



Universidade de São Paulo
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto
Departamento de Química
Programa de Pós-Graduação em Química

“Estudo teórico relacionado a candidatos de vacinas para dependência de cocaína”.

Newton Célio da Silva Junior

Dissertação apresentada à Faculdade de
Filosofia, Ciências e Letras de Ribeirão Preto da
Universidade de São Paulo, como parte das exigências para
a obtenção do título de Mestre em Ciências, Área: **Química**

Ribeirão Preto – SP

2023

Universidade de São Paulo
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto
Departamento de Química
Programa de Pós-Graduação em Química

Newton Célio da Silva Junior

VERSÃO CORRIGIDA

“Estudo teórico relacionado a candidatos de vacinas para dependência de cocaína”.

Dissertação apresentada à Faculdade de
Filosofia, Ciências e Letras de Ribeirão Preto da
Universidade de São Paulo, como parte das exigências para
a obtenção do título de Mestre em Ciências, Área: **Química**

Orientadora: Prof. Dra. Aline Thaís Bruni

Ribeirão Preto – SP

2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Silva-Junior, Newton Célio da

Estudo teórico relacionado a candidatos de vacinas para dependência de cocaína. Ribeirão Preto, 2023.
102p.

Dissertação de Mestrado apresentada à Faculdade de Filosofia, Ciência e Letras de Ribeirão Preto/USP – Área de concentração: Química.

Orientadora: Bruni, Aline Thais

1. Vacina. 2. Dependência de Cocaína. 3. Métodos *in silico*. 4. HLA de Classe II. 5. Química quântica. 6. Quimiometria.

Nome: SILVA JUNIOR, Newton Célio da

Título: Estudo teórico relacionado a candidatos de vacina para dependência de cocaína

Dissertação apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, como parte das exigências para a obtenção do título de Mestre em Ciências, Área: **Química**

Aprovado em:

BANCA EXAMINADORA

Prof. Dr.: _____

Instituição: _____

Julgamento: _____

Assinatura: _____

Prof. Dr.: _____

Instituição: _____

Julgamento: _____

Assinatura: _____

Prof. Dr.: _____

Instituição: _____

Julgamento: _____

Assinatura: _____

AGRADECIMENTOS

Agradeço a minha família, especialmente a minha mãe Laura, meu pai Newton e minha avó Genoveva por sempre me oferecerem todo apoio e escuta possíveis. Carrego vocês por onde eu for. Espero um dia poder retribuir à altura todo o amor e cuidado.

À minha orientadora Aline Thaís Bruni que se destaca tanto como pesquisadora quanto ser humano. Esses anos definitivamente seriam mais difíceis se fosse qualquer outra pessoa assumindo esse papel de orientação. Minha admiração eterna por quem você é, te desejo o mundo.

Aos colegas do Laboratório de Estudos Interdisciplinares por todas oportunidades de colaboração e aprendizado. E aos meus amigos, Francisco e Bárbara, que foram a família que escolhi e nunca hesitaram em estar por mim mesmo quando eu não estava tão presente, amo vocês.

Ao Departamento de Química da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto por toda a minha formação.

À CAPES pelo financiamento desse mestrado e todas as oportunidades oriundas a partir dele.

RESUMO

SILVA-JUNIOR, Newton Célio da. Estudo teórico relacionado a candidatos de vacina para dependência de cocaína. 2023, 102 f. Dissertação de Mestrado, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

A cocaína é uma substância estimulante que pode causar dependência cujo uso tem crescido em várias regiões do mundo. Nesse contexto, é importante pensar em estratégias de redução de danos, isto é, práticas que levem à redução dos problemas associados ao uso dessa droga. Atualmente, no entanto, o tratamento psicossocial vigente aliado a medicamentos que agem sobre os sítios de ligação da cocaína não representa uma boa metodologia de redução de danos visto a alta taxa de evasão enfrentada. Assim, o desenvolvimento de uma vacina para a dependência de cocaína surge como uma abordagem promissora para o tratamento de dependentes químicos dessa substância e, até mesmo, para a prevenção de overdoses. Neste trabalho investigou-se as vacinas para dependência de cocaína conjugadas testadas *in vivo* – compostas de uma combinação proteína carreadora e hapteno - através de métodos computacionais, colocando em destaque o sistema Antígeno Leucocitário Humano (HLA) de classe II. Esse sistema, caracterizado por um grau elevado de polimorfismo, é o responsável por apresentar os antígenos para as células T CD4+, sendo necessário a ativação destas para que a vacina seja capaz de elicitar uma resposta imune adequada. Considerando a variedade de alelos encontrados na população, uma vacina desejável é aquela que dê origem a epítomos de células T reconhecidos com alta afinidade pelas moléculas de HLA da classe II. Destarte, as vacinas já elaboradas foram avaliadas diante desse parâmetro de cobertura populacional com base em Quimiometria sobre cálculos de descritores físico-químicos e biológicos, além de simulações de *docking* e dinâmica molecular para entender o caráter e a estabilidade das interações estabelecidas nesse sistema químico. A análise de componentes principais (PCA) conseguiu distinguir a proteína subunidade B da toxina da Cólera (rCTB) das demais proteínas carreadoras, pelos peptídeos gerados a partir desta apresentarem índices de instabilidade mais altos. Apesar disso, os dados obtidos não foram suficientes para gerar agrupamentos relacionados aos alelos ou haptenos. A proteína rCTB seria a pior escolha como proteína carreadora pela maior parte dos resíduos de aminoácidos envolvidos na reação de conjugação da vacina não estarem acessíveis ao solvente, além dos peptídeos apresentarem baixa promiscuidade. A albumina do soro bovino (BSA), por sua vez, se destacou por gerar peptídeos que se ligavam a até quatro alelos diferentes, além de ser a proteína carreadora cujos peptídeos, em média, apresentaram menor energia de interação intermolecular nos *dockings* para o mecanismo de apresentação dos peptídeos não conjugados. Levando em consideração a possível apresentação de peptídeos conjugados, notou-se que a presença do hapteno majoritariamente favorece a formação de ligações de van der Waals e ligações hidrogênio, sendo um mecanismo mais energeticamente favorável. A dinâmica molecular demonstrou que os complexos peptídeo-hapteno-HLA de classe II são estáveis ao longo do tempo, mas que existe uma variação da quantidade total e na força das ligações hidrogênio quando se varia a formulação da vacina conjugada e o alelo sob estudo. As evidências apontam que o mecanismo de apresentação de peptídeos conjugados é de fato possível, no entanto, não é possível afirmar que a diferença de interações intermoleculares possa ser um fator determinante na imunogenicidade das vacinas visto que não se sabe a frequência biológica desse mecanismo.

Palavras-chaves: cocaína, vacinas conjugadas, métodos *in silico*, HLA de classe II, quimiometria, *docking*, dinâmica molecular.

ABSTRACT

SILVA-JUNIOR, Newton Célio da. Theoretical study applied to vaccine candidates for Cocaine Use Disorder (CUD). 2023, 102 f. Dissertação de Mestrado, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto.

Cocaine is an addictive stimulant drug, whose use is increasing in several regions of the world. In this scenario, it is important to think about harm reduction principles, that is, strategies aimed at reducing negative consequences associated with drug use. Nowadays, however, psychosocial treatment combined with interventions on cocaine binding sites does not represent a good harm reduction methodology due to high evasion. Thus, the development of a vaccine for cocaine use disorder emerges as a promising approach for the treatment of drug users and even for overdose prevention. In this work, *in vivo* evaluated conjugate vaccines for cocaine dependence were investigated through computational methods, highlighting Human Leukocyte Antigen (HLA) class II roles. This system, characterized by a high degree of polymorphism, is responsible for presenting antigens to CD4⁺ T cells, which in turn elicit an adequate immune response when activated. Considering human allelic variation, a desirable vaccine would generate high affinity T cell epitopes to bind as many alleles as possible. Therefore, the vaccines were evaluated based on Chemometrics performed on physicochemical and biological descriptors, in addition to *docking* and molecular dynamics simulations in order to understand the nature and stability of interactions in this chemical system, always keeping up with population coverage parameter. Principal Component Analysis (PCA) was able to distinguish cholera toxin subunit B protein (rCTB) from the other carrier proteins, because the peptides generated from it presented higher instability index. Despite this, data were not sufficient to generate clusters related to alleles or haptens. rCTB protein would be the worst choice as a carrier protein by virtue of most of amino acid residues involved in conjugation reaction are not accessible to solvent, besides peptides having low promiscuity. On the other hand, Bovine Serum Albumine (BSA) stood out for generating peptides that bound up to four different alleles, as well being the carrier protein whose peptide, on average, exhibited lower intermolecular energy in *dockings* for the presentation mechanism of unconjugated peptides. Taking into consideration the possible presentation of conjugated peptides, it was observed that hapten presence mostly favors formation of van der Waals interaction and hydrogen bonds, as a result, being more energetically favorable. Molecular dynamics demonstrated that peptide-hapten-HLA class II complexes are stable over time, but there is variation among total amount and force of hydrogen bonds when varying formulation of conjugate vaccine and allele under study. Evidences indicate that the mechanism of presentation of conjugated peptides is indeed possible, however, it is not possible to state that difference in intermolecular interactions can be a determining factor in vaccine immunogenicity, since biological frequency of this phenomena is not known.

Keywords: cocaine, conjugate vaccines, *in silico* methods, HLA class II, chemometrics, *docking*, molecular dynamics.

Lista de Equações

Equação 1: Função de pontuação semiempírica do Autodock 4.2	27
Equação 2: Equação de Schrödinger independente do tempo, não-relativística	30
Equação 3: Operador Hamiltoniano	30
Equação 4: Função de Variação Linear	31
Equação 5: Princípio Variacional	31
Equação 6: Hamiltoniano Hartree	31
Equação 7: Funcional de Densidade	33
Equação 8: Funcionais Híbridos	34
Equação 9: Funcional B3LYP	34
Equação 10: Forma funcional de um típico campo de força GROMOS	35
Equação 11: RMSD	37
Equação 12: Energia Livre de Ligação Autodock	52

Lista de Figuras

Figura 1 - Farmacologia da cocaína	17
Figura 2 - Maturação de afinidade no centro geminal.....	22
Figura 3. Procedimento Computacional	38
Figura 4 – Reação de conjugação entre o hapteno (1) e a proteína carreadora (2) para formação de uma carbodiimida	39
Figura 5– Molécula de cocaína em comparação com os seus análogos estruturalmente modificados com linkers (haptenos).....	40
Figura 6 – Distribuição das amostras em duas dimensões definidas pelos eixos das duas componentes principais da PCA. Três variáveis de entrada, três classes (proteínas carreadoras)	46
Figura 7– Distribuição dos peptídeos em três dimensões definidas pelos eixos das três primeiras componentes principais da PCA. Três variáveis de entrada, três classes (proteínas carreadoras)	47
Figura 8– Representação gráfica dos <i>loadings</i> : posição das variáveis em relação às componentes principais	47
Figura 9– Distribuição dos peptídeos em duas dimensões definidas pelos eixos das duas componentes principais da PCA. Quatro variáveis de entrada, dez classes (alelos).....	48
Figura 10 – Sete supertipos definidos pela semelhança entre o conjunto de peptídeos que se ligam a cada alelo	49
Figura 11 – Distribuição dos peptídeos em três dimensões definidas pelos eixos das componentes principais da PCA, cinco variáveis de entrada, seis classes (supertipos).....	49
Figura 12 – Distribuição das combinações peptídeos-haptenos em três dimensões definidas pelos eixos das componentes principais da PCA, seis variáveis de entrada provenientes do <i>docking</i> e oito classes (alelos)	54
Figura 13 – Distribuição das combinações peptídeos-haptenos em três dimensões definidas pelos eixos das componentes principais da PCA. Seis variáveis de entrada provenientes do <i>docking</i> , três classes (haptenos).....	55

Figura 14 – Distribuição das combinações peptídeos-haptenos em três dimensões definidas pelos eixos das componentes principais da PCA. Seis variáveis de entrada provenientes do <i>docking</i> , três classes (proteínas carreadoras).....	56
Figura 15. Número de ligações hidrogênio a cada frame da dinâmica molecular entre o alelo DRA*0101/2-DRB1*0301 e o seu respectivo peptídeo-hapteno de maior afinidade conforme determinando anteriormente por <i>docking</i>	57
Figura 16. Número de ligações hidrogênio a cada frame da dinâmica molecular entre o alelo DRA*0101/2-DRB1*0101 e o seu respectivo peptídeo-hapteno de maior afinidade conforme determinando anteriormente por <i>docking</i>	58
Figura 17. Número de ligações hidrogênio a cada frame da dinâmica molecular entre o alelo DQA1*0102-DQB1*0602 e o seu respectivo peptídeo-hapteno de maior afinidade conforme determinando anteriormente por <i>docking</i>	58
Figura 18 - RMSD da cadeia principal da combinação peptídeo-hapteno de maior afinidade para o alelo DRA*0101/2-DRB1*0301 a cada frame da dinâmica molecular.....	59
Figura 19. RMSD da cadeia principal da combinação peptídeo-hapteno de maior afinidade para o alelo DRA*0101/2-DRB1*0101 a cada frame da dinâmica molecular	60
Figura 20. RMSD da cadeia principal da combinação peptídeo-hapteno de maior afinidade para o alelo DQA1*0102-DQB1*0602 a cada frame da dinâmica molecular	60
Figura A1. Sequência primária dos cinco monômeros (DEFGH) que formam a subunidade B pentamérica da toxina da Cólera	93
Figura A2 - Esquema de cores e rótulos referentes a cada um dos alelos.....	94
Figura A3 - Sobreposição das respostas observadas dos alelos em relação aos peptídeos originados pela proteína BSA. Os quadros coloridos em frente a cada sequência mostram a quais alelos cada uma delas se liga com alta afinidade (IC_{50} teórico < 1000 nM).....	95
Figura A4 - Sobreposição das respostas observadas dos alelos em relação aos peptídeos originados pela proteína KLH. Os quadros coloridos em frente a cada sequência mostram a quais alelos cada uma delas se liga com alta afinidade (IC_{50} teórico < 1000 nM).....	96

