W	K	W	L	W	K	A	I	W	V	W	L	V	Y	-21.1
HCmGB1-64	H	V	L	Y	A	D	V	L	W	Y	M	Y	R	M	I	-15.9
HCmGB1-65	L	W	V	W	L	K	K	R	W	K	M	Y	H	M	V	0.6
HCmGB1-66	H	H	L	L	G	D	M	L	V	W	E	V	L	T	Y	-23.0
HCmGB1-67	M	Y	M	K	A	W	V	A	A	R	W	W	R	V	I	-26.7
HCmGB1-68	M	Y	M	K	I	E	I	I	A	R	W	K	I	H	K	-14.5
HCmGB1-69	K	K	W	E	A	W	V	A	I	Y	V	W	R	V	I	-8.2
HCmGB1-70	W	V	K	Y	I	E	I	I	V	K	E	W	R	V	I	-1.4

Tabela 39 – Sétima geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-41	M	Y	M	K	L	D	K	L	A	R	W	V	H	T	V	-40.6
HCmGB1-31	K	K	W	E	L	D	K	L	W	W	M	Y	L	M	Y	-39.9
HCmGB1-32	W	V	K	Y	A	W	V	A	I	Y	V	W	R	V	I	-39.2
HCmGB1-61	M	Y	M	K	L	D	K	L	A	R	W	Y	L	M	Y	-37.1
HCmGB1-21	H	W	L	W	L	D	K	L	W	W	M	Y	L	M	Y	-34.7
HCmGB1-1	L	H	V	L	G	K	M	R	W	W	M	Y	L	M	Y	-31.3
HCmGB1-11	H	W	L	W	L	D	K	L	V	K	E	V	H	T	V	-31.1
HCmGB1-51	M	Y	M	K	I	E	I	I	A	R	W	W	R	V	I	-27.8
HC-E2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-27.5
HCmGB1-2	M	Y	M	K	A	W	V	A	A	R	W	K	I	H	K	-26.9
HCmGB1-4	M	Y	M	K	A	W	V	A	A	R	W	W	R	V	I	-26.7

Tabela 40 – Mutantes filhos da sétima geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-71	H	W	L	W	L	D	K	L	A	R	W	W	R	V	I	-34.6
HCmGB1-72	K	Y	W	K	L	D	K	L	W	R	M	Y	H	M	V	-26.5
HCmGB1-73	M	Y	M	K	A	W	V	A	A	R	W	W	I	V	K	-19.1
HCmGB1-74	L	H	L	L	G	D	M	L	W	W	M	Y	L	M	Y	-14.3
HCmGB1-75	H	W	V	W	L	K	K	R	W	W	M	Y	L	M	Y	-13.7
HCmGB1-76	M	K	M	E	L	D	K	L	A	W	W	V	L	T	Y	-12.5
HCmGB1-77	W	V	K	Y	A	W	V	A	A	R	W	Y	L	M	Y	-11.6
HCmGB1-78	M	Y	M	K	A	W	V	A	A	R	W	K	R	H	I	-8.2
HCmGB1-79	M	Y	M	K	L	D	K	L	I	Y	V	W	R	V	I	-1.3
HCmGB1-80	M	Y	M	K	I	E	I	I	V	K	E	V	H	T	V	0.9

Tabela 41 – Oitava geração de AM anti-HC-E2 mGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-41	M	Y	M	K	L	D	K	L	A	R	W	V	H	T	V	-40.6
HCmGB1-31	K	K	W	E	L	D	K	L	W	W	M	Y	L	M	Y	-39.9
HCmGB1-32	W	V	K	Y	A	W	V	A	I	Y	V	W	R	V	I	-39.2
HCmGB1-61	M	Y	M	K	L	D	K	L	A	R	W	Y	L	M	Y	-37.1
HCmGB1-21	H	W	L	W	L	D	K	L	W	W	M	Y	L	M	Y	-34.7
HCmGB1-71	H	W	L	W	L	D	K	L	A	R	W	W	R	V	I	-34.6
HCmGB1-1	L	H	V	L	G	K	M	R	W	W	M	Y	L	M	Y	-31.3
HCmGB1-11	H	W	L	W	L	D	K	L	V	K	E	V	H	T	V	-31.1
HCmGB1-51	M	Y	M	K	I	E	I	I	A	R	W	W	R	V	I	-27.8
HC-E2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-27.5
HCmGB1-2	M	Y	M	K	A	W	V	A	A	R	W	K	I	H	K	-26.9

Tabela 42 – Mutantes filhos da oitava geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
HCmGB1-81	M	Y	M	K	A	E	V	I	A	R	W	W	I	V	K	-46.4
HCmGB1-82	H	W	L	W	L	D	K	L	W	W	M	W	R	V	I	-37.7
HCmGB1-83	L	W	V	W	L	K	K	R	V	K	M	Y	H	M	V	-35.3
HCmGB1-84	H	W	L	W	L	D	K	L	A	R	W	Y	L	M	Y	-31.8
HCmGB1-85	M	V	M	Y	A	D	V	L	A	Y	W	Y	R	M	I	-27.5
HCmGB1-86	K	Y	W	K	L	D	K	L	A	R	M	Y	H	M	V	-26.8
HCmGB1-87	H	H	L	L	G	D	M	L	W	W	E	V	L	T	Y	-22.5
HCmGB1-88	M	Y	M	K	I	W	I	A	A	R	W	K	R	H	I	-20.0
HCmGB1-89	M	K	M	E	L	D	K	L	W	W	W	V	L	T	Y	-16.6
HCmGB1-90	W	Y	K	K	L	W	K	A	I	R	V	W	L	V	Y	-14.7

APÊNDICE F – Energias do AG para os mutantes SGB1 anti-RBD

Tabela 44 – Primeira geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-2	E	I	Y	T	M	L	Y	I	L	I	W	I	A	R	M	-34.0
SGB1-5	I	M	E	N	K	Y	H	E	N	L	R	K	S	E	L	-32.8
SGB1-10	N	K	T	A	L	W	N	S	T	H	N	Q	L	V	I	-32.6
SGB1-12	S	N	S	H	T	I	A	G	P	N	P	P	R	S	S	-30.3
SGB1-11	T	T	V	Y	I	M	D	T	W	S	A	H	P	T	H	-27.4
SGB1-7	H	R	L	E	V	D	V	Q	R	Q	M	N	Y	H	N	-26.0
SGB1-4	W	H	R	I	R	T	K	W	Y	K	Q	S	I	Y	W	-24.4
SGB1-1	Y	L	W	M	E	A	L	L	Q	W	H	A	G	W	Y	-23.8
SGB1-8	Q	Q	Q	R	H	E	T	M	K	D	I	D	N	Q	Q	-21.3
SGB1-6	R	E	I	Q	W	N	W	R	A	T	E	M	T	I	T	-21.0
SGB1-9	M	S	N	L	P	H	E	V	S	V	L	G	E	M	R	-16.4
SGB1-3	K	V	H	V	A	V	R	Y	V	Y	Y	L	V	K	K	-12.4