Lista de tabelas

Tabela 1- Combinações hapteno-proteína carreadora para as quatro vacinas conjugadas a serem analisadas.....	39
Tabela 2. Nomenclatura dos alelos e entradas correspondentes no <i>PDB</i>	43
Tabela 3 – Valores de informação acumulada, erro padrão de calibração (SEV), erro padrão de validação (SEC) e q^2/r^2 (validação e calibração) para o modelo PLS de 3 fatores para o IC_{50} calculado pelo SMM-align e NN-align.	50
Tabela A1. Peptídeos de alta afinidade com moléculas do sistema HLA de classe II formados a partir da proteína BSA, considerando as lisinas quimicamente acessíveis, processamento por catepsinas e descritores físico-químicos e/ou biológicos calculados para cada um deles. Separados em blocos de acordo com o alelo ao qual se liga.	74
Tabela A2. Peptídeos de alta afinidade com moléculas do sistema HLA de classe II formados a partir da proteína subunidade B da toxina da Cólera, considerando as lisinas quimicamente acessíveis, processamento por catepsinas e descritores físico-químicos e/ou biológicos calculados para cada um deles. Separados em blocos de acordo com o alelo ao qual se liga..	80
Tabela A3. Peptídeos de alta afinidade com moléculas do sistema HLA de classe II formados a partir da proteína KLH, considerando as lisinas quimicamente acessíveis, processamento por catepsinas e descritores físico-químicos e/ou biológicos calculados para cada um deles. Separados em blocos de acordo com o alelo ao qual se liga.	81
Tabela A4 – Energias (kcal/mol) obtidas para as quatro combinações peptídeo-hapteno de maior afinidade (IC_{50} teórico < 1000 nM) para osn alelos com estrutura disponível no PDB.....	98
Tabela A5 – Energias (kcal/mol) obtidas para os peptídeos de maior afinidade (IC_{50} teórico < 1000 nM) para os alelos com estrutura disponível no PDB.	101
Tabela A6. Análise relativa à formação de ligações hidrogênio durante a dinâmica molecular dos alelos com sua combinação peptídeo-hapteno de maior afinidade, incluindo o número de ligações hidrogênio e os resíduos de aminoácidos da molécula de MHC da class II envolvidos nas ligações com maior ocupação.	103

Lista de abreviações e siglas

APC - Antigen Presenting Cells

B3LYP - Becke 3-parameter Lee–Yang–Parr

BSA – Bovine Serum Albumin

CLIP - Class II-associated Invariant Chain Peptide

DFT - Density-Functional Theory

DNA - Deoxyribonucleic Acid

FDA - Food and Drug Administration

GAFF - General Amber Force Field

GRAVY - Grande Average of Hydropathicity

HLA - Human Leukocyte Antigen

IEDB - Immune Epitope Database and Analysis Resource

KLH – Keyhole Limpet Hemocyanin

LCAO – Linear Combination of Atomic Orbitals

LDA - Local-density Approximation

LSDA - Local Spin Density Approximation

MHC - Major Histocompatibility Complex

PCA - Principal Component Analysis

PCR - Polymerase Chain Reaction

PDB - Protein Data Bank

PDBePISA - Protein Data Bank in Europe: Proteins, Interfaces, Structures and Assemblies

pI - Ponto Isoelétrico

PLS - Partial Least Squares

PME – Particle-Mesh Ewald

PROSPER - Protease substrate specificity server

rCTB – Recombinant Cholera Toxin B Subunit

RMSD - Root-mean-square deviation

UNODC - United Nation Office on Drugs and Crime

VMD - Visual Molecular Dynamics

Sumário

1. Capítulo 1: INTRODUÇÃO GERAL E CARACTERIZAÇÃO DO PROBLEMA	13
1.1 Breve histórico do uso da cocaína e cenário atual	13
1.2 Extração e produção da cocaína	15
1.3 Formas de administração, biodisponibilidade e farmacologia da cocaína.....	16
1.4 Vacina para dependência de cocaína: características e mecanismo de ação	20
1.5 Objetivos.....	24
2. Capítulo 2: TÉCNICAS <i>IN SILICO</i> PARA INVESTIGAÇÃO DO PROBLEMA	25
2.1 <i>Docking</i>	25
2.2 Quimiometria	28
2.3 DFT	30
2.4 Dinâmica Molecular	35
Capítulo 3: PROCEDIMENTOS COMPUTACIONAIS	38
4. Capítulo 4: RESULTADOS E DISCUSSÕES	46
5. CONCLUSÕES	62
6. REFERÊNCIAS BIBLIOGRÁFICAS	64
7. APÊNDICE.....	74

1. Capítulo 1: INTRODUÇÃO GERAL E CARACTERIZAÇÃO DO PROBLEMA

1.1 Breve histórico do uso da cocaína e cenário atual

Historicamente, o uso de cocaína foi observado pela primeira vez pelas tribos nômades que dariam origem ao império Inca para lidar com o mal-estar ocasionado pelas altitudes elevadas [1]. Ainda hoje, o hábito de mascar folhas de coca é culturalmente existente entre a população andina, fazendo-se presente em celebrações tradicionais e religiosas [2].

A cocaína foi isolada das folhas de coca apenas em 1859 por Albert Niemann, vindo a ganhar destaque em 1884, quando o oftalmologista Karl Koller descobriu seu efeito anestésico tópico [3]. Ao passo que a utilização de cocaína como anestésico se popularizava em diversas áreas da medicina, também crescia o medo dessa substância pela dependência desenvolvida, principalmente, em quem a usava para tratamento de vício em morfina. Esses casos levaram, em sua maioria, ao encerramento do uso da cocaína com esse objetivo pela comunidade médica [4].

Na mesma época de popularização da cocaína no campo medicinal, era comum encontrar bebidas que empregassem quantidades que não produziam toxicidade aos humanos. Com a criação de uma técnica que convertia as folhas de coca em cocaína pré-refinada, facilitaram-se o transporte e o armazenamento pelo fato de o produto não ser perecível. Dessa forma, o preço dos amplos estoques de cocaína se tornou baixo, o que possibilitou que os produtores de bebidas adicionassem quantidades maiores nas suas mercadorias, resultando em casos de toxicidade e mortes [5].

A legalidade do comércio de cocaína começou a findar a partir da Convenção Internacional de Ópio de 1912, na qual restrições que viriam a ser crescentes já começaram a ser aplicadas, levando no fim das contas à proibição da droga [6]. Junto a isso, o conhecimento difundido entre a população acerca dos efeitos da cocaína e o surgimento de estimulantes mais duradouros e baratos contribuíram para a redução no consumo dessa substância [7]. Contudo, uma tendência contrária começou a ser observada a partir do surto de usuários de “crack” em 1983 nas Bahamas, uma forma de cocaína base livre capaz de estabelecer o vício com maior rapidez [8].

Segundo o último relatório do Escritório das Nações Unidas sobre Drogas e Crime (UNODC-*United Nations Office for Drugs and Crimes*) publicado em 2021, mas com dados

sobre uso de cocaína não atualizados desde 2019, cerca de 0,4 % da população global na faixa de 15-64 anos foi usuária de cocaína durante os 12 meses anteriores à publicação, o que equivale a aproximadamente 20 milhões de pessoas.

Os tratamentos psicossociais vigentes, embora necessários, não são eficazes por si só para o tratamento da dependência para usuários que assim desejarem, fato comprovado pela alta taxa de evasão [9]. Não existem medicamentos aprovados pelo órgão estadunidense FDA (*Food and Drug Administration*) para tratamento da dependência de cocaína ainda que existam diversos estudos clínicos que busquem tratamentos capazes de estabelecer abstinência e/ou evitar relapso [10]. Diferentemente da maioria das intervenções terapêuticas propostas, a vacina para dependência de cocaína representa uma intervenção farmacodinâmica afetando diretamente a distribuição da droga no organismo humano, ao fim impedindo que ela atinja seu sítio de ação [11].

A primeira tentativa de desenvolvimento de uma vacina para transtorno por uso de substâncias ocorreu no início dos anos 70, sendo voltada para casos de abuso de morfina e heroína [12]. O potencial de criação de uma vacina para a dependência de cocaína só foi explorado a partir dos anos 90, iniciando-se com a imunização por cocaína conjugada à hemocianina de lapa buraco de fechadura [13].

Apesar das diversas formulações de vacinas já testadas, os resultados sempre decaem em eficácia quando se passa para a etapa de estudos clínicos, isto é, realizados em humanos [14]. Nesse contexto, a TA-CD, única vacina que avançou para a fase II nos estudos clínicos, foi capaz de elicitar altos níveis de anticorpos apenas em 38% dos participantes [15]. Essa variação na resposta pode estar relacionada ao mecanismo de funcionamento de uma vacina conjugada, o qual envolve apresentação pelo sistema Antígeno Leucocitário Humano de classe II (HLA de classe II) [16]. Esse sistema é conhecido pelo seu alto grau de polimorfismo que, no sítio de ligação, define a afinidade e especificidade de cada molécula [17].

1.2 Extração e produção da cocaína

A cocaína é uma substância de ocorrência natural em plantas nativas da América do Sul. Trata-se de um alcaloide, isto é, um composto orgânico nitrogenado com propriedades básicas [2]. Quimicamente, pode ser encontrada em duas formas principais: a forma de sal e a de base livre. Não existe um consenso quanto ao processo de fabricação voltado para o tráfico uma vez que vários aditivos podem ser inclusos com o objetivo de aumentar o rendimento e, conseqüentemente, o lucro com essa atividade [18]. Contudo, no geral, a primeira etapa envolve a extração das moléculas de cocaína das folhas da *Erythroxylum novogranatense* ou *Erythroxylum coca* [19] por meio de solventes orgânicos. O resultado dessa extração é o que se denomina pasta base de cocaína, a qual já pode ser utilizada para obtenção de efeitos recreativos, mas também é transformada em outras formas mais puras. Cabe ressaltar que embora, comumente, os termos pasta de cocaína e base de cocaína sejam usados como sinônimos, pasta de cocaína tende a se referir ao primeiro intermediário comercial obtido no processo de manufatura, contendo ainda muitos produtos residuais ao passo que base de cocaína é utilizada para denominar o segundo intermediário, apresentando pureza maior que a pasta [20].

Por existirem muitas formas de cocaína base livre e pouca disseminação quanto a composição dos produtos e como devem ser nomeados, é comum que alguns termos sejam usados de forma equivocada por civis, autoridades e pela mídia. Essa categorização pode ser feita com base nos ingredientes utilizados e modo de produção, sendo necessário entender as possíveis mudanças introduzidas por diluição e adulterantes [21].

1.3 Formas de administração, biodisponibilidade e farmacologia da cocaína

Entre as formas de base livre, podem-se citar o Oxi e o Crack que, por possuírem baixo ponto de fusão, são consumidos pelo fumo. Já o cloridrato de cocaína, a forma de sal, possui alto ponto de fusão, sendo aspirado pelas vias aéreas ou injetado diretamente na corrente sanguínea ao ser solubilizado em água. Cada uma dessas vias de administração resulta em velocidades de absorção e biodisponibilidade diferentes [22].

Por meio das vias de administração oral e nasal a cocaína apresenta uma velocidade de absorção mais baixa, quando comparada àquelas observadas por fumar ou injetar. No caso da administração oral, a lentidão pode estar relacionada ao fato da cocaína estar ionizada no meio ácido estomacal, sendo bem absorvida somente ao alcançar o intestino delgado, ambiente mais alcalino [23]. Existem evidências de que após a administração nasal a maior parte da cocaína que alcança a circulação sistêmica seja devida à absorção oral uma vez que parte da dose utilizada seria engolida e, conseqüentemente, absorvida pelo sistema gastrointestinal [24].

A biodisponibilidade da cocaína fumada é cerca de 70% já que ocorre a pirólise e condensação dessa substância no utensílio utilizado para administração, geralmente, um cachimbo [25,26]. Já a biodisponibilidade da cocaína injetada na corrente sanguínea é praticamente 100% já que não é submetida ao metabolismo de primeira passagem no fígado, responsável pela baixa biodisponibilidade da rota oral [27].

Após injetada a cocaína produz efeito dentro de 15-30 segundos, ao passo que quando absorvida pelo alvéolos pulmonares ou membranas nasofaríngeas, atinge o seu sítio de ação em cerca de 6 segundos [28].

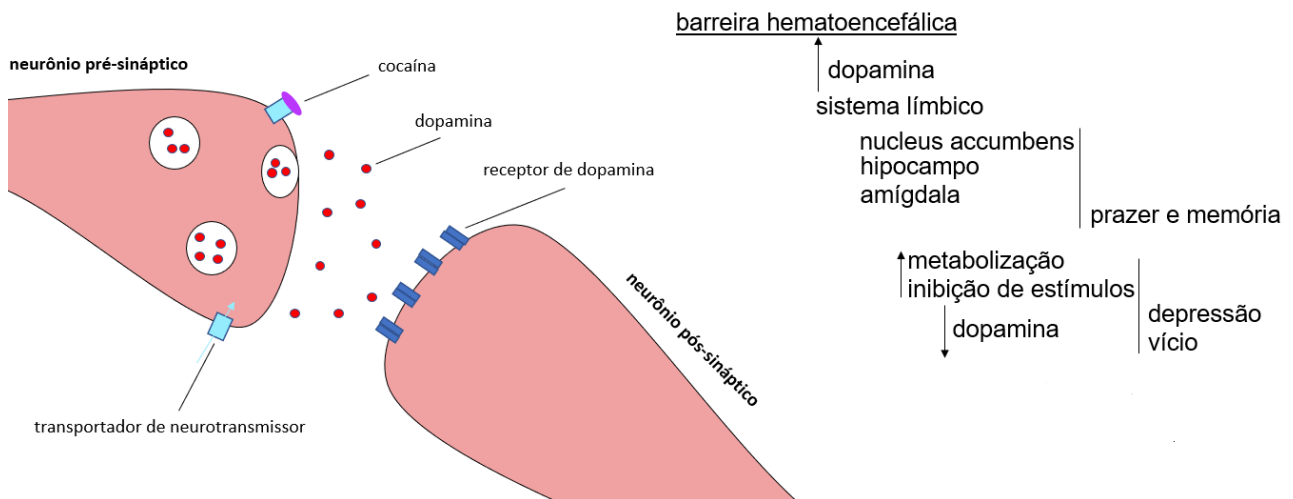
Uma das etapas essenciais à ação da cocaína é atravessar a barreira hematoencefálica, a partir da qual vai se encaminhar para o sistema límbico, mais especificamente para a região do *nucleus accumbens*. Nesse local, a cocaína bloqueia o transportador de dopamina, impedindo a captura desse neurotransmissor e aumentando a sua concentração na fenda sináptica, desencadeando uma intensa sensação de prazer [29]. A utilização da cocaína é associada a essa sensação no hipocampo e amígdala, centros de memória [30].

Na tentativa de reestabelecer concentrações normais de dopamina, mecanismos de metabolização de dopamina e inibição de estímulos [31] entre os neurônios são ativados. Alguns desses mecanismos são duradouros, reduzindo a dopamina a níveis muito baixos até um ponto denominado depleção da dopamina que ocasiona o sentimento de depressão

experimentado por usuários [32]. Também devido a algumas alterações de caráter duradouro, para obter a mesma intensidade de efeito recreativo de uma utilização anterior, gradativamente o usuário precisa consumir quantidades maiores dessa substância.