Tabela 45 – Mutantes filhos da primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-13	Y	L	W	M	E	A	L	L	Y	K	Q	S	I	Y	W	-36.5
SGB1-14	Q	Q	Q	R	H	E	T	M	A	T	E	M	T	I	T	-36.1
SGB1-15	E	I	Y	T	M	L	Y	I	N	L	R	K	S	E	L	-26.2
SGB1-16	S	N	S	H	T	I	A	G	T	H	N	Q	L	V	I	-26.0
SGB1-17	I	M	E	N	K	Y	H	E	L	I	W	I	A	R	M	-24.8
SGB1-18	T	T	V	Y	I	M	D	T	R	Q	M	N	Y	H	N	-23.8
SGB1-19	H	R	L	E	V	D	V	Q	W	S	A	H	P	T	H	-21.4
SGB1-20	K	V	H	V	A	V	R	Y	S	V	L	G	E	M	R	-14.5
SGB1-21	M	S	N	L	P	H	E	V	V	Y	Y	L	V	K	K	-8.2
SGB1-22	R	E	I	Q	W	N	W	R	K	D	I	D	N	Q	Q	-7.9
SGB1-23	W	H	R	I	R	T	K	W	Q	W	H	A	G	W	Y	-0.7
SGB1-24	N	K	T	A	L	W	N	S	P	N	P	P	R	S	S	-41.2

Tabela 46 – Segunda geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-13	Y	L	W	M	E	A	L	L	Y	K	Q	S	I	Y	W	-36.5
SGB1-14	Q	Q	Q	R	H	E	T	M	A	T	E	M	T	I	T	-36.1
SGB1-2	E	I	Y	T	M	L	Y	I	L	I	W	I	A	R	M	-34.0
SGB1-5	I	M	E	N	K	Y	H	E	N	L	R	K	S	E	L	-32.8
SGB1-10	N	K	T	A	L	W	N	S	T	H	N	Q	L	V	I	-32.6
SGB1-12	S	N	S	H	T	I	A	G	P	N	P	P	R	S	S	-30.3
SGB1-11	T	T	V	Y	I	M	D	T	W	S	A	H	P	T	H	-27.4
SGB1-15	E	I	Y	T	M	L	Y	I	N	L	R	K	S	E	L	-26.2
SGB1-16	S	N	S	H	T	I	A	G	T	H	N	Q	L	V	I	-26.0
SGB1-7	H	R	L	E	V	D	V	Q	R	Q	M	N	Y	H	N	-26.0
SGB1-4	I	M	E	N	K	Y	H	E	L	I	W	I	A	R	M	-24.8
SGB1-1	W	H	R	I	R	T	K	W	Y	K	Q	S	I	Y	W	-24.4

Tabela 47 – Mutantes filhos da segunda geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-28	E	I	Y	T	K	Y	H	E	L	I	W	K	S	E	L	-41.7
SGB1-29	I	M	E	N	K	Y	H	E	Y	K	Q	S	I	Y	W	-35.8
SGB1-30	Q	Q	Q	R	H	E	T	M	Y	K	Q	S	I	Y	W	-33.5
SGB1-31	S	N	S	H	T	I	A	G	R	Q	M	N	Y	H	N	-33.4
SGB1-32	N	K	T	A	T	I	A	G	T	H	N	P	R	S	S	-32.8
SGB1-33	E	I	Y	T	M	L	Y	I	W	S	A	H	P	T	H	-31.6
SGB1-34	W	H	R	I	R	T	K	W	L	I	W	I	A	R	M	-27.3
SGB1-35	T	T	V	Y	I	M	D	T	N	L	R	K	S	E	L	-19.9
SGB1-36	H	R	L	E	V	D	V	Q	T	H	N	Q	L	V	I	-16.9

Tabela 48 – Terceira geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-28	E	I	Y	T	K	Y	H	E	L	I	W	K	S	E	L	-41.7
SGB1-13	Y	L	W	M	E	A	L	L	Y	K	Q	S	I	Y	W	-36.5
SGB1-14	Q	Q	Q	R	H	E	T	M	A	T	E	M	T	I	T	-36.1
SGB1-29	I	M	E	N	K	Y	H	E	Y	K	Q	S	I	Y	W	-35.8
SGB1-2	E	I	Y	T	M	L	Y	I	L	I	W	I	A	R	M	-34.0
SGB1-30	Q	Q	Q	R	H	E	T	M	Y	K	Q	S	I	Y	W	-33.5
SGB1-31	S	N	S	H	T	I	A	G	R	Q	M	N	Y	H	N	-33.4
SGB1-32	N	K	T	A	T	I	A	G	T	H	N	P	R	S	S	-32.8
SGB1-5	I	M	E	N	K	Y	H	E	N	L	R	K	S	E	L	-32.8

Tabela 49 – Mutantes filhos da terceira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-37	Q	Q	Q	R	H	E	T	M	R	Q	M	N	Y	H	N	-41.3
SGB1-38	E	I	Y	T	K	Y	H	E	A	T	E	M	T	I	T	-35.1
SGB1-39	Y	L	W	M	E	A	L	L	L	I	W	K	S	E	L	-32.6
SGB1-40	I	M	E	N	K	Y	H	E	T	H	N	P	R	S	S	-30.4
SGB1-41	E	I	Y	T	K	Y	H	E	L	I	W	S	I	Y	W	-25.2
SGB1-42	Y	L	W	M	H	E	T	M	Y	K	Q	M	T	I	T	-22.6
SGB1-43	S	N	S	H	L	W	N	S	N	L	R	I	A	R	M	-21.1
SGB1-44	Q	Q	Q	R	E	A	L	L	A	T	E	S	I	Y	W	-19.1
SGB1-45	N	K	T	A	T	I	A	G	N	L	R	K	S	E	L	-18.1
SGB1-46	I	M	E	N	M	L	Y	I	Y	K	Q	I	A	R	M	-17.6
SGB1-47	I	M	E	N	M	L	Y	I	P	N	P	Q	L	V	I	-17.3
SGB1-48	S	N	S	H	T	I	A	G	Y	K	Q	S	I	Y	W	-17.2

Tabela 50 – Quarta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-28	E	I	Y	T	K	Y	H	E	L	I	W	K	S	E	L	-41.7
SGB1-37	Q	Q	Q	R	H	E	T	M	R	Q	M	N	Y	H	N	-41.3
SGB1-13	Y	L	W	M	E	A	L	L	Y	K	Q	S	I	Y	W	-36.5
SGB1-14	Q	Q	Q	R	H	E	T	M	A	T	E	M	T	I	T	-36.1
SGB1-29	I	M	E	N	K	Y	H	E	Y	K	Q	S	I	Y	W	-35.8
SGB1-38	E	I	Y	T	K	Y	H	E	A	T	E	M	T	I	T	-35.1
SGB1-2	E	I	Y	T	M	L	Y	I	L	I	W	I	A	R	M	-34.0
SGB1-30	Q	Q	Q	R	H	E	T	M	Y	K	Q	S	I	Y	W	-33.5
SGB1-31	S	N	S	H	T	I	A	G	R	Q	M	N	Y	H	N	-33.4

Tabela 51 – Mutantes filhos da quarta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5
SGB1-50	Y	L	W	M	H	E	T	M	Y	K	Q	N	Y	H	N	-41.0
SGB1-51	I	M	E	N	H	E	T	M	Y	K	Q	M	T	I	T	-35.7
SGB1-52	Q	Q	Q	R	T	I	A	G	Y	K	Q	N	Y	H	N	-34.0
SGB1-53	S	N	S	H	M	L	Y	I	P	N	P	I	A	R	M	-28.7
SGB1-54	Y	L	W	M	K	Y	H	E	A	T	E	K	S	E	L	-26.3
SGB1-55	I	M	E	N	L	W	N	S	N	L	R	Q	L	V	I	-22.4
SGB1-56	Q	Q	Q	R	K	Y	H	E	A	T	E	S	I	Y	W	-19.2
SGB1-57	S	N	S	H	H	E	T	M	R	Q	M	S	I	Y	W	-22.3
SGB1-58	E	I	Y	T	E	A	L	L	L	I	W	M	T	I	T	-17.9
SGB1-59	E	I	Y	T	K	Y	H	E	L	I	W	I	A	R	M	-3.0
SGB1-60	E	I	Y	T	M	L	Y	I	A	T	E	M	T	I	T	-0.4

Tabela 52 – Quinta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-28	E	I	Y	T	K	Y	H	E	L	I	W	K	S	E	L	-41.7
SGB1-37	Q	Q	Q	R	H	E	T	M	R	Q	M	N	Y	H	N	-41.3
SGB1-50	Y	L	W	M	H	E	T	M	Y	K	Q	N	Y	H	N	-41.0
SGB1-13	Y	L	W	M	E	A	L	L	Y	K	Q	S	I	Y	W	-36.5
SGB1-14	Q	Q	Q	R	H	E	T	M	A	T	E	M	T	I	T	-36.1
SGB1-29	I	M	E	N	K	Y	H	E	Y	K	Q	S	I	Y	W	-35.8
SGB1-51	I	M	E	N	H	E	T	M	Y	K	Q	M	T	I	T	-35.7
SGB1-38	E	I	Y	T	K	Y	H	E	A	T	E	M	T	I	T	-35.1

Tabela 53 – Mutantes filhos da quinta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-61	I	Q	E	R	H	Y	T	E	Y	T	Q	S	T	Y	T	-43.2
SGB1-62	Y	L	W	M	H	E	T	M	Y	K	Q	S	I	Y	W	-41.2
SGB1-63	Y	L	W	M	E	A	L	L	N	L	R	I	A	R	M	-38.8
SGB1-64	Y	L	W	M	E	A	L	L	Y	K	Q	N	Y	H	N	-38.2
SGB1-65	Q	Q	Q	R	E	A	L	L	P	N	P	Q	L	V	I	-36.7
SGB1-66	E	I	Y	T	K	Y	H	E	R	Q	M	N	Y	H	N	-34.7
SGB1-67	Q	Q	Q	R	H	E	T	M	L	I	W	K	S	E	L	-28.0
SGB1-68	I	M	E	N	H	E	T	M	A	T	E	M	T	I	T	-27.3
SGB1-69	I	M	E	N	M	L	Y	I	A	T	E	M	T	I	T	-24.0
SGB1-70	S	N	S	H	L	W	N	S	R	Q	M	S	I	Y	W	-20.3
SGB1-71	E	I	Y	T	K	Y	H	E	Y	K	Q	M	T	I	T	-17.3
SGB1-72	Q	M	Q	N	K	E	H	M	A	K	E	M	I	I	W	-8.0

Tabela 54 – Sexta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-61	I	Q	E	R	H	Y	T	E	Y	T	Q	S	T	Y	T	-43.2
SGB1-28	E	I	Y	T	K	Y	H	E	L	I	W	K	S	E	L	-41.7
SGB1-37	Q	Q	Q	R	H	E	T	M	R	Q	M	N	Y	H	N	-41.3
SGB1-62	Y	L	W	M	H	E	T	M	Y	K	Q	S	I	Y	W	-41.2
SGB1-50	Y	L	W	M	H	E	T	M	Y	K	Q	N	Y	H	N	-41.0
SGB1-63	Y	L	W	M	E	A	L	L	N	L	R	I	A	R	M	-38.8
SGB1-64	Y	L	W	M	E	A	L	L	Y	K	Q	N	Y	H	N	-38.2
SGB1-65	Q	Q	Q	R	E	A	L	L	P	N	P	Q	L	V	I	-36.7

Tabela 55 – Mutantes filhos da sexta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-74	I	Q	E	R	H	Y	T	E	L	I	W	K	S	E	L	-38.2
SGB1-75	Y	L	W	M	E	A	L	L	P	N	P	Q	L	V	I	-30.0
SGB1-76	Y	L	W	M	E	A	L	L	Y	K	Q	I	A	R	M	-26.5
SGB1-77	Q	L	Q	M	H	E	T	M	R	K	M	N	I	H	W	-26.4
SGB1-78	S	N	S	H	E	A	L	L	P	N	P	S	I	Y	W	-26.2
SGB1-79	Y	L	W	M	M	L	Y	I	A	T	E	I	A	R	M	-19.4
SGB1-80	E	I	Y	T	K	Y	H	E	Y	T	Q	S	T	Y	T	-17.6
SGB1-81	Q	Q	Q	R	E	A	L	L	Y	K	Q	N	Y	H	N	-17.4
SGB1-82	Y	L	W	M	H	E	T	M	N	L	R	N	Y	H	N	-10.4
SGB1-83	I	M	E	N	E	A	L	L	N	L	R	M	T	I	T	-3.8
SGB1-84	Q	Q	Q	R	L	W	N	S	R	Q	M	Q	L	V	I	0.1

Tabela 56 – Sétima geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-61	I	Q	E	R	H	Y	T	E	Y	T	Q	S	T	Y	T	-43.2
SGB1-28	E	I	Y	T	K	Y	H	E	L	I	W	K	S	E	L	-41.7
SGB1-37	Q	Q	Q	R	H	E	T	M	R	Q	M	N	Y	H	N	-41.3
SGB1-62	Y	L	W	M	H	E	T	M	Y	K	Q	S	I	Y	W	-41.2
SGB1-50	Y	L	W	M	H	E	T	M	Y	K	Q	N	Y	H	N	-41.0
SGB1-63	Y	L	W	M	E	A	L	L	N	L	R	I	A	R	M	-38.8
SGB1-74	I	Q	E	R	H	Y	T	E	L	I	W	K	S	E	L	-38.2

Tabela 57 – Mutantes filhos da sétima geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
SGB1-86	E	Q	Y	R	H	Y	T	E	L	Q	W	K	Y	E	N	-45.7
SGB1-87	I	M	E	N	E	A	L	L	N	L	R	S	I	Y	W	-42.3
SGB1-88	Q	Q	Q	R	M	L	Y	I	R	Q	M	I	A	R	M	-31.7
SGB1-89	Y	L	W	M	H	E	T	M	Y	K	Q	S	I	H	W	-28.7
SGB1-89	I	Q	E	R	E	A	L	L	Y	T	Q	M	T	I	T	-28.2
SGB1-91	Y	L	W	M	H	Y	T	E	A	T	E	S	T	Y	T	-21.9
SGB1-92	S	N	S	H	H	E	T	M	P	N	P	S	Y	Y	N	-21.1
SGB1-93	Y	Q	W	R	H	A	T	L	N	I	R	I	S	R	L	-22.0
SGB1-94	Q	I	Q	T	K	E	H	M	R	I	M	N	S	H	L	-20.9
SGB1-95	I	L	E	M	E	Y	L	E	L	L	W	K	A	E	M	-11.4
SGB1-96	Y	L	W	M	H	E	T	M	Y	K	Q	N	Y	Y	N	-7.0

Tabela 58 – Oitava geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-86	E	Q	Y	R	H	Y	T	E	L	Q	W	K	Y	E	N	-45.7
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-61	I	Q	E	R	H	Y	T	E	Y	T	Q	S	T	Y	T	-43.2
SGB1-87	I	M	E	N	E	A	L	L	N	L	R	S	I	Y	W	-42.3
SGB1-28	E	I	Y	T	K	Y	H	E	L	I	W	K	S	E	L	-41.7
SGB1-37	Q	Q	Q	R	H	E	T	M	R	Q	M	N	Y	H	N	-41.3
SGB1-62	Y	L	W	M	H	E	T	M	Y	K	Q	S	I	Y	W	-41.2