A cocaína também age sobre o sistema nervoso simpático, levando a um aumento na vasoconstrição e frequência cardíaca, o que pode desencadear uma overdose dependendo da dose administrada [33]. Efeitos agudos no sistema cardiovascular incluem arritmia e infarto agudo do miocárdio enquanto existem evidências de efeitos crônicos como cardiomiopatia, aterosclerose e doenças das artérias coronárias [34].

Figura 1 - Farmacologia da cocaína



Fonte:Elaborada pelo autor

Legenda: Após atravessar a barreira hematoencefálica, a cocaína promove aumento de dopamina na sinapse por bloquear os transportadores desse neurotransmissor. Esse efeito ocorre no sistema límbico, mais especificamente no *nucleus accumbens*, sendo responsável pela intensa sensação de prazer. Essa sensação é associada ao uso da droga nos centros de memória hipocampo e amígdala. Para reestabelecer níveis normais de dopamina, alguns mecanismos de metabolização e inibição de estímulos entre os neurônios são ativados, o que por sua vez pode levar à redução da dopamina a níveis abaixo do normal (depleção da dopamina). Esse estado é associado ao sentimento de depressão experimentado por alguns usuários, bem como à necessidade de utilizar mais cocaína, podendo levar ao vício.

O uso crônico da cocaína promove mudanças no Sistema Nervoso Central, de forma a ocasionar fenômenos como tolerância e sensibilização. A tolerância se refere justamente à necessidade de doses cada vez maiores para se obter os efeitos subjetivos associados a uma utilização anterior. Existem evidências de que seja desenvolvida por mecanismos farmacodinâmicos, uma vez que não há diferenças significativas nas taxas de absorção e metabolização da cocaína entre usuários crônicos e usuários ocasionais [35]. Além disso, observou-se que o uso contínuo reduz a capacidade da cocaína em inibir os transportadores de

dopamina e mesmo após um longo período de abstinência, uma única dose de cocaína é capaz de reestabelecer esse efeito, o que possivelmente levaria a um episódio de “recaída” [36].

Embora esse modelo de autoadministração a longo prazo que estabelece uma condição hipodopaminérgica seja capaz de explicar alguns comportamentos relacionados ao vício, também existe um modelo de autoadministração intermitente que leva à sensibilização, possivelmente um mecanismo com bases neurobiológicas diferentes das observadas na tolerância [37].

A sensibilização seria o aumento da atividade motora e das respostas comportamentais estereotipadas após a exposição repetida à cocaína, relacionada a mudanças neuroadaptativas. Essas mudanças podem ser acompanhadas a partir de padrões de descarga (oscilações) que dependem da atividade coordenada dos neurônios, isto é, da arquitetura anatômica de uma população de neurônios [38]. A intensidade de uma oscilação de baixa frequência (~ 0,8 Hz) nos disparos dos neurônios da área tegmental ventral apresentou correlação com o nível de sensibilização [39]. Fatores ambientais como locais, situações e objetos podem desencadear essa mesma exacerbação [40,41].

A cocaína é metabolizada por hidrólise espontânea e esterases teciduais para benzoilecgonina e através da bunitilcolinesterase no plasma em éster metílico de ecgonina [42]. Esses são os metabólitos majoritários, sendo ambos farmacologicamente inativos. Além deles, destaca-se a formação de norcocaína, um metabólito minoritário ativo da cocaína, promovida pela desmetilação da cocaína por enzimas hepáticas do citocromo P-450 [43]. A norcocaína pode ser metabolizada para N-hidroxinorcocaína, nitróxido de norcocaína e íon norcocaína nitrosônio. Essas espécies químicas geradas são reativas e produzem, principalmente, espécies reativas de oxigênio, responsáveis por efeitos de toxicidade observados no fígado, sistema nervoso central, sistema cardiovascular, sistema genital e nos rins [44]. Um fator que também contribui para a toxicidade dessas espécies reativas de oxigênio é a interrupção de uma parte do sistema de defesa antioxidante, a glutatona, existindo evidências de que a redução em sua atividade esteja atrelada aos comportamentos induzidos por cocaína [45].

A detecção da cocaína ou seus metabólitos pode ser realizada em fluidos corporais como saliva, urina, sangue e nos fios de cabelo, onde se acumulam. O limite de tempo para detecção de cocaína, benzoilecgonina ou cocaetileno, um metabólito derivado do uso conjunto de cocaína e álcool, nos fios de cabelo é de 6 meses após o início da abstinência [46]. Na saliva, a detecção de cocaína ou benzoilecgonina, após uso de uma dose de 100 mg de cocaína, pode ser realizada

de 1-2 dias [47]. Não é viável detectar cocaína a longo prazo em amostras de sangue ou plasma devido ao baixo tempo de meia-vida nessas matrizes [48], mas através da benzoilecgonina também é possível obter resultados na faixa de 1-2 dias após o uso. Na urina, o tempo de detecção assim como em outros fluidos depende da dose administrada, podendo a benzoilecgonina ser detectada quando administrada doses extremamente altas até 3 semanas depois [49].

1.4 Vacina para dependência de cocaína: características e mecanismo de ação

Uma vacina para dependência de cocaína ao elicitar a formação de anticorpos de alta afinidade faz com que as moléculas de cocaína sejam retidas na corrente sanguínea por não conseguirem ultrapassar a barreira hematoencefálica, sendo esse efeito dependente da concentração de anticorpos e da concentração plasmática da cocaína [50]. As moléculas retidas são metabolizadas majoritariamente em benzoilecgonina e éster metílico de ecgonina, ambos farmacologicamente inativos [51]. Contudo, a norcocaína, que se trata de um metabólito minoritário, é capaz de suscitar efeitos semelhantes ao da cocaína. Isso faz com que seja desejável que os anticorpos gerados pelo processo de imunização sejam capazes de reconhecer a cocaína, bem como seus metabólitos ativos [52].

Nesse sentido, uma vacina para dependência de cocaína consegue ajudar um usuário a manter abstinência caso seja de sua vontade abandonar o vício. Além disso, evita possíveis casos de overdose. Essas funcionalidades vão ao encontro dos objetivos da nova Política Nacional sobre Drogas determinada pelo decreto N° 9761, de 11 de Abril de 2019 na qual é instaurado que “as ações, os programas, os projetos [...] objetivarão que as pessoas mantenham-se abstinentes em relação ao uso de drogas” e “buscar equilíbrio entre as diversas frentes que compõem de forma intersistêmica a Pnad [...] em políticas de redução da demanda (prevenção, promoção e manutenção da abstinência [...], suporte social e redução dos riscos e danos sociais e à saúde [...])”

As vacinas para dependência de cocaína pertencem à classe de vacinas conjugadas uma vez que a cocaína não é uma molécula imunogênica por si só. A cocaína é, então, um hapteno, uma molécula de baixo peso molecular que, apesar de ser reconhecida por alguns anticorpos específicos, não desencadeiam etapas como geração de células B de memória [53]. Dessa forma, conjuga-se a cocaína a uma proteína carreadora, que se trata de uma proteína que já se sabe gerar uma resposta imune. Essa conjugação envolve a formação de ligação covalente entre a molécula de cocaína e algum resíduo de aminoácido na superfície da proteína. As proteínas carreadoras se diferenciam em estrutura, quantidade de anticorpos e quantidade de haptenos que podem carregar [54].

Inicialmente, a estrutura hapteno-proteína carreadora é capturada por uma célula dendrítica folicular, uma célula apresentadora de antígenos (APC), e encaminhada para os linfonodos, onde dois tipos celulares primordiais à resposta humoral estão presentes: célula B e linfócito T auxiliar [55]. Então, a estrutura hapteno-proteína carreadora é apresentada ao

receptor da célula B que se trata de um anticorpo que reconhece a estrutura do hapteno ou da proteína carreadora. Paralelamente, as moléculas de HLA da Classe II são sintetizadas no retículo endoplasmático e associadas a uma proteína denominada cadeia invariante. O complexo de Golgi é responsável por transportá-las em vesículas exocíticas que se fundem com os endossomos. Ao início do transporte, ocorre a clivagem da cadeia invariante deixando apenas um pequeno fragmento que ocupa o sítio de ligação do HLA da Classe II, o peptídeo de classe invariante associado à classe II (CLIP) [56].

A partir daí a estrutura hapteno-proteína carreadora é submetida ao processamento mediado por proteases, como catepsinas, em ambiente ácido, após fusão do endossomo com lisossomo [57]. Dessa forma, são gerados peptídeos que podem ou não estar conjugados a molécula de cocaína. O CLIP é removido pela ação da molécula HLA-DM, possibilitando que os peptídeos gerados a partir dos antígenos se liguem à fenda do HLA da Classe II.

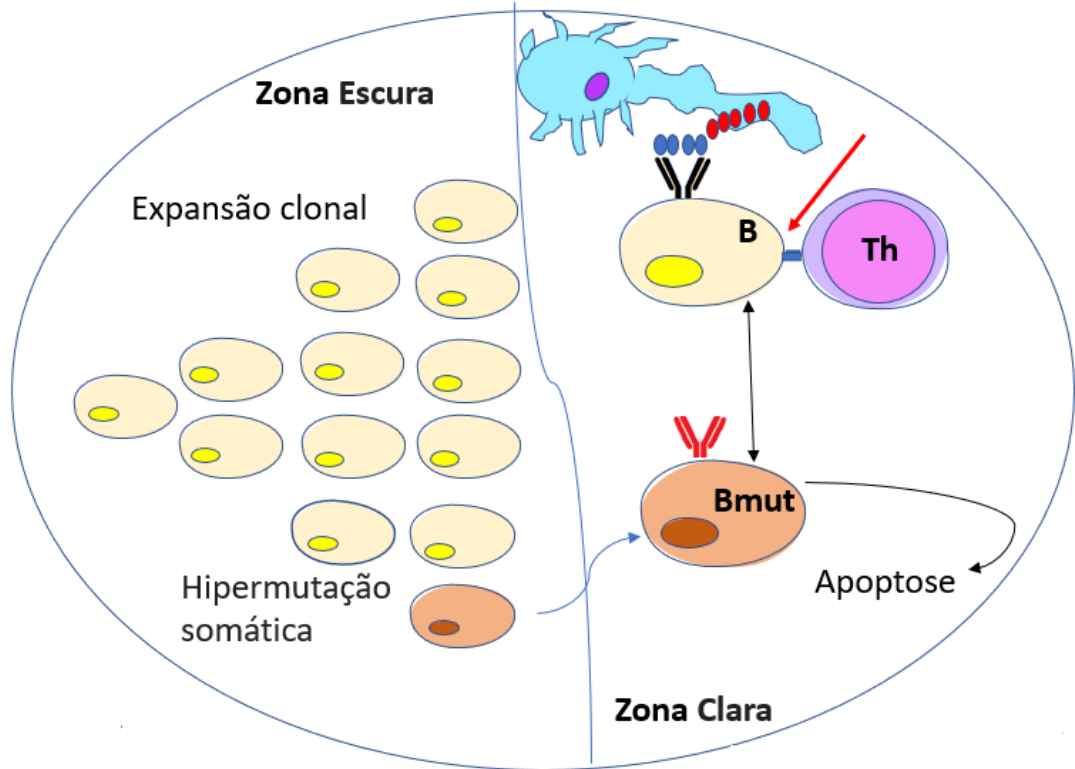
O complexo peptídeo-HLA da Classe II migra para a superfície da célula B, onde vai ser apresentado para o receptor do linfócito T auxiliar. Assim, dois mecanismos de apresentação são possíveis: o de um peptídeo não conjugado à cocaína e o de um peptídeo conjugado à cocaína [58]. O primeiro possui comprovação enquanto o segundo foi observado apenas uma vez durante um experimento com vacinas para polissacarídeos, que também são vacinas conjugadas [59].

Quão maior a estabilidade do complexo peptídeo-HLA da Classe II, maior a chance do peptídeo ser capaz de estimular o linfócito T auxiliar, o qual é responsável por liberar citocinas, sinalizadores que regulam o processo de maturação de afinidade [60].

As etapas até aqui descritas ocorrem na zona clara de uma estrutura denominada centro germinal. A célula B que contém o anticorpo que reconhece alguma porção da estrutura hapteno-proteína carreadora migra para a zona escura, região em que ocorre a expansão clonal dessa célula B, formando células idênticas que sejam igualmente capazes de estimular a liberação de citocinas. Ao mesmo tempo, algumas dessas células passam por hipermutação somática, uma mutação nos genes que altera aminoácidos localizados no sítio de ligação do receptor da célula B [61]. Essas células B mutadas retornam a zona clara para serem testadas. Caso sejam capazes de reconhecer alguma porção da sequência da estrutura hapteno-proteína carreadora e promover, por fim, a liberação de citocinas por linfócitos T auxiliares, essas células mutadas também são submetidas a expansão clonal. Caso contrário, elas sofrem apoptose celular. Nesse sentido, essas mutações auxiliam na produção de anticorpos de maior afinidade já que as células

B primárias e as células B mutadas competem entre si, o que dá nome ao processo “maturação de afinidade” [62].

Figura 2 - Maturação de afinidade no centro germinal



Fonte: Elaborada pelo autor.

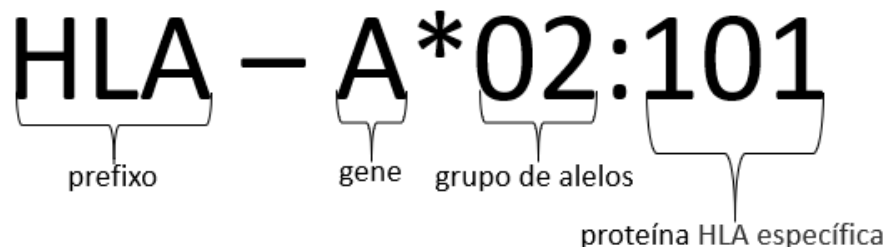
Legenda: Na zona clara, as células B com anticorpos distintos em sua superfície competem pela estrutura hapteno-proteína carreadora apresentada por uma célula apresentadora de antígenos (em azul). As células B que possuem maior afinidade pelo antígeno e conseguem estimular a liberação de citocinas pelas células T auxiliares migram para a zona escura onde são submetidas a expansão clonal e hipermutação somática, esta última gerando células B com mutações nos anticorpos. Essas novas células B migram para a zona clara, passando pela seleção já descrita. As células B que não forem capazes de promover a ativação de células T auxiliares sofrem apoptose.