Tabela 59 – Mutantes filhos da oitava geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-97	I	L	E	M	E	Y	L	E	Y	T	Q	S	T	Y	T	-44.3
SGB1-98	I	M	E	N	M	L	Y	I	L	Q	W	K	Y	E	N	-37.3
SGB1-99	S	N	S	H	L	W	N	S	Y	Q	Q	Q	L	V	I	-35.8
SGB1-100	Y	Q	W	R	H	A	T	L	A	T	E	M	T	I	T	-26.4
SGB1-101	Y	Q	W	R	L	W	N	S	P	N	P	Q	L	V	I	-24.5
SGB1-102	I	M	E	N	E	A	L	L	L	I	W	K	S	E	L	-23.4
SGB1-103	Y	Q	W	R	H	E	T	M	R	Q	M	S	I	Y	W	-22.6
SGB1-104	Q	Q	Q	R	E	A	L	L	Y	Q	Q	S	Y	Y	N	-21.4
SGB1-105	E	Q	Y	R	H	Y	T	E	N	L	R	I	A	R	M	-17.5
SGB1-106	Y	Q	W	R	H	E	T	M	R	Q	Q	S	Y	Y	N	-12.9
SGB1-107	Q	L	Q	M	H	E	T	M	Y	K	M	N	I	H	W	-12.2
SGB1-108	E	I	Y	T	K	Y	H	E	N	L	R	S	I	Y	W	-10.9

Tabela 60 – Nona geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-86	E	Q	Y	R	H	Y	T	E	L	Q	W	K	Y	E	N	-45.7
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-97	I	L	E	M	E	Y	L	E	Y	T	Q	S	T	Y	T	-44.3
SGB1-61	I	Q	E	R	H	Y	T	E	Y	T	Q	S	T	Y	T	-43.2
SGB1-87	I	M	E	N	E	A	L	L	N	L	R	S	I	Y	W	-42.3
SGB1-28	E	I	Y	T	K	Y	H	E	L	I	W	K	S	E	L	-41.7
SGB1-37	Q	Q	Q	R	H	E	T	M	R	Q	M	N	Y	H	N	-41.3

Tabela 61 – Mutantes filhos da nona geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-109	S	Q	S	R	L	W	N	S	P	Q	P	Q	L	V	I	-74.1
SGB1-110	E	Q	Y	R	M	L	Y	I	L	Q	W	I	A	R	M	-50.5
SGB1-111	I	M	E	N	H	Y	T	E	N	L	R	K	Y	E	N	-40.7
SGB1-112	E	Q	Y	R	H	Y	T	E	R	Q	W	K	Y	E	N	-37.4
SGB1-113	I	M	E	N	E	A	L	L	Y	T	Q	S	T	Y	T	-28.2
SGB1-114	I	L	E	M	E	A	L	L	Y	T	Q	M	T	I	T	-27.5
SGB1-115	Y	N	W	H	L	W	N	S	Y	N	Q	Q	L	V	I	-27.2
SGB1-116	Y	Q	W	R	E	A	L	L	Y	Q	Q	S	I	Y	W	-26.0
SGB1-117	Q	Q	Q	R	H	E	T	M	R	Q	M	S	Y	Y	N	-23.8
SGB1-118	I	Q	E	R	H	Y	T	E	N	L	R	S	I	Y	W	-14.9
SGB1-119	Q	I	Q	T	K	E	H	M	L	I	M	N	S	H	L	-11.2
SGB1-120	Y	L	W	M	E	Y	L	E	A	T	E	S	T	Y	T	-9.1

Tabela 62 – Décima geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-109	S	Q	S	R	L	W	N	S	P	Q	P	Q	L	V	I	-74.1
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
SGB1-110	E	Q	Y	R	M	L	Y	I	L	Q	W	I	A	R	M	-50.5
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-86	E	Q	Y	R	H	Y	T	E	L	Q	W	K	Y	E	N	-45.7
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6
SGB1-97	I	L	E	M	E	Y	L	E	Y	T	Q	S	T	Y	T	-44.3
SGB1-61	I	Q	E	R	H	Y	T	E	Y	T	Q	S	T	Y	T	-43.2
SGB1-87	I	M	E	N	E	A	L	L	N	L	R	S	I	Y	W	-42.3

Tabela 63 – Mutantes filhos da Décima geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-121	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-75.6
SGB1-122	Y	Q	W	R	E	E	L	M	Y	Q	Q	S	I	Y	W	-68.5
SGB1-123	Q	Q	Q	R	H	A	T	L	R	Q	M	S	Y	Y	N	-49.4
SGB1-124	I	L	E	M	E	Y	L	E	A	T	Q	S	T	Y	T	-34.6
SGB1-125	E	Q	Y	R	M	L	Y	I	Y	Q	Q	Q	L	V	I	-32.8
SGB1-126	Y	L	W	M	E	A	L	L	Y	T	E	M	T	I	T	-28.8
SGB1-127	Y	Q	W	R	L	W	N	S	L	Q	W	I	A	R	M	-23.9
SGB1-128	S	Q	S	R	L	W	N	S	P	N	P	Q	L	V	I	-22.7
SGB1-129	I	M	E	N	H	Y	T	E	N	L	R	S	T	Y	T	-22.5
SGB1-130	S	N	S	H	L	W	N	S	P	Q	P	Q	L	V	I	-14.6
SGB1-131	E	M	Y	N	M	Y	Y	E	L	L	W	K	A	E	M	-21.5
SGB1-132	I	Q	E	R	H	L	T	I	N	Q	R	I	Y	R	N	-3.8

Tabela 64 – Décima primeira geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-121	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-75.6
SGB1-109	S	Q	S	R	L	W	N	S	P	Q	P	Q	L	V	I	-74.1
SGB1-122	Y	Q	W	R	E	E	L	M	Y	Q	Q	S	I	Y	W	-68.5
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
SGB1-110	E	Q	Y	R	M	L	Y	I	L	Q	W	I	A	R	M	-50.5
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-123	Q	Q	Q	R	H	A	T	L	R	Q	M	S	Y	Y	N	-49.4
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5
SGB1-26	I	M	E	N	M	L	Y	I	N	L	R	I	A	R	M	-46.0
SGB1-86	E	Q	Y	R	H	Y	T	E	L	Q	W	K	Y	E	N	-45.7
SGB1-27	Y	L	W	M	E	A	L	L	A	T	E	M	T	I	T	-45.6

Tabela 65 – Mutantes filhos da Décima primeira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-133	Y	Q	W	R	M	W	Y	S	Y	Q	Q	Q	A	V	M	-54.1
SGB1-134	Q	M	Q	N	M	A	Y	L	N	L	M	S	A	Y	M	-48.7
SGB1-135	Y	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-48.2
SGB1-136	E	Q	Y	R	E	A	L	L	L	Q	W	M	T	I	T	-34.9
SGB1-137	Y	Q	Q	R	H	A	T	L	Y	Q	Q	S	Y	Y	N	-33.5
SGB1-138	I	Q	W	R	E	E	L	M	Y	Q	Q	S	I	Y	W	-29.6
SGB1-139	S	Q	S	R	L	W	N	S	A	Q	A	Q	L	V	I	-26.0
SGB1-140	Y	L	W	M	H	Y	T	E	A	T	E	K	Y	E	N	-25.2
SGB1-141	E	Q	Y	R	L	L	N	I	L	Q	W	I	L	R	I	-22.5
SGB1-142	Q	Q	W	R	H	E	T	M	R	Q	M	S	Y	Y	N	-22.2
SGB1-143	S	N	S	H	L	W	N	S	A	N	A	Q	L	V	I	-16.6
SGB1-144	I	Q	E	R	E	L	L	I	R	Q	R	I	I	R	W	-14.7

Tabela 66 – Décima segunda geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-121	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-75.6
SGB1-109	S	Q	S	R	L	W	N	S	P	Q	P	Q	L	V	I	-74.1
SGB1-122	Y	Q	W	R	E	E	L	M	Y	Q	Q	S	I	Y	W	-68.5
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
SGB1-133	Y	Q	W	R	M	W	Y	S	Y	Q	Q	Q	A	V	M	-54.1
SGB1-110	E	Q	Y	R	M	L	Y	I	L	Q	W	I	A	R	M	-50.5
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-123	Q	Q	Q	R	H	A	T	L	R	Q	M	S	Y	Y	N	-49.4
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-134	Q	M	Q	N	M	A	Y	L	N	L	M	S	A	Y	M	-48.7
SGB1-135	Y	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-48.2
SGB1-49	Q	Q	Q	R	E	A	L	L	R	Q	M	S	I	Y	W	-46.5

Tabela 67 – Mutantes filhos da Décima segunda geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-145	Y	Q	E	R	E	A	L	L	Y	Q	Q	S	I	Y	W	-62.7
SGB1-146	Y	Q	W	R	E	E	L	L	Y	T	Q	S	I	Y	W	-44.7
SGB1-147	I	Q	E	R	E	A	L	M	Y	Q	Q	S	I	Y	W	-36.0
SGB1-148	Q	M	Q	N	H	E	T	M	N	L	M	S	Y	Y	N	-34.8
SGB1-149	Y	Q	W	R	M	W	Y	S	Y	Q	Q	Q	L	V	M	-29.1
SGB1-150	Q	Q	Q	R	E	A	L	L	R	T	M	S	I	Y	W	-27.9
SGB1-151	Q	Q	Q	R	M	A	Y	L	R	Q	M	S	A	Y	M	-26.0
SGB1-152	I	M	E	N	M	L	Y	I	Y	Q	Q	S	I	Y	W	-25.2
SGB1-153	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	A	V	I	-22.4
SGB1-154	Y	Q	W	R	M	A	Y	L	Y	Q	Q	S	A	Y	M	-21.4
SGB1-155	I	Q	E	R	E	E	L	M	N	L	R	I	A	R	M	-18.0
SGB1-156	E	Q	Y	R	H	L	T	I	L	Q	W	I	Y	R	N	-14.9

Tabela 68 – Décima terceira geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-121	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-75.6
SGB1-109	S	Q	S	R	L	W	N	S	P	Q	P	Q	L	V	I	-74.1
SGB1-122	Y	Q	W	R	E	E	L	M	Y	Q	Q	S	I	Y	W	-68.5
SGB1-145	Y	Q	E	R	E	A	L	L	Y	Q	Q	S	I	Y	W	-62.7
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
SGB1-133	Y	Q	W	R	M	W	Y	S	Y	Q	Q	Q	A	V	M	-54.1
SGB1-110	E	Q	Y	R	M	L	Y	I	L	Q	W	I	A	R	M	-50.5
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-123	Q	Q	Q	R	H	A	T	L	R	Q	M	S	Y	Y	N	-49.4
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-134	Q	M	Q	N	M	A	Y	L	N	L	M	S	A	Y	M	-48.7
SGB1-135	Y	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-48.2

Tabela 69 – Mutantes filhos da Décima terceira geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-157	Y	Q	E	R	E	A	L	L	Y	Q	Q	Q	L	V	I	-64.7
SGB1-158	Y	Q	W	R	M	W	Y	S	L	Q	W	I	A	R	M	-45.6
SGB1-159	E	Q	Y	R	M	L	Y	I	Y	Q	Q	Q	A	V	M	-44.2
SGB1-160	Y	Q	E	R	E	A	L	L	N	L	M	S	A	Y	M	-41.0
SGB1-161	Y	Q	W	R	L	W	N	S	Y	Q	Q	S	I	Y	W	-36.8
SGB1-162	Y	Q	W	R	E	E	L	M	Y	T	Q	S	L	Y	W	-36.6
SGB1-163	I	Q	E	R	E	A	L	L	Y	Q	Q	S	I	Y	W	-35.4
SGB1-164	Y	Q	W	R	H	E	T	L	Y	Q	Q	S	Y	Y	N	-31.1
SGB1-165	Q	M	Q	N	M	A	Y	L	Y	T	Q	S	I	Y	W	-23.5
SGB1-166	I	Q	E	R	E	L	L	I	R	L	R	I	I	R	W	-20.4
SGB1-167	Q	M	Q	N	M	A	Y	L	N	Q	M	S	A	Y	M	-14.7
SGB1-168	Q	Q	Q	R	H	A	T	M	R	Q	M	S	Y	Y	N	-14.1

Tabela 70 – Décima quarta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-121	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-75.6
SGB1-109	S	Q	S	R	L	W	N	S	P	Q	P	Q	L	V	I	-74.1
SGB1-122	Y	Q	W	R	E	E	L	M	Y	Q	Q	S	I	Y	W	-68.5
SGB1-157	Y	Q	E	R	E	A	L	L	Y	Q	Q	Q	L	V	I	-64.7
SGB1-145	Y	Q	E	R	E	A	L	L	Y	Q	Q	S	I	Y	W	-62.7
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
SGB1-133	Y	Q	W	R	M	W	Y	S	Y	Q	Q	Q	A	V	M	-54.1
SGB1-110	E	Q	Y	R	M	L	Y	I	L	Q	W	I	A	R	M	-50.5
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.23
SGB1-123	Q	Q	Q	R	H	A	T	L	R	Q	M	S	Y	Y	N	-49.4
SGB1-73	Y	Q	W	R	H	E	T	M	Y	Q	Q	S	Y	Y	N	-49.2
SGB1-134	Q	M	Q	N	M	A	Y	L	N	L	M	S	A	Y	M	-48.7

Tabela 71 – Mutantes filhos da Décima quarta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-169	Y	Q	W	R	M	W	N	S	Y	Q	Q	Q	A	V	M	-55.0
SGB1-170	Y	Q	E	R	E	A	L	L	R	Q	Q	S	I	Y	W	-51.2
SGB1-171	Y	Q	W	R	E	E	L	M	Y	Q	Q	S	L	Y	W	-44.5
SGB1-172	Y	Q	W	R	L	W	Y	S	Y	Q	Q	Q	L	V	I	-38.7
SGB1-173	Y	Q	E	R	E	A	L	L	Y	Q	Q	Q	I	V	W	-36.7
SGB1-174	Q	Q	Q	R	M	A	Y	L	L	Q	M	S	A	Y	M	-36.2
SGB1-175	Q	Q	Q	R	H	A	T	L	N	Q	M	S	Y	Y	N	-36.2
SGB1-176	Y	Q	E	R	E	A	L	L	Y	Q	Q	S	L	Y	I	-34.4
SGB1-177	Q	Q	Q	R	E	A	L	L	Y	T	M	S	I	Y	W	-33.6
SGB1-178	I	Q	E	R	E	A	L	L	Y	T	Q	S	L	Y	W	-31.9
SGB1-179	Y	M	W	N	M	E	Y	M	Y	L	Q	S	A	Y	M	-24.7
SGB1-180	E	Q	Y	R	H	L	T	I	R	Q	W	I	Y	R	N	-8.7