As moléculas de HLA da classe II possuem sítio de ligação aberto o que possibilita que se liguem a um repertório de peptídeos de diversos tamanhos [63]. Nos humanos, elas são codificadas em três loci principais: HLA- DP, -DQ e -DR. Cada molécula é constituída por duas cadeias, α e β , cada uma produzida pela tradução de um gene diferente. O polimorfismo desses genes é uma característica bem notável, principalmente para os genes que codificam a cadeia β [64]. Para o gene DRB que codifica a cadeia β da molécula de HLA-DR, já foram descobertos

e nomeados 4419 alelos possíveis que se traduzem em 2903 moléculas diferentes [65]. Os aminoácidos que se alteram de uma sequência para outra podem estar presentes no sítio de ligação ou nas regiões que o flanqueiam, influenciando diretamente nas interações intermoleculares que podem ser estabelecidas e, conseqüentemente, no repertório de peptídeos reconhecido por cada molécula. Destarte, deseja-se que uma vacina seja promíscua no sentido de gerar peptídeos que se liguem com afinidade ao maior número de alelos possíveis. Isso garante uma cobertura maior sobre a população [66].

Os genes do HLA se expressam de maneira codominante, isto é, os alelos herdados do pai e da mãe se expressam igualmente. Cada indivíduo herda um par de genes HLA-DQ (DQA1 e DQB1, codificando as cadeias α e β), um par de genes HLA-DP (DPA1 e DPB1, codificando as cadeias α e β), o gene HLA-DRA α (DRA1) e um ou dois genes HLA-DR β (normalmente, DRB1 ou DRB3, -4 ou -5). Então, um indivíduo heterozigoto pode herdar seis ou oito alelos de MHC de classe II, três ou quatro de cada progenitor. Além disso, algumas moléculas DQ α codificadas em um cromossomo podem se associar com moléculas DQ β codificadas em outro cromossomo, resultando em um número de moléculas expressas de HLA de classe II maior que o esperado [56].

A nomenclatura dos alelos ocorre conforme a seguinte regra [67]:



A numeração do grupo de alelos e da proteína HLA específica surgem, respectivamente, dos métodos de tipificação sorológico e molecular. O método sorológico mais comum é o método de citotoxicidade celular mediada por anticorpo e dependente de complemento. Essa técnica reconhece similaridades estruturais entre alelos, agrupando-os, porém, não sendo capaz de distingui-los individualmente. Dessa forma, para esse fim, são utilizadas técnicas baseadas em PCR e sequenciamento de DNA que conseguem determinar a sequência primária de cada proteína [68].

1.5 Objetivos

O objetivo geral é estudar potenciais candidatos a vacinas para a dependência de cocaína já testadas *in vivo* de forma a compreender e definir propriedades bioquímicas que contribuam para o sucesso desse tipo de vacina.

Os objetivos específicos podem ser elencados como:

- Entender as limitações e potencialidades relacionadas à formulação estrutural de vacinas conjugadas, principalmente aquelas pertinentes à eficácia *in vivo* e à interação com as moléculas HLA de classe II.
- Determinar a natureza e a estabilidade das interações intermoleculares estabelecidas na apresentação do antígeno pelas moléculas de HLA de classe II;
- Investigar os mecanismos de apresentação de peptídeos conjugados e não conjugados;
- Guiar o desenvolvimento de uma vacina para dependência de cocaína, oferecendo percepções sobre os caminhos a serem ou não seguidos.

2. Capítulo 2: TÉCNICAS *IN SILICO* PARA INVESTIGAÇÃO DO PROBLEMA

2.1 Docking

O *docking* consiste em uma técnica que prediz a conformação e orientação de um possível ligante dentro do sítio de ligação de uma molécula alvo. Essa predição pode ser compreendida por meio de dois processos principais: amostragem conformacional do ligante e avaliação das interações intermoleculares [69]. A amostragem conformacional do ligante é realizada por meio de um algoritmo de busca que se ramifica em três categorias: sistemático, estocástico ou determinístico. O algoritmo de busca sistemático explora basicamente todo o espaço conformacional por alterar os parâmetros estruturais de forma incremental, o que pode levar a um problema de explosão combinatória [70]. Esse problema pode ser contornado ao substituir a busca sistemática exaustiva por uma busca baseada em fragmentos, na qual um fragmento é tratado como rígido e ancorado a uma região complementar no sítio de ligação e os outros são acrescentados a ele enquanto são submetidos ao algoritmo de busca [71]. Outra alternativa possível é a geração prévia de um *ensemble* conformacional, restringindo assim o número de poses candidatas [72].

O algoritmo de busca estocástico altera aleatoriamente os graus de liberdade (translacional, rotacional e conformacional) do ligante gerando poses que são avaliadas por um critério probabilístico. Apesar de gerar poses diversificadas, esse algoritmo não garante a convergência para um mínimo global, sendo necessário então que seja performado múltiplas vezes para garantia de um resultado ótimo [73]. No algoritmo de busca determinístico, o estado do ligante a cada iteração é determinado pelo estado anterior, levando a uma minimização da energia que pode aprisionar o sistema em um mínimo local [74]. Como exemplo de algoritmo de busca estocástico, pode-se citar o algoritmo genético em que valores de variáveis são utilizados para descrição da translação e orientação do ligante, sendo sua manipulação e interpretação associados a conceitos biológicos [75].

Grolmusz e Ördög (1994) [76] explicam como as ideias baseadas na linguagem da genética e da evolução podem ser aplicadas ao desenvolvimento de um algoritmo

Um indivíduo é um ponto no espaço de busca, e seu genótipo é a fita de números (ou vetor) que o descreve. O fenótipo é a coleção de atributos do indivíduo, e sua aptidão é o valor da função correspondente ao indivíduo. Uma população é simplesmente uma coleção de indivíduos. O algoritmo primeiro seleciona uma população de indivíduos (geralmente aleatórios) que formam a primeira geração,

então entra em um ciclo de derivar a *n*-ésima geração de indivíduos a partir da geração anterior. Toda geração possui o mesmo tamanho fixo. Um ciclo do algoritmo realiza uma seleção e aplica operadores de variação aos indivíduos da população atual.

No *docking* molecular, o genótipo corresponde aos parâmetros que descrevem a translação, orientação e conformação do ligante em relação à molécula receptora, ao passo que o fenótipo equivale às coordenadas atômicas. A aptidão é avaliada por uma função de energia que avalia as interações envolvidas no processo de ligação, sendo utilizada como critério durante a seleção para decidir quais poses permanecem a cada iteração. As poses de cada geração podem sofrer mutação (alteração aleatória de algum parâmetro do genótipo) ou crossover (quando um par de poses gera novas poses com parâmetros herdados de ambos precursores) [77].

Quanto às funções de pontuação, elas podem se encaixar em cinco categorias: i) baseadas em físicas, ii) em métodos de mecânica quântica, iii) em conhecimento, iv) em aprendizado de máquina ou v) empíricas [78].

As funções de pontuação baseadas em física se dividem em duas outras categorias: baseadas em campo de força e em métodos de mecânica quântica. As funções de pontuação baseadas em campo de força consistem em equações de energia potencial que descrevem as interações eletrostáticas e de van der Waals entre os pares de átomos proteína-ligante. Naturalmente, essas funções não levam em consideração a contribuição entrópica associada à mudança torcional do ligante e ao efeito do solvente, sendo aprimoradas pela criação de modelos que levam em consideração esses fatores entrópicos [79].

As funções de pontuação baseadas em métodos de mecânica quântica geralmente levam a uma melhor acurácia em comparação à mecânica molecular já que descrevem fenômenos que requerem o tratamento explícito dos elétrons: formação de ligação covalente, polarização, transferência de carga, rearranjo de ligações químicas e ligações de halogênio. Devido ao alto custo computacional desses cálculos, é mais comum que seja empregada uma versão híbrida de mecânica quântica e mecânica molecular, sendo a primeira aplicada somente ao ligante e seus arredores [80].

Já as funções de pontuação empíricas consistem em equações em que são atribuídos coeficientes a cada fator energético, sendo estes otimizados a partir de uma regressão realizada sobre dados de afinidade proteína-ligante de um conjunto de treinamento bem caracterizado [81]. Representa uma alternativa computacionalmente eficiente devido à simplicidade dos termos e permite que descritores de qualquer tipo sejam correlacionados com

a energia livre de ligação [82]. O AutoDock, por exemplo, emprega uma função de pontuação classificada em seu guia como semiempírica por possuir a estrutura associada a um campo de força, porém sendo ponderada por constantes ajustadas a valores experimentais (equação 1) [83].

$$S = W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{hb} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{ele} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \\ + W_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{\left(-\frac{r_{ij}^2}{2\sigma^2} \right)}$$

Equação 1: Função de pontuação semiempírica do Autodock 4.2

Na Equação 1, as constantes W_{vdw} , W_{hb} , W_{ele} e W_{sol} foram calibradas conforme a energia livre empírica obtida de constantes de ligação obtidas experimentalmente.

O primeiro termo é um potencial 6/12 (Lennard-Jones) que descreve interações intermoleculares dispersivas (forças de Van der Waals) e repulsivas cujos parâmetros são baseados no campo de força Amber. O segundo termo, por sua vez, é um potencial 10/12 voltado para as ligações hidrogênio, no qual os parâmetros C e D são ajustados para atribuir o máximo de energia potencial de 5 kcal mol⁻¹ em ligações hidrogênio de 1,9 Å com oxigênio ou nitrogênio, e o máximo de energia potencial de 1 kcal mol⁻¹ a 2,5 Å com enxofre. A função E(t) é responsável por averiguar a direcionalidade com base no ângulo t da geometria ideal de uma ligação hidrogênio. O terceiro termo descreve o potencial de Coulomb enquanto o último é um potencial de dessolvatação baseado no volume dos átomos (V) que circunvizinham um dado átomo e o isola do solvente, ponderado por um parâmetro de solvente (S) e um termo exponencial dependente de distância.

As funções de pontuação baseadas em conhecimento derivam os potenciais desejados por meio das estruturas tridimensionais de um grande conjunto de complexos proteína-ligante de acordo com o inverso do princípio estatístico de Boltzmann. Assim, assume-se que a frequência de um par de átomos em diferentes distâncias está relacionada à posição ótima dos átomos dos ligantes em relação aos átomos da molécula receptora e se converte essa frequência em um potencial de força média [84].

Por último, existem também as funções de pontuação baseadas em aprendizado de máquina que incluem a utilização de diversos algoritmos como máquina de vetores de suporte, floresta aleatória, rede neural artificial, e aprendizagem profunda [78].

2.2 Quimiometria

A Análise de Componentes Principais (PCA) é uma técnica de aprendizagem não supervisionada aplicada sobre um conjunto de dados de alta dimensão com o intuito de diminuir a dimensionalidade e possibilitar a observação de tendências e padrões [85].

As variáveis originais são combinadas linearmente para formar variáveis latentes, também conhecidas como componentes principais. Os pesos que ponderam as variáveis originais são otimizados de modo a maximizar a variância observada na variável latente, o que significa que a maior parte da informação relevante está sendo retida. Cada variável latente é construída ortogonalmente às demais, de forma sempre a abranger a variação ainda não explicada. Trata-se de um problema de álgebra linear em que se busca o autovetor de maior valor da matriz de covariância [86].

Os elementos dos autovetores (pesos) são denominados *loadings* por meio dos quais é possível determinar a contribuição de cada variável original para a construção das componentes principais enquanto os elementos da combinação linear são denominados *scores*, sendo correspondentes ao valor das amostras (coordenadas) no novo espaço dimensional [87]. A proximidade de duas amostras no espaço definido pelas componentes principais indica uma similaridade entre elas, assim como a posição dos vetores unitários das variáveis originais define uma correlação positiva ou negativa observada na representação gráfica dos *loadings* [88].

Os dados podem ser submetidos a um pré-processamento antes da análise primária, o qual consiste em uma operação matemática com o intuito de remover fontes indesejáveis de variação que podem afetar o resultado da análise de dados, como dados faltantes, ruído, variáveis com grande diferença de intensidade, linha de base, efeitos multiplicativos e mudanças de pico [89]. O pré-processamento pode ser realizado sobre as amostras (linhas da matriz de dados) ou sobre as variáveis (colunas da matriz de dados). Os principais tipos de pré-processamento sobre as variáveis incluem centrar os dados na média e autoescalar. Dados centrados na média são obtidos quando a média de cada variável é subtraída de seus respectivos elementos ao passo que autoescalar envolve centrar na média seguido pela divisão de cada elemento de uma variável pelo seu desvio padrão [90].

A regressão por mínimos quadrados parciais (PLS) consiste em performar uma PCA sobre a matriz de variáveis independentes e, depois, utilizar as componentes principais obtidas como regressores da matriz de variáveis dependentes. Como as componentes principais são ortogonais, esse tipo de regressão é adequada para os casos em que o número de observações é

maior que o número de preditores, o que implica em multicolinearidade entre as variáveis [91]. Contudo, não há garantia que as primeiras componentes principais que explicam a variação da matriz de variáveis independentes sejam relevantes para explicar a informação da matriz de variáveis dependentes. A estratégia empregada na PLS é a decomposição simultânea de ambas matrizes de forma que as componentes extraídas expliquem o máximo da covariância entre elas [92].

O cálculo do erro de validação pode ser feito por meio de amostras independentes que não foram utilizadas durante a etapa de calibração do modelo de regressão ou pelo emprego de validação cruzada. A validação cruzada consiste em deixar n-amostras fora do conjunto de validação, utilizando-as posteriormente para mensurar a capacidade preditiva do modelo. Esse procedimento é repetido até que todas as amostras sejam empregadas no conjunto de validação.

2.3 DFT

A mecânica quântica provém de uma descrição de sistemas químicos a partir da aplicação de operadores apropriados sobre uma função de onda ψ cujo resultado são variáveis observáveis.

O produto da função de onda ψ com seu complexo conjugado ψ^* ($\psi|\psi^*$) possui unidade de densidade de probabilidade. A probabilidade que um sistema químico seja encontrado em uma região do espaço multidimensional é igual à integral de ($|\psi|^2$) sobre essa região do espaço [93,94].

Alguns postulados definem restrições quanto ao que constitui uma função de onda aceitável: a integral normalizada de $|\psi|^2$ sobre todo o espaço deve ser igual a 1, o que significa que a função de onda deve ser quadraticamente integrável. Além disso, ψ deve ser contínua e ter um único valor associado a cada ponto do domínio [95].

O operador que retorna a energia E do sistema é denominado Hamiltoniano H , a partir do qual podemos descrever a equação de Schrödinger independente do tempo, não-relativística (equação 2).

$$H\psi = E\psi$$

Equação 2: Equação de Schrödinger independente do tempo, não-relativística

A forma mais comum do operador Hamiltoniano, capaz de descrever sistemas com múltiplos elétrons interagindo, leva em consideração cinco contribuições para a energia total do sistema: energia cinética dos elétrons; energia cinética dos núcleos; atração dos elétrons e núcleos; repulsão intereletrônica e repulsão internuclear [96] (equação 3).

$$H = - \sum_i \frac{\hbar}{2m_e} \nabla_i^2 - \sum_k \frac{\hbar}{2m_k} \nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i < j} \frac{e^2}{r_{ij}} + \sum_{k < l} \frac{e^2 Z_k Z_l}{r_{kl}}$$

Equação 3: Operador Hamiltoniano

onde i e j percorrem os elétrons, k e l percorrem os núcleos, \hbar é a constante de Planck dividida por 2π , m_e é a massa do elétron, m_k é a massa do núcleo k , ∇^2 é o operador Laplaciano, e é a carga de um elétron, Z é o número atômico e r_{ab} é a distância entre partículas a - b .