Tabela 72 – Décima quinta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-121	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-75.6
SGB1-109	S	Q	S	R	L	W	N	S	P	Q	P	Q	L	V	I	-74.1
SGB1-122	Y	Q	W	R	E	E	L	M	Y	Q	Q	S	I	Y	W	-68.5
SGB1-157	Y	Q	E	R	E	A	L	L	Y	Q	Q	Q	L	V	I	-64.7
SGB1-145	Y	Q	E	R	E	A	L	L	Y	Q	Q	S	I	Y	W	-62.7
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
SGB1-169	Y	Q	W	R	M	W	N	S	Y	Q	Q	Q	A	V	M	-55.0
SGB1-133	Y	Q	W	R	M	W	Y	S	Y	Q	Q	Q	A	V	M	-54.1
SGB1-170	Y	Q	E	R	E	A	L	L	R	Q	Q	S	I	Y	W	-51.2
SGB1-110	E	Q	Y	R	M	L	Y	I	L	Q	W	I	A	R	M	-50.5
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
SGB1-123	Q	Q	Q	R	H	A	T	L	R	Q	M	S	Y	Y	N	-49.4

Tabela 73 – Mutantes filhos da Décima quinta geração de AM. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-181	Y	Q	E	R	M	W	N	S	Y	Q	Q	Q	A	V	M	-52.9
SGB1-182	Y	Q	E	R	E	A	L	L	Y	Q	Q	Q	L	Y	I	-51.1
SGB1-183	Y	Q	E	R	E	A	L	L	Y	Q	Q	Q	A	V	M	-49.9
SGB1-184	I	Q	D	R	D	A	L	L	Y	Q	Q	S	I	Y	W	-49.5
SGB1-185	Q	Q	Q	R	M	A	Y	L	L	Q	W	S	A	Y	M	-32.5
SGB1-186	Y	Q	E	R	E	A	L	L	Y	Q	Q	S	I	V	W	-30.3
SGB1-187	Y	Q	W	R	M	W	Y	S	R	Q	Q	S	I	Y	W	-29.1
SGB1-188	Q	Q	Q	R	H	A	T	L	Y	Q	M	S	Y	Y	N	-27.8
SGB1-189	I	Q	D	R	D	A	L	L	Y	T	Q	S	I	Y	W	-21.4
SGB1-190	Y	M	W	N	M	E	Y	M	N	L	Q	S	A	Y	M	-21.0
SGB1-191	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-20.8
SGB1-192	E	Q	Y	R	H	L	T	I	R	Q	M	I	Y	R	N	-19.8

Tabela 74 – Décima sexta geração de AM anti-RBD SGB1. Todos os valores de ΔG estão em quilocaloria por mol.

ID\resíduo	2	4	6	8	13	15	17	19	42	44	46	49	51	53	55	ΔG
SGB1-121 _{MD}	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-83.4
SGB1-121 _{histag}	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-81.5
SGB1-121	I	Q	E	R	E	A	L	L	Y	T	Q	S	I	Y	W	-75.6
SGB1-109	S	Q	S	R	L	W	N	S	P	Q	P	Q	L	V	I	-74.1
SGB1-122	Y	Q	W	R	E	E	L	M	Y	Q	Q	S	I	Y	W	-68.5
SGB1-157	Y	Q	E	R	E	A	L	L	Y	Q	Q	Q	L	V	I	-64.7
SGB1-145	Y	Q	E	R	E	A	L	L	Y	Q	Q	S	I	Y	W	-62.7
SGB1-25	S	N	S	H	L	W	N	S	P	N	P	Q	L	V	I	-62.2
SGB1-85	Y	Q	W	R	L	W	N	S	Y	Q	Q	Q	L	V	I	-56.7
SGB1-169	Y	Q	W	R	M	W	N	S	Y	Q	Q	Q	A	V	M	-55.0
SGB1-133	Y	Q	W	R	M	W	Y	S	Y	Q	Q	Q	A	V	M	-54.0
SGB1-181	Y	Q	E	R	M	W	N	S	Y	Q	Q	Q	A	V	M	-53.0
SGB1-170	Y	Q	E	R	E	A	L	L	R	Q	Q	S	I	Y	W	-51.2
SGB1-182	Y	Q	E	R	E	A	L	L	Y	Q	Q	Q	L	Y	I	-51.1
SGB1-110	E	Q	Y	R	M	L	Y	I	L	Q	W	I	A	R	M	-50.5
FabRBD	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	-50.2
GB1-REF	T	K	I	N	K	E	T	E	E	T	D	T	T	T	T	-19.5

APÊNDICE G – Arquivos de trajetórias SGB1-121

Os arquivos originais de trajetória do SGB1-121_{MD}, com o pós-processamento das trajetórias de DM dos mutantes SGB1-121_{MD} e SGB1-121_{his – tag}, nos quais foram removidas moléculas de água, íons, rotação e translação, podem ser baixados no seguinte link:

https://drive.google.com/drive/folders/1g8T-TXYrbSTAmngsnp7Ug-UVfEKx45Do?usp=drive_link

APÊNDICE H – Códigos usados pelo AG

Script da função de pontuação

```
#!/bin/bash

set -e

declare -a arr=("GLY" "ALA" "LEU" "ILE" "VAL" "SER" "THR" "TYR" "ASN" "GLN" "ASP" "GLU" "ARG" "LYS" "HIS" "TRP" "MET" "PRO" )

declare -a aminoacidos=( 2 4 6 8 13 15 17 19 42 44 46 49 51 53 55 )

./zera.exe

for aa in "${aminoacidos[@]}"
do
    for str in "${arr[@]}"
    do
        for ((j=1;j<=20;j=j+1)); do
            ./pymol.sh $str $aa $j

            ./pontua.exe GB1.pdb molde.pdb $aa $j $str

        done
    done
done

rm -f \#*\#

done

done
```

Figura 34 – Script shell usado na função de pontuação.

Fonte: elaborado pelo Autor (2024).

Script da população inicial de mutantes do AG

```
#!/bin/bash
export PATH=$PATH:/usr/apps/pymol/
populacao=18
for ((i=1;i<=$populacao;i=i+1)); do
cd pasta$i
input="amino.dat"
while IFS= read -r line
do
n="$(cut -d " " -f 1 <<"$line")"
residuo="$(cut -d " " -f 2 <<"$line")"
rotamero="$(cut -d " " -f 5 <<"$line")"
pos="$(cut -d " " -f 3 <<"$line")"
if [ $n -eq $i ]
then
mv antibody.pdb select1.pdb
./pymol.sh $residuo $pos $rotamero
echo "$n $residuo $pos $rotamero $sum"
rm select1.pdb
mv select.pdb antibody.pdb
fi
if [ $n -eq $i ]
then
mv complex.pdb select1.pdb
./pymol.sh $residuo $pos $rotamero
echo "$n $residuo $pos $rotamero $sum"
rm select1.pdb
mv select.pdb complex.pdb
fi
done < "$input"
cd ..
done
```

Figura 35 – Script shell usado para gerar a população inicial de mutantes do AG.

Fonte: elaborado pelo Autor (2024).