A correlação do movimento entre as partículas faz com que seja difícil expressar funções de ondas precisas. Uma forma de simplificar esse problema é adotar a aproximação de Born-Oppenheimer [97], a qual assume que o movimento dos núcleos é mais lento que o movimento dos elétrons. Assumindo-se posições fixas para os núcleos, o termo da energia cinética nuclear se torna independente dos elétrons, a correlação do termo potencial de atração elétron-núcleo é

eliminada e o termo potencial de repulsão núcleo-núcleo se torna apenas uma constante para determinada geometria [98].

Uma função tentativa Φ pode ser construída a partir de uma função de variação linear, isto é, uma combinação linear de funções de onda já conhecidas φ_i (conjunto de bases) ponderadas por coeficientes a_i (equação 4). Essa abordagem dá origem a métodos como a Combinação Linear de Orbitais Atômicos (LCAO), muito utilizada para construção de orbitais moleculares [99].

$$\Phi = \sum_{i=1}^N a_i \varphi_i$$

Equação 4: Função de Variação Linear

Determinado o conjunto de bases para descrever o sistema químico, buscam-se por coeficientes que possam minimizar a energia para todas as combinações lineares possíveis do conjunto de bases. A condição para uma função estar em seu mínimo é que suas derivadas em relação a suas variáveis independentes seja zero.

O Princípio Variacional (equação 5) provê um parâmetro que permite a escolha de uma função Φ que descreva o estado fundamental de um sistema (mesmo que as funções ψ_i sejam desconhecidas). Uma vez que qualquer função de onda sempre apresente energia maior ou igual àquela do estado fundamental, a qualidade das funções de onda, determinadas por quaisquer métodos, podem ser determinada pelo seu valor de energia associado, sendo que valores menores indicam que a função Φ se aproxima mais do comportamento real do sistema [100].

$$\frac{\int \Phi H \Phi dr}{\int \Phi^2 dr} \geq E_0$$

Equação 5: Princípio Variacional

A partir do formalismo de um elétron, o qual assume que a energia de um sistema de muitos elétrons pode ser determinada simplesmente através da soma de orbitais ocupados com um único elétron, desenvolveu-se um modelo de aproximação capaz de lidar com os efeitos eletrônicos: as funções de onda produto Hartree [101].

O hamiltoniano Hartree assimila a repulsão intereletrônica por meio de um termo que representa a interação de um elétron com um campo potencial médio representado pelos demais elétrons [102] (equação 6).

$$h_i = -\frac{1}{2} \nabla_i^2 - \sum_{k=1}^M \frac{Z_k}{r_{ik}} + \sum_{j \neq i} \int \frac{\rho_j}{r_{ij}} dr$$

Equação 6: Hamiltoniano Hartree

onde i e j percorrem os elétrons, k percorre os núcleos, ∇ é o operador Laplaciano, Z é o número atômico, r_{ab} é a distância entre partículas a e b e ρ_j é a densidade da carga associada ao elétron j .

Hartree propôs o algoritmo iterativo chamado método do campo autoconsistente (SCF). O primeiro passo do processo SCF é a suposição das funções de onda para todos os orbitais moleculares ocupado, utilizando-as para montar os operadores hamiltonianos de um elétron necessários. A solução da equação de Schrödinger fornece um novo conjunto de funções de onda. Dessa forma, os hamiltonianos de um elétron passam a ser formados por funções de onda mais acuradas para obter o campo potencial médio, sendo esse processo repetido para obter gradativamente melhores funções de onda. O algoritmo só chega ao fim quando a diferença entre dois conjuntos de funções de onda sucessivos se encaixa em determinado critério limite, idealmente até a função de início e de saída do algoritmo seja as mesmas [103,104].

A princípio, a função de onda produto-Hartree não satisfazia ao princípio de Pauli que determina que as funções de onda sejam antissimétricas, isto é, o sinal das funções de onda eletrônicas deve mudar sempre que as coordenadas de dois elétrons são trocadas. Isso também quer dizer que não pode existir mais de um elétron descrito pelos mesmos números quânticos. Modificações para contornar esse problema foram realizadas por Fock e, por essa razão, o método passou a ser denominado Hartree-Fock SCF [99].

Em termos práticos, o método Hartree-Fock SCF não é aplicado a moléculas grandes, principalmente, devido à transformação das integrais de repulsão de dois elétrons de quatro índices, envolvida no cálculo de efeitos de correlação [105]. Cada índice é executado sobre o número total de funções de base, sendo assim, existem a princípio N^4 integrais para serem avaliadas. Esse comportamento que escala conforme o tamanho do conjunto de bases representa um empecilho para a teoria Hartree-Fock [106].

A partir desse entrave, existem duas abordagens principais: as teorias semiempíricas e a teoria ‘*ab initio*’. As teorias semiempíricas propõem a aproximação de integrais de dois elétrons por meio da parametrização de valores experimentais para simplificar a solução das equações HF [107]. Enquanto isso, a teoria ‘*ab initio*’ acredita que as equações HF no limite de um conjunto de bases infinito são o caminho para a solução exata da equação de Schrödinger, sendo necessário uma correção à energia de correlação eletrônica que não é tratada pela teoria HF por si só [108].

Uma alternativa ao uso das funções de onda para determinação da energia e demais propriedades de um sistema químico é a utilização da densidade eletrônica por meio da teoria denominada Teoria do Funcional de Densidade (DFT). Hohenberg–Kohn conseguiram

demonstrar que a construção de um hamiltoniano é possível pois toda a informação química de um sistema está contida na densidade eletrônica: a integral da densidade define o número de elétrons, os máximos da densidade definem as posições dos núcleos e a altura dos máximos define as cargas nucleares [109]. Além disso, eles demonstraram que a densidade eletrônica obedece ao princípio variacional. Kohn e Sham foram responsáveis por definir o funcional de densidade [110] (equação 7).

$$E[p(r)] = T_{ni}[p(r)] + V_{ne}[p(r)] + V_{ee}[p(r)] + \Delta T[p(r)] + \Delta V_{ee}[p(r)]$$

Equação 7: Funcional de Densidade

Os termos se referem respectivamente à: energia cinética para elétrons não interagentes, interação elétron-núcleo, repulsão clássica elétron-elétron, correção para a energia cinético devido à natureza interativa dos elétrons e correção não clássica para o termo de repulsão elétron-elétron. Os dois últimos termos geralmente são tratados juntos num termo denominado energia de troca e correlação, E_{xc} . As diferentes possibilidades de tratamento dessa energia dão origem aos vários funcionais existentes.

Uma das abordagens para obtenção de ε_{xc} é a Aproximação de Densidade Local (LDA), na qual se assume que o valor de ε_{xc} em qualquer posição \mathbf{r} pode ser calculado exclusivamente através do valor de p naquela posição [111]. Entre os funcionais de densidade que se encaixam nessa definição, aquele com mais aplicações é derivado da análise de um gás de elétrons homogêneo, onde a densidade possui o mesmo valor em toda posição. Quando esse funcional leva em consideração a polarização de spin, geralmente recebe a nomenclatura Aproximação de Spin-Densidade Local (LSDA) [112].

Como tipicamente a densidade eletrônica está longe de assumir uma uniformidade espacial, a abordagem LDA apresenta limitações. Assim, desenvolveram-se funcionais que calculam a energia de troca e correlação a partir do valor da densidade e do gradiente da densidade, denominados “Aproximação do Gradiente Generalizado” (GGA) [113]. Geralmente, esses funcionais são construídos com a correção do gradiente sendo adicionada ao valor do funcional LDA.

Além das aproximações propostas acima, existem os funcionais híbridos que são considerados métodos de conexão adiabática. Nesses métodos, a energia de troca e correlação é obtida por meio da contribuição da energia exata de um sistema de elétrons fracamente interagentes, dada pelo cálculo da energia Hartree-Fock E_x^{HF} utilizando os orbitais Kohn-Sham, acrescida de uma contribuição E_{xc}^{DFT} determinada por algum funcional DFT. Essas contribuições são ajustadas por uma constante α a ser otimizada [114] (equação 8).

$$E_{xc} = (1 - \alpha)E_{xc}^{DFT} + \alpha E_x^{HF}$$

Equação 8: Funcionais Híbridos

Como extensão dessa proposta e forma de aumentar a exatidão dos cálculos, outros funcionais que levam em consideração mais parâmetros ajustáveis foram criados, como o B3LYP [115,116] (equação 9).

$$E_{xc}^{B3LYP} = (1 - a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^B + (1 - c)E_c^{LSDA} + cE_c^{LYP}$$

Equação 9: Funcional B3LYP

onde $a = 0,20$, $b = 0,72$, $c = 0,81$, E_x^{LSDA} é a energia de troca para o funcional VWN3 LSDA, E_x^{HF} é a energia de troca obtida pelo método HF, ΔE_x^B é a correção da energia de troca proposta por Becke, E_c^{LSDA} é a energia de correlação para o funcional VWN3 LSDA e E_c^{LYP} é a energia de correlação para o funcional Lee-Yang-Parr.

Esse funcional se popularizou principalmente pela boa capacidade preditiva de propriedades moleculares, sendo bem validado para sistemas orgânicos [117]. E mesmo após o desenvolvimento de funcionais que corrigem as desvantagens apresentadas no B3LYP, ele continua se destacando pelo relativo baixo custo computacional que compensa seus defeitos [118].

2.4 Dinâmica Molecular

A Dinâmica Molecular consiste em um método de simulação computacional que permite acompanhar o movimento de todos os átomos de um sistema molecular ao longo do tempo por meio de um modelo físico que rege as interações interatômicas.

Um dos princípios da Dinâmica Molecular é a solução numérica da equação de movimento de Newton cujo resultado provém os novos vetores de posição a cada passo da simulação [119]. Inicialmente, calcula-se a força atuante sobre cada átomo por intermédio da primeira derivada da energia potencial em relação às posições dos átomos. Dessa forma, determina-se a aceleração com base nas forças calculadas e na massa já conhecida dos átomos. Integrando-se a aceleração, obtém-se a velocidade cuja integral, por sua vez, resulta no deslocamento das partículas. Essas novas variáveis definem a energia potencial e cinética do sistema, a partir das quais será realizada sucessivas integrações da equação de movimento que definirão uma trajetória durante o tempo de simulação definido [120].

A energia potencial do sistema é definida a partir de um campo de força, uma função matemática que estabelece a dependência dessa grandeza com as coordenadas das partículas. Um típico campo de força possui termos referentes a comprimento e ângulos de ligação, ângulos diedros, torções, interações van der Waals e de Coulomb [121]. Como exemplo, temos os campos de forças GROMOS [122,123] que assumem a forma definida pela equação 10.

$$\begin{aligned}
 E_{potencial} = & \sum_{ligações} \frac{K_{b,i}}{4} (b_i^2 - b_{i,0}^2)^2 + \sum_{ângulos} \frac{K_{\theta,i}}{2} (\cos \theta_i - \cos \theta_{i,0})^2 \\
 & + \sum_{impróprios} \frac{K_{\xi,i}}{2} (\xi_i - \xi_{i,0})^2 + \sum_{diedros} K_{\varphi,i} [1 + \cos \delta_i \cos(m_i \varphi_i)] \\
 & + \sum_{pares\ de\ átomos} 4\varepsilon_{i,j} \left[\left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^6 \right] + \frac{q_i q_j}{4\pi \varepsilon_0 r_{i,j}} + RF_{i,j}
 \end{aligned}$$

Equação 10: Forma funcional de um típico campo de força GROMOS

Na equação 10, o primeiro termo avalia a variação do comprimento das ligações covalentes em relação ao seu valor de equilíbrio, $K_{b,i}$ é uma constante de força em $\text{kJmol}^{-1}\text{nm}^{-4}$. O segundo termo representa a deformação do ângulo entre átomos ligados covalentemente em relação ao valor de equilíbrio, $K_{\theta,i}$ é uma constante elástica em kJmol^{-1} .

O terceiro termo leva em consideração ângulos impróprios que basicamente envolvem qualquer configuração angular não baseada em quatro átomos ligados sucessivamente, sendo

$K_{\xi,i}$ a constante que define a altura da barreira de rotação. O quarto termo abrange a torção própria dos ângulos diedros, $K_{\phi,i}$ é a barreira de rotação, m é o número de mínimos para a torção e δ é a diferença de fase do ângulo diedro (0 ou 180°). O sexto termo é a interação para átomos não ligados covalentemente, definida pelo potencial de Lennard-Jones. O penúltimo termo é o potencial de Coulomb e o último termo é a contribuição do método *reaction field* [124], utilizado para tratar as interações eletrostáticas de longo alcance. A partir de certo limite, o sistema começa a ser tratado como homogêneo por meio da representação de um contínuo dielétrico, capaz de ser polarizado pela distribuição de cargas do “soluto”.

A fim de evitar efeitos de superfície que poderiam ser ocasionados por um tratamento isolado de um sistema finito, introduzem-se as chamadas condições periódicas de contorno. Essas condições consistem na replicação da caixa que contém o sistema químico, gerando imagens idênticas e equidistantes deste em todas as direções do espaço. Assim, durante as simulações, quando um átomo se move na caixa original, se move da mesma maneira em sua imagem periódica, o que leva a uma conservação da quantidade de átomos na caixa principal [125].

Os sistemas moleculares podem englobar fenômenos que variam em escala de tempo, o que é importante ser observado durante a determinação dos passos de integração que correspondem ao intervalo de tempo, Δt . Desse modo, Δt deve ser menor que o tempo correspondente ao(s) evento(s) de interesse bioquímico, sendo comum que assumam valores na faixa dos femtossegundos [126]. Com o intuito de aumentar os passos temporais de integração e, conseqüentemente, reduzir o custo computacional dessa técnica, podem-se aplicar restrições moleculares que restringem os graus de liberdade que correspondem aos movimentos de maior frequência como os vibracionais [127].

As simulações são capazes de auxiliar na avaliação da flexibilidade de uma biomolécula com estrutura determinada, o que permite a identificação de outras conformações que podem estar envolvidas no desempenho de sua função [128]. Por outro lado, conseguem refinar estruturas modeladas computacionalmente para que se aproximem da estrutura nativa [129]. As poses de ligações derivadas de *docking* molecular podem ser complementarmente analisadas por dinâmica molecular, sendo capaz de aumentar o poder preditivo de ligantes [130].

Além disso, é possível determinar como o sistema sob análise se comporta quando submetido a alguma perturbação, o que inclui os ligantes ligados a ele, mutações ou

modificações pós-traducionais, presença de moléculas no ambiente, estado de protonação, temperatura, voltagem através da membrana, entre outros [131].

Um dos passos da Dinâmica Molecular é o equilíbrio NPT e NVT que qualificam o sistema, respectivamente, em um *ensemble* isotérmico-isobárico e *ensemble* canônico. *Ensembles* são uma coleção de todos os possíveis estados de um sistema com estados microscópicos distintos, mas que representam um único estado termodinâmico [132]. O ensemble canônico garante que um número constante de átomos N, assim como um volume constante V e uma temperatura constante T sejam mantidos durante a simulação. Já o ensemble isotérmico-isobárico assegura que um número constante de átomos N, assim como uma pressão constante P e uma temperatura constante T sejam mantidos durante a simulação. Dessa maneira, a calibração permite uma comparação mais direta com os processos do mundo real que acontecem sob condições semelhantes, além de prover um jeito de lidar com as flutuações do sistema [133].

Em termos de análise de resultados, uma das técnicas mais comuns para determinar se o sistema químico realmente atingiu o equilíbrio (convergência) é a Raiz Quadrada do Desvio Quadrático Médio (RMSD) entre duas estruturas estáticas (equação 9).

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_i^X - r_i^Y)^2}$$

Equação 11: RMSD

onde N é igual ao número de átomos, i é o i-ésimo átomo, r^X é a estrutura alvo (os frames da dinâmica molecular) e r^Y é a estrutura de referência (ponto de início da simulação).

Além disso, é comum realizar uma investigação sobre a formação de ligações hidrogênio entre o ligante e a proteína receptora, pois se trata de uma força intermolecular que confere bastante estabilidade ao complexo final.

5. CONCLUSÕES

Em relação aos objetivos desse trabalho, conclui-se que os métodos computacionais foram hábeis em auxiliar no estudo aqui proposto. Na primeira parte desse estudo, embora os descritores físico-químicos e biológicos calculados para os peptídeos tenham sido suficientes para distinguir a proteína subunidade B da toxina da Cólera das demais proteínas carreadoras, eles não foram adequados para definir uma separação entre os alelos. Posteriormente, os descritores provenientes do *docking*, voltados ao caráter da interação peptídeo-hapteno-molécula de HLA da classe II, não foram suficientes para definir a formação de agrupamentos baseados nas proteínas carreadoras, hapteno e alelos.

Buscando compreender as limitações e potencialidades relacionadas à formulação estrutural, as análises realizadas sobre os peptídeos não conjugados aos haptenos levaram à conclusão que a subunidade B da toxina da Cólera seria a pior escolha para proteína carreadora visto que a maior parte dos resíduos de lisina não se encontra acessível ao solvente e os peptídeos com possibilidade de ligação aos haptenos apresentam baixa promiscuidade. Por outro lado, a proteína BSA se destaca por gerar peptídeos que se ligam a quatro alelos diferentes, fator pertinente à eficácia *in vivo*.

Aplicando o mesmo critério sobre as simulações de *docking*, segundo o mecanismo em que só o peptídeo ocupa o sítio de ligação da molécula HLA de classe II, a proteína BSA também se destaca como a melhor proteína carreadora por apresentar, em média, menor energia intermolecular na interação com todos os alelos. Considerando o possível mecanismo de apresentação da estrutura peptídeo-hapteno, a melhor combinação entre as analisadas, segundo o critério de afinidade, seria KLH-GND. Em termos da natureza e estabilidade das interações intermoleculares durante a apresentação do antígeno, os resultados indicaram que os haptenos predominantemente não prejudicam a apresentação, inclusive favorecendo ligações de van der Waals e ligações hidrogênio. O único alelo em que isso não se demonstrou verdade foi o DRA*0101/2-DRB1*0401.

Com o intuito de investigar a factibilidade do mecanismo de apresentação do peptídeo não conjugado, analisou-se o comportamento do sistema químico durante a dinâmica molecular. No geral, para os casos verificados, o complexo peptídeo-hapteno-HLA de Classe II foi estável ao longo do tempo, indicando que esse mecanismo de apresentação é de fato possível. As ligações hidrogênio formadas variaram consideravelmente em quantidade total e ocorrência nos frames da trajetória de um caso para o outro. Em termos energéticos, essa discrepância teria impacto direto na imunogenicidade da resposta imune. A mesma afirmação

não poderia ser feita em termos biológicos, uma vez que não se sabe com certeza a proporção desse tipo de apresentação.

Como um todo, o trabalho forneceu percepções sobre os caminhos a serem seguidos (utilização de BSA como proteína carreadora considerando mecanismo de apresentação de peptídeo não conjugado e KLH considerando mecanismo de apresentação conjugado) ou não (utilização da subunidade B da toxina da cólera como proteína carreadora) durante a elaboração de uma vacina para dependência de cocaína, além de construir caminhos para pesquisa ainda mais profunda sobre a relação de moléculas HLA de classe II com vacinas conjugadas. Destarte, é importante entender todas as nuances envolvidas no mecanismo da apresentação de vacinas conjugadas, mais especificamente naquelas voltadas para a dependência de cocaína, para que finalmente seja possível atingir a eficácia clínica desejada, e, posteriormente, sejam utilizadas como estratégia de redução de danos.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] BRAIN, P. F.; COWARD, G. A. A review of the history, actions, and legitimate uses of cocaine. **Journal of substance abuse**, v. 1, n. 4, p. 431–451, 1989.
- [2] BAUER, I. Travel medicine, coca and cocaine: Demystifying and rehabilitating *Erythroxylum* - A comprehensive review. **Tropical Diseases, Travel Medicine and Vaccines**, v. 5, n. 1, p. 1–14, 2019.
- [3] MOON, J. S.; COLEMAN, M. L. Coca: The Sun God's Anesthetic Leaves. **Anesthesiology**, v. 135, p. 284–291, 2021.
- [4] GRINSPOON, L.; BAKALAR, J. B. Coca and cocaine as medicines: An historical review. **Journal of Ethnopharmacology**, v. 3, n. 2–3, p. 149–159, 1981.
- [5] KARCH, S. B. Cocaine: History, use, abuse. **Journal of the Royal Society of Medicine**, v. 92, n. 8, p. 393–397, 1999.
- [6] UNODC. **World Drug Report**. [s.l.: s.n.]. v. 45
- [7] FERREIRA, P. E. M.; MARTINI, R. K. Cocaína: lendas, história e abuso. **Revista Brasileira de Psiquiatria**, v. 23, n. 2, p. 96–99, 2001.
- [8] ALLEN, D. F. et al. Epidemic Free-base Cocaine Abuse. **The Lancet**, p. 459–462, 1986.
- [9] LAPPAN, S. N.; BROWN, A. W.; HENDRICKS, P. S. Dropout rates of in-person psychosocial substance use disorder treatments: a systematic review and meta-analysis. **Addiction**, v. 115, n. 2, p. 201–217, 2020.
- [10] CHAN, B. et al. Pharmacotherapy for Cocaine Use Disorder—a Systematic Review and Meta-analysis. **Journal of General Internal Medicine**, v. 34, n. 12, p. 2858–2873, 2019.
- [11] ORSON, F. M. et al. The future potential for cocaine vaccines. **Expert Opinion on Biological Therapy**, v. 14, n. 9, p. 1271–1283, 2014.
- [12] BONESE, K. F. et al. Changes in heroin self-administration by a rhesus monkey after morphine immunisation. **Nature**, v. 252, p. 708–710, 1974.
- [13] BAGASRA, O.; FORMAN, L. J.; HOWEEDY, A. A potential vaccine for cocaine abuse prophylaxis. **Immunopharmacology**, v. 23, p. 173–179, 1992.
- [14] TRUONG, T. T.; KOSTEN, T. R. Current status of vaccines for substance use disorders: A brief review of human studies. **Journal of the Neurological Sciences**, v. 434, n. July 2021, p. 120098, 2021.
- [15] MARTELL, B. A. et al. Cocaine Vaccine for the Treatment of Cocaine Dependence in Methadone-Maintained Patients. **ARCH GEN PSYCHIATRY**, v. 66, n. 10, p. 1116–1123, 2009.
- [16] POLLARD, A. J.; BIJKER, E. M. A guide to vaccinology: from basic principles to new developments. **Nature Reviews Immunology**, v. 21, n. 2, p. 83–100, 2021.
- [17] DUFFY, E. B.; DRAKE, J. R.; HARTON, J. A. Evolving Insights for MHC Class II Antigen Processing and Presentation in Health and Disease. **Current Pharmacology Reports**, v. 3, n. 5, p. 213–220, 2017.
- [18] RIBEIRO, M. et al. Adulterants in crack cocaine in Brazil. **Trends in Psychiatry and Psychotherapy**, v. 41, n. 2, p. 186–190, 2019.
- [19] PLOWMAN, T. The identification of coca (*Erythroxylum* species): 1860–1910. **Botanical Journal of the Linnean Society**, v. 84, n. 4, p. 329–353, 1982.
- [20] GOMES DE CASTRO NETO, A. et al. Cocaine and Its Variations in Forms of Presentation and Addiction. **Psychopathology - An International and Interdisciplinary Perspective**, 2020.
- [21] CARPENTIER, C. et al. **Cocaine Insights 2. Cocaine – a spectrum of products**. Vienna: [s.n.].
- [22] HATSUMAKI, D. K.; FISCHMAN, M. W. Crack Cocaine and Cocaine Hydrochloride. Are the Differences

Myth or Reality? **JAMA**, v. 276, n. 19, p. 1580–1588, 1986.

[23] WILKINSON, P. et al. Intranasal and oral cocaine kinetics. **Clinical Pharmacology & Therapeutics**, v. 27, p. 386–394, 1980.

[24] FATTINGER, K. et al. Nasal mucosal versus gastrointestinal absorption of nasally administered cocaine. **European Journal of Clinical Pharmacology**, v. 56, n. 4, p. 305–310, 2000.

[25] PEREZ-REYES, M. et al. Free-base cocaine smoking. **Clin. Pharmacol. Ther.**, v. 32, p. 459–465, 1982.

[26] CONE, E. J. Pharmacokinetics and pharmacodynamics of nicotine. **Journal of Analytical Toxicology**, v. 19, p. 459–478, 1995.

[27] COE, M. A. et al. Bioavailability and pharmacokinetics of oral cocaine in humans. **Journal of Analytical Toxicology**, v. 42, n. 5, p. 285–292, 2018.

[28] CUNHA-OLIVEIRA, T. et al. Medical Toxicology of Drugs of Abuse. In: **Principles of Addiction**. [s.l.: s.n.]. v. 1p. 159–175.

[29] TRIFILIEFF, P.; MARTINEZ, D. **Cocaine: Mechanism and Effects in the Human Brain**. [s.l.] Elsevier, 2013.

[30] GOODMAN, J.; PACKARD, M. G. Memory systems and the addicted brain. **Frontiers in Psychiatry**, v. 7, n. FEB, p. 1–9, 2016.

[31] ASHOK, A. H. et al. Association of stimulant use with dopaminergic alterations in users of cocaine, amphetamine, or methamphetamine a systematic review and meta-analysis. **JAMA Psychiatry**, v. 74, n. 5, p. 511–519, 2017.

[32] DACKIS, C. A.; GOLD, M. S. New concepts in cocaine addiction: The dopamine depletion hypothesis. **Neuroscience and Biobehavioral Reviews**, v. 9, n. 3, p. 469–477, 1985.

[33] REBEC, G. V. Cocaine and Amphetamines. **eLS**, p. 1–8, 2012.

[34] KIM, S. T.; PARK, T. Acute and chronic effects of cocaine on cardiovascular health. **International Journal of Molecular Sciences**, v. 20, n. 3, p. 1–22, 2019.

[35] MENDELSON, J. H. et al. Cocaine tolerance: Behavioral, cardiovascular, and neuroendocrine function in men. **Neuropsychopharmacology**, v. 18, n. 4, p. 263–271, 1998.

[36] SICILIANO, C. A.; FORDAHL, S. C.; JONES, S. R. Cocaine self-administration produces long-lasting alterations in dopamine transporter responses to cocaine. **Journal of Neuroscience**, v. 36, n. 30, p. 7807–7816, 2016.

[37] CARR, C. C.; FERRARIO, C. R.; ROBINSON, T. E. Intermittent access cocaine self-administration produces psychomotor sensitization: effects of withdrawal, sex and cross-sensitization. **Psychopharmacology**, v. 237, n. 6, p. 1795–1812, 2020.

[38] BUZSÁKI, G.; DRAGUHN, A. Neuronal Oscillations in Cortical Networks. **Science**, v. 304, n. June, p. 1926–1929, 2004.

[39] LIU, C. L. et al. Cocaine-induced locomotor sensitization associates with slow oscillatory firing of neurons in the ventral tegmental area. **Scientific Reports**, v. 8, n. 1, p. 1–11, 2018.

[40] NESTLER, E. J. The neurobiology of cocaine addiction. **Science & practice perspectives / a publication of the National Institute on Drug Abuse, National Institutes of Health**, v. 3, n. 1, p. 4–10, 2005.

[41] SOLINAS, M. et al. Reversal of cocaine addiction by environmental enrichment. **Proceedings of the National Academy of Sciences of the United States of America**, v. 105, n. 44, p. 17145–17150, 2008.

[42] INABA, T.; STEWART, D. J.; KALOW, W. Metabolism of cocaine in man. **Clinical Pharmacology and Therapeutics**, v. 23, n. 5, p. 547–552, 1978.

[43] SCHINDLER, C. W.; GOLDBERG, S. R. Accelerating cocaine metabolism as an approach to the treatment of cocaine abuse and toxicity. **Future Medicinal Chemistry**, v. 4, n. 2, p. 163–175, 2012.

[44] KOVACIC, P. Role of oxidative metabolites of cocaine in toxicity and addiction: Oxidative stress and

electron transfer. **Medical Hypotheses**, v. 64, n. 2, p. 350–356, 2005.

[45] BEISER, T.; YAKA, R. The Role of Oxidative Stress in Cocaine Addiction. **Journal of Neurology & Neuromedicine**, v. 4, n. 1, p. 17–21, 2019.

[46] SUÁREZ-GARCÍA, A. et al. Duration of detection of cocaine and metabolites in hair after discontinuation of abuse. **Microchemical Journal**, v. 153, n. October 2019, p. 104335, 2020.

[47] SCHEIDWEILER, K. B. et al. Pharmacokinetics of cocaine and metabolites in human oral fluid and correlation with plasma concentrations after controlled administration. **Therapeutic Drug Monitoring**, v. 32, n. 5, p. 628–637, 2010.