Script da nova população de mutantes do AG

```
#!/bin/bash
export PATH=$PATH:/usr/apps/pymol/
populacao=12
for ((i=1;i<=$populacao;i=i+1)); do
  cd pasta$i
  input="permutacao.dat"
  tamline=31
  while IFS= read -r line
  do
    n="$(cut -d " " -f 1 <<<"$line")"
    for ((k=3,l=4;k<=$tamline;k=k+2,l=l+2)); do
      residuo="$(cut -d " " -f $k <<<"$line")"
      pos="$(cut -d " " -f $l <<<"$line")"
      echo ">>>> $n $residuo $pos "
      if [ $n -eq $i ]
      then
        mv antibody.pdb select1.pdb
        ./pymol.sh $residuo $pos
        echo "$n $residuo $pos "
        rm select1.pdb
        mv select.pdb antibody.pdb
      fi
      if [ $n -eq $i ]
      then
        mv complex.pdb select1.pdb
        ./pymol.sh $residuo $pos
        echo "$n $residuo $pos"
        rm select1.pdb
        mv select.pdb complex.pdb
      fi
    done
  done < "$input"
  cd ..
done
```

Figura 36 – Script shell usado para gerar a população de mutantes do AG.

Fonte: elaborado pelo Autor (2024).

Script shell usado na DM do algoritmo genético

```
#!/bin/bash
source /usr/apps/amber18/amber.sh
set -e
geracao=1
gerMAX=2
populacao=10
while [ $geracao -lt $gerMAX ]; do
  for ((i=1;i<=$populacao;i=i+1)); do
    cd pasta$i
    /usr/apps/amber18/bin/tleap -f script1.in
    /usr/apps/amber18/bin/tleap -f script2.in
    /usr/apps/amber18/bin/tleap -f script3.in
    echo -e "minimizando $i"
    /usr/apps/amber18/bin/pmemd.cuda -O -i min.in -o min.out -p cpxa.prmtop -c cpxa.inpcrd -r min.ncrst
    echo -e "equilibracao $i"
    /usr/apps/amber18/bin/pmemd.cuda -O -i heat.in -o heat.out -p cpxa.prmtop -c min.ncrst -r heat.ncrst -x heat.nc
    echo -e "MD $i"
    /usr/apps/amber18/bin/pmemd.cuda -O -i prod.in -o prod.out -p cpxa.prmtop -c heat.ncrst -r prod.ncrst -x prod.nc
    /usr/apps/amber18/bin/pmemd.cuda -O -i prod2.in -o prod2.out -p cpxa.prmtop -c prod.ncrst -r prod2.ncrst -x prod.nc
    echo -e "MMPBSA $i"
    amber.python /usr/apps/amber18/bin/MMPBSA.py -O -i mmpbsa.in -o FINAL_RESULTS_MMPBSA.dat -sp cpxa.prmtop -cp cpxv.prmtop -rp antibody.prmtop -lp antigen.prmtop -y prod.nc
    rm -rf prod.nc
    cd ..
  done
  let geracao=geracao+1;
done
```

Figura 37 – Script shell usado na DM do algoritmo genético.

Fonte: elaborado pelo Autor (2024).

Códigos auxiliares usados na no AG

Devido à natureza modular do AG, diversas partes dos códigos foram escritas em linguagem C++, o que facilita a manipulação de arquivos de texto ou em formato PDB. Os códigos a seguir desempenham funções auxiliares dentro do AG, como a função de pontuação e a função de permutação:

/*Código de permutação para gerar mutantes filhos*/

```
#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <math.h>

struct vetor {
    char line[1000];
};

struct vetor L[100000];
struct vetor L2[100000];

float separar(int i, char * string) {
    char str[10];

    for (int j = i - 3, k = 0; j < i + 3; j++, k++)
        str[k] = string[j];
    str[7] = '\0';

    return atof(str);
}

int igual(char * str1, char * str2) {
    int ok = 1;
    for (int i = 0; i < 100; i++)
        if (str1[i] != str2[i])
            ok = 0;
    return ok;
}

int verifica(struct vetor * Li, int N, char * str) {
    int ok = 0;
    for (int i = 0; i < N; i++) {
        if (igual(str, Li[i].line) == 1)
            ok = 1;
    }
    return ok;
}

int main() {
    FILE * fp;
    char c;
    int i = 0;
    int N2 = 0;
```

```
fp = fopen("ranking_ordem.dat", "r"); /* Arquivo ASCII, para teitura */
if (!fp) {
    printf("Erro na abertura do arquivo");
}
char line[150] = ""; /* initialize all to 0 ('\0') */

while (fgets(L[i].line, sizeof(line), fp)) {
    /* tsin is FILE* input */

    L[i].line[101] = '\0';

    i++;
} /* imprime o caracter lido */

fclose(fp);

fp = fopen("permutacao.dat", "w");
int corte = 51;
char m;
char str1[200], str2[200];

for (int j = 0; j < 12; j = j + 2) {

    strcpy(str1, L[j].line);
    strcpy(str2, L[j + 1].line);

    for (int l = 0; l < corte; l++) {
        str2[l] = L[j].line[l];
        str1[l] = L[j + 1].line[l];
    }

    if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
        printf("copia erro: segunda permutação\n");
        for (int l = 0; l < 24; l++) {
            char c = str2[l];
            str2[l] = str1[l];
            str1[l] = c;
        }

        for (int l = 72; l < 100; l++) {
            char c = str2[l];
            str2[l] = str1[l];
            str1[l] = c;
        }
    }
}
```

```
if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: terceira permutação\n");
    printf("%s\n", str1);
    printf("%s\n", str2);
    printf("%d\n", j);

    for (int l = 5; l < 9; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
    for (int l = 17; l < 21; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
    for (int l = 30; l < 35; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
    for (int l = 44; l < 48; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
    for (int l = 58; l < 62; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
    for (int l = 72; l < 76; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
    for (int l = 86; l < 90; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
```

```
printf("copia erro: quarta permutação\n");
for (int l = 50; l < 58; l++) {
    char c = str2[l];
    str2[l] = str1[l];
    str1[l] = c;
    printf("%s \n %s \n", str1, str2);
}
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: quinta permutação\n");
    for (int l = 59; l < 63; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 6 permutação\n");
    for (int l = 64; l < 71; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 7 permutação\n");
    for (int l = 71; l < 78; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 8 permutação\n");
    for (int l = 78; l < 85; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 9 permutação\n");
```

```
    for (int l = 85; l < 92; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}
if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 10 permutação\n");
    for (int l = 92; l < 99; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 11 permutação\n");
    for (int l = 0; l < 4; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 12 permutação\n");
    for (int l = 4; l < 10; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 13 permutação\n");
    for (int l = 10; l < 16; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;
    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 14 permutação\n");
    for (int l = 16; l < 22; l++) {
        char c = str2[l];
        str2[l] = str1[l];
```

```
        str1[l] = c;

    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 15 permutação\n");
    for (int l = 22; l < 29; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;

    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 16 permutação\n");
    for (int l = 29; l < 36; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;

    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 17 permutação\n");
    for (int l = 36; l < 43; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;

    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: 18 permutação\n");
    for (int l = 43; l < 50; l++) {
        char c = str2[l];
        str2[l] = str1[l];
        str1[l] = c;

    }
}

if (verifica(L, i, str1) == 1 || verifica(L, i, str2) == 1) {
    printf("copia erro: nenhuma permutação efetiva\n%s\n", str1, str2);
}

strcpy(L2[j].line, str1);
strcpy(L2[j + 1].line, str2);

}
```

```
for (int j = 0; j < 12; j++)  
    fprintf(fp, "%d %s\n", j + 1, L2[j].line);  
  
fclose(fp);  
return 0;  
}
```