[48] BARROSO, M.; GALLARDO, E.; QUEIROZ, J. A. Bioanalytical methods for the determination of cocaine and metabolites in human biological samples. **Bioanalysis**, v. 1, n. 5, p. 977–1000, 2009.

[49] VANDEVENNE, M.; VANDENBUSSCHE, H.; VERSTRAETE, A. Detection time of drugs of abuse in urine. **Acta Clinica Belgica**, v. 55, n. 6, p. 323–333, 2000.

[50] ORSON, F. M. et al. Substance abuse vaccines. **Annals of the New York Academy of Sciences**, v. 1141, p. 257–269, 2008.

[51] KOLBRICH, E. A. et al. Major and minor metabolites of cocaine in human plasma following controlled subcutaneous cocaine administration. **Journal of Analytical Toxicology**, v. 30, n. 8, p. 501–510, 2006.

[52] WANG, Q. et al. Contribution of the active metabolite, norcocaine, to cocaine's effects after intravenous and oral administration in rats: Pharmacodynamics. **Psychopharmacology**, v. 153, n. 3, p. 341–352, 2001.

[53] GEFEN, T. et al. The effect of haptens on protein-carrier immunogenicity. **Immunology**, v. 144, n. 1, p. 116–126, 2015.

[54] LEMUS, R.; KAROL, M. H. **Conjugation of Haptens**. [s.l.: s.n.]. v. 138

[55] ROCHE, P. A.; FURUTA, K. The ins and outs of MHC class II-mediated antigen processing and presentation. **Nature Reviews Immunology**, v. 15, n. 4, p. 203–216, 2015.

[56] ABBAS, A. K.; LICHTMAN, A. H. **Imunologia Básica: Funções e Distúrbios do Sistema Imunológico**. 3 ed. ed. Rio de Janeiro: Editora Elsevier Ltda, 2009.

[57] BLUM, J. S.; WEARSCH, P. A.; CRESSWELL, P. **Pathways of antigen processing**. [s.l.: s.n.]. v. 31

[58] RAPPUOLI, R. et al. On the mechanisms of conjugate vaccines. v. 116, n. 1, p. 2018–2020, 2019.

[59] AVCI, F. Y. et al. A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. **Nature Medicine**, v. 17, n. 12, p. 1602–1609, 2011.

[60] LAZARSKI, C. A. et al. The kinetic stability of MHC class II:Peptide complexes is a key parameter that dictates immunodominance. **Immunity**, v. 23, n. 1, p. 29–40, 2005.

[61] WAGNER, A.; SOUZA, S. DE; CRUVINEL, W. D. M. Immune System - Part II Basis of the immunological response mediated by T and B lymphocytes. **Definitions**, v. 55, n. 11, 2020.

[62] RAJEWSKY, K. **Clonal selection and learning in the antibody system** *Nature*, 1996.

[63] RUNDESKY, A. Y. et al. Sequence analysis of peptides bound to MHC class II molecules. **Nature**, v. 622–627, n. 17, p. 603–605, 1991.

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

[65] ROBINSON, J. et al. IPD-IMGT/HLA Database. **Nucleic Acids Research**, v. 48, n. D1, p. D948–D955, 2020.

[66] CUNHA-NETO, E. MHC-restricted antigen presentation and recognition: Constraints on gene, recombinant and peptide vaccines in humans. **Brazilian Journal of Medical and Biological Research**, v. 32, n. 2, p. 199–205, 1999.

- [67] MARSH, S. G. E. et al. Nomenclature for factors of the HLA system, 2010. **Tissue Antigens**, v. 75, n. 4, p. 291–455, 2010.
- [68] FERNANDES, A. P. M. et al. Como entender a associação entre o sistema HLA e as doenças auto-imunes endócrinas. **Arquivos Brasileiros de Endocrinologia & Metabologia**, v. 47, n. 5, p. 601–611, 2003.
- [69] KITCHEN, D. B. et al. Docking and scoring in virtual screening for drug discovery: Methods and applications. **Nature Reviews Drug Discovery**, v. 3, n. 11, p. 935–949, 2004.
- [70] PRIETO-MARTÍNEZ, F. D.; ARCINIEGA, M.; MEDINA-FRANCO, J. L. Acoplamiento Molecular: Avances Recientes y Retos. **TIP Revista Especializada en Ciencias Químico-Biológicas**, v. 21, n. May, p. 0–23, 2018.
- [71] FERREIRA, L. G. et al. **Molecular docking and structure-based drug design strategies**. [s.l.: s.n.]. v. 20
- [72] EVANGELISTA FALCON, W. et al. Ensemble Docking in Drug Discovery: How Many Protein Configurations from Molecular Dynamics Simulations are Needed to Reproduce Known Ligand Binding? **Journal of Physical Chemistry B**, v. 123, n. 25, p. 5189–5195, 2019.
- [73] TORRES, P. H. M. et al. Key topics in molecular docking for drug design. **International Journal of Molecular Sciences**, v. 20, n. 18, p. 1–29, 2019.
- [74] GUEDES, I. A.; MAGALHÃES, C. S. DE; DARDENNE, L. E. Receptor-ligand molecular docking. **Biophysical Reviews**, v. 6, n. 1, p. 75–87, 2014.
- [75] JONES, G. et al. Development and Validation of a Genetic Algorithm for Flexible Docking. **J. Mol. Biol.**, v. 267, p. 724–748, 1997.
- [76] ÖRDÖG, R.; GROLMUSZ, V. Evaluating genetic algorithms in protein-ligand docking. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 4983 LNBI, p. 402–413, 2008.
- [77] MORRIS, G. M. et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. **Journal of computational chemistry**, v. 19, n. 14, p. 1639–1662, 1998.
- [78] LI, J.; FU, A.; ZHANG, L. An Overview of Scoring Functions Used for Protein–Ligand Interactions in Molecular Docking. **Interdisciplinary Sciences – Computational Life Sciences**, v. 11, n. 2, p. 320–328, 2019.
- [79] SAPUNDZHI, F.; PRODANOVA, K.; LAZAROVA, M. **Survey of the scoring functions for protein-ligand docking**. AIP Conference Proceedings. **Anais...American Institute of Physics Inc.**, 13 Nov. 2019
- [80] PECINA, A. et al. SQM/COSMO Scoring Function: Reliable Quantum-Mechanical Tool for Sampling and Ranking in Structure-Based Drug Design. **ChemPlusChem**, 2020.
- [81] GUEDES, I. A.; PEREIRA, F. S. S.; DARDENNE, L. E. Empirical scoring functions for structure-based virtual screening: Applications, critical aspects, and challenges. **Frontiers in Pharmacology**, v. 9, n. SEP, p. 1–18, 2018.
- [82] PASON, L. P.; SOTRIFFER, C. A. Empirical Scoring Functions for Affinity Prediction of Protein-ligand Complexes. **Molecular Informatics**, v. 35, n. 11–12, p. 541–548, 2016.
- [83] YAN, Z.; WANG, J. Scoring Functions of Protein-Ligand Interactions. In: **Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery**. [s.l.] Medical Information Science Reference, 2016. p. 220–245.
- [84] MUEGGE, I. A knowledge-based scoring function for protein-ligand interactions: Probing the reference state. **Perspectives in Drug Discovery and Design**, v. 20, p. 99–114, 2000.
- [85] LEVER, J.; KRZYWINSKI, M.; ALTMAN, N. Principal component analysis. **Nature Methods**, v. 14, n. 7, p. 641–642, 2017.
- [86] BRO, R.; SMILDE, A. K. Principal component analysis. **Analytical Methods**, v. 6, n. 9, p. 2812–2831, 2014.
- [87] JOLLIFE, I. T.; CADIMA, J. Principal component analysis: A review and recent developments. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v.

374, n. 2065, 2016.

[88] CORDELLA, C. B. Y. PCA : The Basic Building Block of Chemometrics. **Analytical chemistry**, n. December, p. 1–46, 2012.

[89] MISHRA, P. et al. New data preprocessing trends based on ensemble of multiple preprocessing techniques. **TrAC - Trends in Analytical Chemistry**, v. 132, p. 116045, 2020.

[90] G. DUMANCAS, G. et al. Chemometrics: From Data Preprocessing to Fog Computing. **International Journal of Fog Computing**, v. 2, n. 1, p. 1–42, 2019.

[91] D. TOBIAS, R. An Introduction to Partial Least Squares Regression. **SUGI Proceedings**, p. 1–8, 1995.

[92] ABDI, H. **Partial Least Squares (PLS) Regression**. [s.l.: s.n.].

[93] GAO, S. Interpreting the wave function - what are electrons? and how do they move? v. 22, p. 13–23, 2013.

[94] LUIS LEVADA, C. et al. Review of the Schrödinger Wave Equation. **IOSR Journal of Applied Chemistry (IOSR-JAC)**, v. 11, n. 4, p. 1–07, 2018.

[95] PAULA, J. DE; ATKINS, P. **Physical Chemistry**. 8th. ed. New York: W. H. Freeman and Company, 2006. v. 272

[96] CRAMER, C. J. **Essentials of Computational Chemistry Theories and Models Second Edition**. [s.l.: s.n.].

[97] BORN, M.; OPPENHEIMER, R. Zur Quantentheorie der Molekeln. **Annalen der Physik**, v. 389, n. 457–484, 1927.

[98] KLAR, H. The Born-Oppenheimer Approximation Revisited. **Journal of Applied Mathematics and Physics**, v. 08, n. 08, p. 1507–1514, 2020.

[99] LEVINE, I. N. **Quantum Chemistry**. 7th. ed. New York: Pearson Education, 2014.

[100] ECHENIQUE, P.; ALONSO, J. L. A mathematical and computational review of Hartree-Fock SCF methods in quantum chemistry. **Molecular Physics**, v. 105, n. 23–24, p. 3057–3098, 2007.

[101] SZABO, A.; OSTLUND, N. S. **Modern Quantum Chemistry**. New York: Dover Publications, 1996.

[102] MUELLER, M. **Fundamentals of Quantum Chemistry**. New York: Kluwer Academic Publishers, 2001.

[103] MARTINEZ, J. P. Beyond computational difficulties: Survey of the two decades from the elaboration to the extensive application of the Hartree-Fock method. **Studies in History and Philosophy of Science Part B - Studies in History and Philosophy of Modern Physics**, v. 60, p. 123–135, 2017.

[104] BLINDER, S. M. Basic Concepts of Self-Consistent-Field Theory. **American Journal of Physics**, v. 33, n. 6, p. 431–443, 1965.

[105] YAMAMOTO, S.; NAGASHIMA, U. Four-index integral transformation exploiting symmetry. **Computer Physics Communications**, v. 166, n. 1, p. 58–65, 2005.

[106] PHAM, B. Q.; GORDON, M. S. Compressing the Four-Index Two-Electron Repulsion Integral Matrix using the Resolution-of-the-Identity Approximation Combined with the Rank Factorization Approximation. **Journal of Chemical Theory and Computation**, v. 15, n. 4, p. 2254–2264, 2019.

[107] JOHANSSON, M. P.; KAILA, V. R. I.; SUNDHOLM, D. Ab Initio, Density Functional Theory, and Semi-Empirical Calculations. **Methods in Molecular Biology**, v. 924, p. 3–27, 2013.

[108] YOUNG, D. C. **COMPUTATIONAL CHEMISTRY: A practical guide for applying techniques to real-world problems**. New York: [s.n.].

[109] CAPELLE, K. A Bird 's-Eye View of Density-Functional Theory. v. 36, n. 4, p. 1318–1343, 2006.

[110] KOHN, W.; SHAM, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. **Physical Review**, v. 140, p. 289–295, 1965.

[111] ZIESCHE, P.; KURTH, S.; PERDEW, J. P. Density functionals from LDA to GGA. **Computational Materials Science**, v. 11, n. 2, p. 122–127, 1998.

- [112] MIETZSCH, N. The validity of the local density approximation for smooth short range interaction potentials. **Journal of Mathematical Physics**, v. 61, n. 11, 2020.
- [113] MARDIROSSIAN, N.; HEAD-GORDON, M. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. **Molecular Physics**, v. 115, n. 19, p. 2315–2372, 2017.
- [114] ARBUZNIKOV, A. V. HYBRID EXCHANGE CORRELATION FUNCTIONALS AND POTENTIALS: CONCEPT ELABORATION. **Journal of Structural Chemistry**, v. 48, p. S1–S31, 2007.
- [115] LEE, C.; TANG, W.; PARR, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. **Physical Review B**, v. 37, n. 2, p. 785–789, 1998.
- [116] BECKE, A. D. Density-functional thermochemistry. III. The role of exact exchange. **The Journal of Chemical Physics**, v. 98, n. 7, p. 5648–5652, 1993.
- [117] TORRES, E.; DILABIO, G. A. A (nearly) universally applicable method for modeling noncovalent interactions using B3LYP. **Journal of Physical Chemistry Letters**, v. 3, n. 13, p. 1738–1744, 2012.
- [118] LU, L. Can B3LYP be improved by optimization of the proportions of exchange and correlation functionals? **International Journal of Quantum Chemistry**, v. 115, n. 8, p. 502–509, 2015.
- [119] EOM, K. Computer Simulation of Protein Materials at Multiple Length Scales: From Single Proteins to Protein Assemblies. **Multiscale Science and Engineering**, v. 1, n. 1, p. 1–25, 2019.
- [120] NAMBA, A. M.; SILVA, V. B. DA; SILVA, C. H. T. P. DA. Dinâmica molecular: Teoria e aplicações em planejamento de fármacos. **Ecletica Química**, v. 33, n. 4, p. 13–24, 2008.
- [121] GONZÁLEZ, M. A. Force fields and molecular dynamics simulations. **Collection de la Société Française de la Neutronique**, v. 12, p. 169–200, 2011.
- [122] PETROV, D. et al. A Systematic Framework for Molecular Dynamics Simulations of Protein Post-Translational Modifications. **PLoS Computational Biology**, v. 9, n. 7, 2013.
- [123] OOSTENBRINK, C. et al. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1656–1676, 2004.
- [124] TIRONI, I. G. et al. A generalized reaction field method for molecular dynamics simulations. **The Journal of Chemical Physics**, v. 102, n. 13, p. 5451–5459, 1995.
- [125] BAUER, P.; HESS, B.; LINDAHL, E. GROMACS Documentation. **GROMACS Reference Manual**, p. 1–664, 2022.
- [126] FACCIOLI, P. Molecular dynamics at low time resolution. **Journal of Chemical Physics**, v. 133, n. 16, p. 1–12, 2010.
- [127] BERENDSEN, H. J. C.; GUNSTEREN, W. F. VAN. MOLECULAR DYNAMICS WITH CONSTRAINTS. **The Physics of Superionic Conductors and Electrode Materials**, p. 221–240, 1983.
- [128] SMITH, Z. et al. Discovering Protein Conformational Flexibility through Artificial-Intelligence-Aided Molecular Dynamics. **Journal of Physical Chemistry B**, v. 124, n. 38, p. 8221–8229, 2020.
- [129] HEO, L.; FEIG, M. What makes it difficult to refine protein models further via molecular dynamics simulations? **Proteins: Structure, Function and Bioinformatics**, v. 86, n. S1, p. 177–188, 2017.
- [130] CLARK, A. J. et al. Prediction of Protein-Ligand Binding Poses via a Combination of Induced Fit Docking and Metadynamics Simulations. **Journal of Chemical Theory and Computation**, v. 12, n. 6, p. 2990–2998, 2016.
- [131] HOLLINGSWORTH, S. A.; DROR, R. O. Molecular Dynamics Simulation for All. **Neuron**, v. 99, n. 6, p. 1129–1143, 2018.
- [132] SINGH, A. et al. Application of molecular dynamic simulation to study food proteins: A review. **Critical Reviews in Food Science and Nutrition**, v. 58, n. 16, p. 2779–2789, 2018.
- [133] SMITH, A.; DONG, X.; RAGHAVAN, V. An Overview of Molecular Dynamics Simulation for Food

Products and Processes. **Processes**, v. 10, n. 1, 2022.

[134] KANTAK, K. M. et al. Evaluation of anti-cocaine antibodies and a cocaine vaccine in a rat self-administration model. **Psychopharmacology**, v. 148, n. 3, p. 251–262, 2000.

[135] ETTINGER, R. H.; ETTINGER, W. F.; HARLESS, W. E. Active immunization with cocaine-protein conjugate attenuates cocaine effects. **Pharmacology Biochemistry and Behavior**, v. 58, n. 1, p. 215–220, 1997.

[136] CARRERA, M. R. A. et al. A second-generation vaccine protects against the psychoactive effects of cocaine. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 4, p. 1988–1992, 2001.

[137] FOX, B. S. et al. Efficacy of a therapeutic cocaine vaccine in rodent models. **Nature Medicine**, v. 2, n. 10, p. 1129–1132, 1996.

[138] CARRERA, M. R. A. et al. Suppression of psychoactive effects of cocaine by active immunization. **Nature**, v. 378, n. 6558, p. 727–730, 1995.

[139] THERMO SCIENTIFIC. Crosslinkers Technical Handbook. **Thermo Scientific**, p. 1–53, 2012.

[140] CAI, X. et al. Probing active cocaine vaccination performance through catalytic and noncatalytic hapten design. **Journal of Medicinal Chemistry**, v. 56, n. 9, p. 3701–3709, 2013.

[141] KRISSEL, E.; HENRICK, K. Inference of macromolecular assemblies from crystalline state. **Journal of Molecular Biology**, v. 372, n. 3, p. 774–797, 2007.

[142] SANCHEZ-TRINCADO, J. L.; GOMEZ-PEROSANZ, M.; RECHE, P. A. Fundamentals and Methods for T- and B-Cell Epitope Prediction. **Journal of Immunology Research**, v. 2017, 2017.

[143] MANOURY, B. Proteases: Essential actors in processing antigens and intracellular toll-like receptors. **Frontiers in Immunology**, v. 4, n. SEP, p. 1–5, 2013.

[144] SONG, J. et al. PROSPER: An Integrated Feature-Based Tool for Predicting Protease Substrate Cleavage Sites. **PLoS ONE**, v. 7, n. 11, 2012.

[145] WANG, P. et al. Peptide binding predictions for HLA DR, DP and DQ molecules. **BMC Bioinformatics**, v. 11, n. 1, p. 568, 2010.

[146] NIELSEN, M.; LUND, O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. **BMC Bioinformatics**, v. 10, p. 296, 2009.

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

[148] SIDNEY, J. et al. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. **Immunome Research**, v. 4, n. 1, p. 1–14, 2008.

[149] STURNIOLO, T. et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. **Nature Biotechnology**, v. 17, n. 6, p. 555–561, 1999.

[150] REYNISSON, B. et al. NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. **Nucleic Acids Research**, v. 48, n. W1, p. W449–W454, 2021.

[151] MATEOS-APARICIO, G. Partial least squares (PLS) methods: Origins, evolution, and application to social sciences. **Communications in Statistics - Theory and Methods**, v. 40, n. 13, p. 2305–2317, 2011.

[152] BAIROCH, A.; APWEILER, R. The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. **Nucleic Acids Research**, v. 27, n. 1, p. 49–54, 1999.

[153] SOUTHWOOD, S. et al. Several common HLA-DR types share largely overlapping peptide binding repertoires. **Journal of immunology (Baltimore, Md. : 1950)**, v. 160, n. 7, p. 3363–73, 1998.

[154] GASTEIGER, E. et al. Protein Identification and Analysis Tools on the ExPASy Server. **The Proteomics Protocols Handbook**, p. 571–608, 2005.

- [155] BJELLQVIST, B. et al. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. **Electrophoresis**, v. 14, n. 1, p. 1023–1031, 1993.
- [156] VARSHAVSKY, A. The N-end rule pathway of protein degradation. **Genes to Cells**, v. 2, n. 1, p. 13–28, 1997.
- [157] GURUPRASAD, K.; REDDY, B. V. B.; PANDIT, M. W. Correlation between stability of a protein and its dipeptide composition: A novel approach for predicting in vivo stability of a protein from its primary sequence. **Protein Engineering, Design and Selection**, v. 4, n. 2, p. 155–161, 1990.
- [158] IKAI, A. Thermostability and aliphatic index of globular proteins. **Journal of Biochemistry**, v. 88, n. 6, p. 1895–1898, 1980.
- [159] KYTE, J.; DOOLITTLE, R. F. A simple method for displaying the hydropathic character of a protein. **Journal of Molecular Biology**, v. 157, n. 1, p. 105–132, 1982.
- [160] MAGNAN, C. N. et al. High-throughput prediction of protein antigenicity using protein microarray data. **Bioinformatics**, v. 26, n. 23, p. 2936–2943, 2010.
- [161] LI, F. et al. Procleave: Predicting Protease-specific Substrate Cleavage Sites by Combining Sequence and Structural Information. **Genomics, Proteomics and Bioinformatics**, v. 18, n. 1, p. 52–64, 2020.
- [162] SADEGH-NASSERI, S.; KIM, A. R. MHC class II auto-antigen presentation is unconventional. **Frontiers in Immunology**, v. 6, n. JUL, p. 1–5, 2015.
- [163] SHEN, Y. et al. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. **Journal of Chemical Theory and Computation**, v. 10, n. 10, p. 4745–4758, 2014.
- [164] THÉVENET, P. et al. PEP-FOLD: An updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. **Nucleic Acids Research**, v. 40, n. W1, p. 288–293, 2012.
- [165] MAUPETIT, J.; TUFFERY, P.; DERREUMAUX, P. A coarse-grained protein force field for folding and structure prediction. **Wiley InterScience**, v. 69, n. 2, p. 394–408, 2007.
- [166] NEESE, F. The ORCA program system. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, v. 2, n. 1, p. 73–78, 2012.
- [167] HANWELL, M. D. et al. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. **Journal of Cheminformatics**, v. 4, n. 8, p. 1–17, 2012.
- [168] WANG, J. et al. Development and Testing of a General Amber Force Field. **Journal of Computational Chemistry**, v. 25, n. 9, p. 1157–1174, 2004.
- [169] LILJEDAHL, M. et al. Altered antigen presentation in mice lacking H2-O. **Immunity**, v. 8, n. 2, p. 233–243, 1998.
- [170] GHOSH, P. et al. The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. **Nature**, v. 378, n. 7, p. 457–462, 1995.
- [171] SMITH, K. J. et al. Crystal structure of HLA-DR2 (DRA*0101, DRB1*1501) Complexed with a peptide from human myelin basic protein. **Journal of Experimental Medicine**, v. 188, n. 8, p. 1511–1520, 1998.
- [172] STERN, L. J. et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. **Nature**, v. 368, n. 6468, p. 215–221, 1994.
- [173] LI, Y. et al. Structural basis for the binding of an immunodominant peptide from myelin basic protein in different registers by two HLA-DR2 proteins. **Journal of Molecular Biology**, v. 304, n. 2, p. 177–188, 2000.
- [174] HENNECKE, J.; WILEY, D. C. Structure of a complex of the human α/β T cell receptor (TCR) HA1.7, Influenza hemagglutinin peptide, and major histocompatibility complex class II molecule, HLA-DR4 (DRA*0101 and DRB1*0401): Insight into TCR cross-restriction and alloreactivity. **Journal of Experimental Medicine**, v. 195, n. 5, p. 571–581, 2002.
- [175] KIM, C.-Y. et al. Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. **PNAS**, v. 101, n. 12, p. 4175–4179, 2004.
- [176] SIEBOLD, C. et al. Crystal structure of HLA-DQ0602 that protects against type 1 diabetes and confers

strong susceptibility to narcolepsy. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 7, p. 1999–2004, 2004.

[177] HENDERSON, K. N. et al. A Structural and Immunological Basis for the Role of Human Leukocyte Antigen DQ8 in Celiac Disease. **Immunity**, v. 27, n. 1, p. 23–34, 2007.

[178] PARRY, C. S.; GORSKI, J.; STERN, L. J. Crystallographic Structure of the Human Leukocyte Antigen DRA, DRB3*0101: Models of a Directional Alloimmune Response and Autoimmunity. **Journal of Molecular Biology**, v. 371, n. 2, p. 435–446, 2007.

[179] DAI, S. et al. Crystal structure of HLA-DP2 and implications for chronic beryllium disease. **Proceedings of the National Academy of Sciences of the United States of America**, v. 107, n. 16, p. 7425–7430, 2010.

[180] HSU, K.-C. et al. iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. **BioMed Central**, v. 12, 2011.

[181] MORRIS, G. M. et al. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. **Journal of Computational Chemistry**, v. 30, n. 16, p. 2785–2791, 2009.

[182] TROTT, O.; OLSON, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**, v. 31, n. 2, p. NA-NA, 2009.

[183] SCHMID, N. et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. **European Biophysics Journal**, v. 40, n. 7, p. 843–856, 2011.

[184] BERENDSEN, H. J. C. et al. INTERACTION MODELS FOR WATER IN RELATION TO PROTEIN HYDRATION. In: **Intermolecular Forces**. [s.l.: s.n.]. v. 14p. 331–342.

[185] DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. **The Journal of Chemical Physics**, v. 98, n. 12, p. 10089–10092, 1993.

[186] JANEŽIČ, D.; MERZEL, F. An Efficient Symplectic Integration Algorithm for Molecular Dynamics Simulations. **Journal of Chemical Information and Computer Sciences**, v. 35, n. 2, p. 321–326, 1995.

[187] BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. **Journal of Chemical Physics**, v. 126, n. 1, 2007.

[188] PARRINELLO, M.; RAHMAN, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. **Journal of Chemical Physics**, v. 45, n. 14, 1980.

[189] HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: Visual Molecular Dynamics. **J. Molec. Graphics**, v. 14, p. 33–38, 1996.

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

[191] ZHAO, R.; JIANG, M. Bovine serum albumin as an immunogenic carrier facilitating the development of hapten-specific monoclonal antibodies. 2020.

[192] PEÑA, I.; DOMÍNGUEZ, J. M. Thermally denatured BSA, a surrogate additive to replace BSA in buffers for high-throughput screening. **Journal of Biomolecular Screening**, v. 15, n. 10, p. 1281–1286, 2010.

[193] SRIRAMULU, D. K.; WU, S.; LEE, S. G. Effect of ligand torsion number on the AutoDock mediated prediction of protein-ligand binding affinity. **Journal of Industrial and Engineering Chemistry**, v. 83, p. 359–365, 2020.

[194] HUEY, R. et al. A Semiempirical Free Energy Force Field with Charge-Based Desolvation. **Journal of Computational Chemistry**, v. 28, n. 6, p. 1145–1152, 2007.

[195] DEMHARTER, S. et al. HLA-DM Stabilizes the Empty MHCII Binding Groove: A Model Using Customized Natural Move Monte Carlo. **Journal of Chemical Information and Modeling**, v. 59, n. 6, p. 2894–2899, 2019.

[196] CHATURVEDI, P. et al. The functional role of class II-associated invariant chain peptide (CLIP) in its

ability to variably modulate immune responses. **International Immunology**, v. 12, n. 6, p. 757–765, 2000.

[197] CHAPPEY, O. N.; SANDOUK, P.; SCHERRMANN, J. M. G. **Monoclonal Antibodies in Hapten Immunoassays**.

[198] DERKS, E. P. P. et al. Parallel processing of chemical information in a local area network—II. A parallel cross-validation procedure for artificial neural networks. **Computers & Chemistry**, v. 20, n. 4, p. 439–448, Aug. 1996.

[199] SCHÄFER, A.; HORN, H.; AHLRICHS, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. **The Journal of Chemical Physics**, v. 97, n. 4, p. 2571–2577, Aug. 1992.

[200] SCHÄFER, A.; HUBER, C.; AHLRICHS, R. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. **The Journal of Chemical Physics**, v. 100, n. 8, p. 5829–5835, Apr. 1994.

[201] EICKHORN, K. et al. Auxiliary basis sets to approximate Coulomb potentials. **Chemical Physics Letters**, v. 240, n. 4, p. 283–290, Jun. 1995.

[202] WEIGEND, F. Accurate Coulomb-fitting basis sets for H to Rn. **Physical Chemistry Chemical Physics**, v. 8, n. 9, p. 1057, 2006.

[203] GULDE, R.; POLLAK, P.; WEIGEND, F. Error-Balanced Segmented Contracted Basis Sets of Double- ζ to Quadruple- ζ Valence Quality for the Lanthanides. **Journal of Chemical Theory and Computation**, v. 8, n. 11, p. 4062–4068, Nov. 2012.

[204] STOYCHEV, G. L.; AUER, A. A.; NEESE, F. Automatic Generation of Auxiliary Basis Sets. **Journal of Chemical Theory and Computation**, v. 13, n. 2, p. 554–562, Feb. 2017.

[205] GRIMME, S.; EHRLICH, S.; GOERIGK, L. Effect of the damping function in dispersion corrected density functional theory. **Journal of Computational Chemistry**, 2011.

[206] GRIMME, S. et al. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. **The Journal of Chemical Physics**, v. 132, n. 15, p. 154104, Apr. 2010.

[207] TSUZUKI, S.; UCHIMARU, T. Accuracy of intermolecular interaction energies, particularly those of hetero-atom containing molecules obtained by DFT calculations with Grimme's D2, D3 and D3BJ dispersion corrections. **Physical Chemistry Chemical Physics**, v. 22, n. 39, p. 22508–22519, Oct. 2020.

[208] BJORNSSON, R.; KRAEMER, T.; SPARTA, M.; SCHAPIRO, I.; BERILONE, J.; ROSE, B. **ORCA Input Library**, [s.d.].