Código que calcula a distância e pontua o tipo de interação <Pontua.exe>

```
#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <math.h>

/*Estrutura que armazena informação sobre a linha do arquivo PDB*/

struct vetor {
    char line[100];
};

/*Listas que armazenam os arquivos PDB*/
struct vetor L[10000];

struct vetor L2[100000];

char str[9];

/*função que converte substring em float*/

float separar(int i, char * string) {

    for (int j = i - 3, k = 0; j < i + 3; j++, k++)
        str[k] = string[j];
    str[8] = '\0';

    return atof(str);
}

/*função que converte substring em inteiro*/
int separarint(int i, char * string) {

    for (int j = i - 2, k = 0; j < i + 2; j++, k++)
        str[k] = string[j];
    str[8] = '\0';

    return atoi(str);
}

/*função principal*/

int main(int argc, char * argv[])

{
```

```
FILE * fp;

int Na1 = 0, Na2 = 0;

fp = fopen(argv[1], "r"); /* Arquivo ASCII, para leitura */
if (!fp) {
    printf("Erro na abertura do arquivo");
}
char line[150] = ""; /* initialize all to 0 ('\0') */

while (fgets(L[Na1].line, sizeof(line), fp)) {
    /* tsin is FILE* input */

    if (L[Na1].line[0] == 'E' || L[Na1].line[0] == ' ') break;

    Na1++;
}

fclose(fp);

fp = fopen(argv[2], "r"); /* Arquivo ASCII, para leitura */
if (!fp) {
    printf("Erro na abertura do arquivo");
}
char line2[150] = ""; /* initialize all to 0 ('\0') */

while (fgets(L2[Na2].line, sizeof(line2), fp)) {
    /* tsin is FILE* input */

    if (L2[Na2].line[0] == 'E' || L2[Na2].line[0] == ' ') break;

    Na2++;
} /* */

fclose(fp);

/* abre arquivo output*/
fp = fopen("pontuacao.dat", "a");

/*inicializa variaveis*/
int posicao = atoi(argv[3]), pos, k, pontos = 0, Nres = 0;

/*lê posições do átomos*/
for (pos = 0; pos < Na1; pos++) {
    if (separarint(25, L[pos].line) == posicao && L[pos].line[77] != 'H') {
```

```

Nres++;
for (k = 0; k < Na2; k++) {

    float x1, y1, z1, px, py, pz;

    x1 = separar(34, L2[k].line);
    y1 = separar(42, L2[k].line);
    z1 = separar(51, L2[k].line);
    px = separar(34, L[pos].line);
    py = separar(42, L[pos].line);
    pz = separar(51, L[pos].line);

    /*calcula distâncias entre átomos e pontuações para o tipo de interação*/
    if (sqrt(pow(x1 - px, 2) + pow(y1 - py, 2) + pow(z1 - pz, 2)) >= .00 && sqrt(pow(x1 -
px, 2) + pow(y1 - py, 2) + pow(z1 - pz, 2)) < 0.5) {

        if (L2[k].line[77] == 'C' && L[pos].line[77] == 'C') pontos += 17;
        if (L2[k].line[77] == 'C' && L[pos].line[77] == 'N') pontos -= 95;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'C') pontos -= 95;
        if (L2[k].line[77] == 'C' && L[pos].line[77] == 'O') pontos -= 95;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'C') pontos -= 95;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'O') pontos -= 95;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'N') pontos -= 95;
        if (L2[k].line[77] == 'H' && L[pos].line[77] == 'N' && L[pos].line[17] == 'H' &&
L[pos].line[18] == 'I' && L[pos].line[19] == 'S') pontos += 150;
        if (L2[k].line[77] == 'H' && L[pos].line[77] == 'O' && L[pos].line[17] == 'A' &&
L[pos].line[18] == 'S' && L[pos].line[19] == 'N') pontos += 150;
        if (L2[k].line[77] == 'H' && L[pos].line[77] == 'O' && L[pos].line[17] == 'G' &&
L[pos].line[18] == 'L' && L[pos].line[19] == 'N') pontos += 150;
        if (L2[k].line[77] == 'H' && L[pos].line[77] == 'O' && L[pos].line[17] == 'G' &&
L[pos].line[18] == 'L' && L[pos].line[19] == 'U') pontos += 150;
        if (L2[k].line[77] == 'H' && L[pos].line[77] == 'O' && L[pos].line[17] == 'A' &&
L[pos].line[18] == 'S' && L[pos].line[19] == 'P') pontos += 150;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'O' && L[pos].line[17] == 'S' &&
L[pos].line[18] == 'E' && L[pos].line[19] == 'R') pontos += 150;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'O' && L[pos].line[17] == 'T' &&
L[pos].line[18] == 'Y' && L[pos].line[19] == 'R') pontos += 150;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'N' && L[pos].line[17] == 'T' &&
L[pos].line[18] == 'R' && L[pos].line[19] == 'P') pontos += 150;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'N' && L[pos].line[17] == 'A' &&
L[pos].line[18] == 'S' && L[pos].line[19] == 'N') pontos += 150;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'N' && L[pos].line[17] == 'G' &&
L[pos].line[18] == 'L' && L[pos].line[19] == 'N') pontos += 150;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'N' && L[pos].line[17] == 'A' &&
L[pos].line[18] == 'R' && L[pos].line[19] == 'G') pontos += 150;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'N' && L[pos].line[17] == 'H' &&
L[pos].line[18] == 'I' && L[pos].line[19] == 'S') pontos += 150;
        if (L2[k].line[77] == 'O' && L[pos].line[77] == 'N' && L[pos].line[17] == 'L' &&
L[pos].line[18] == 'Y' && L[pos].line[19] == 'S') pontos += 150;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'O' && L[pos].line[17] == 'S' &&
L[pos].line[18] == 'E' && L[pos].line[19] == 'R') pontos += 150;
    }
}

```

```

        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'O' && L[pos].line[17] == 'T' &&
L[pos].line[18] == 'Y' && L[pos].line[19] == 'R') pontos += 150;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'N' && L[pos].line[17] == 'T' &&
L[pos].line[18] == 'R' && L[pos].line[19] == 'P') pontos += 150;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'N' && L[pos].line[17] == 'A' &&
L[pos].line[18] == 'S' && L[pos].line[19] == 'N') pontos += 150;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'N' && L[pos].line[17] == 'G' &&
L[pos].line[18] == 'L' && L[pos].line[19] == 'N') pontos += 150;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'N' && L[pos].line[17] == 'A' &&
L[pos].line[18] == 'R' && L[pos].line[19] == 'G') pontos += 150;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'N' && L[pos].line[17] == 'H' &&
L[pos].line[18] == 'I' && L[pos].line[19] == 'S') pontos += 150;
        if (L2[k].line[77] == 'N' && L[pos].line[77] == 'N' && L[pos].line[17] == 'L' &&
L[pos].line[18] == 'Y' && L[pos].line[19] == 'S') pontos += 150;

    }

}

}

}

printf(">>>>>>>>>%d\n", Nres);
printf("Resid= %s pos= %d pontuacao= %d rotamero %d\n", argv[5], posicao, (pontos) /
Nres, atoi(argv[4]));
fprintf(fp, "Resid= %s pos= %d pontuacao= %d rotamero %d\n", argv[5], posicao,
(pontos) / Nres, atoi(argv[4]));
fclose(fp);

return 0;
}

```