



UMA BREVE INTRODUÇÃO À SIMULAÇÃO COMPUTACIONAL APLICADA À QUÍMICA

ALUÍSIO MARQUES DA FONSECA
EMMANUEL SILVA MARINHO
JOSÉ CLEITON SOUSA DOS SANTOS
(Organizadores)

REITOR

Hidelbrando dos Santos Soares

VICE-REITOR

Dárcio Ítalo Alves Teixeira

EDITORA DA UECE

Cleudene de Oliveira Aragão

CONSELHO EDITORIAL

Antônio Luciano Pontes

Eduardo Diatahy Bezerra de Menezes

Emanuel Ângelo da Rocha Fragoso

Francisco Horácio da Silva Frota

Francisco Josênio Camelo Parente

Gisafran Nazareno Mota Jucá

José Ferreira Nunes

Liduina Farias Almeida da Costa

Lucili Grangeiro Cortez

Luiz Cruz Lima

Manfredo Ramos

Marcelo Gurgel Carlos da Silva

Marcony Silva Cunha

Maria do Socorro Ferreira Osterne

Maria Salete Bessa Jorge

Silvia Maria Nóbrega-Therrien



UMA BREVE INTRODUÇÃO À SIMULAÇÃO COMPUTACIONAL APLICADA À QUÍMICA

**ALUÍSIO MARQUES DA FONSECA
EMMANUEL SILVA MARINHO
JOSÉ CLEITON SOUSA DOS SANTOS**

(Organizadores)

1ª Edição
Fortaleza - CE
2023



Uma breve introdução à simulação computacional aplicada à química

© 2023 Copyright by Aluisio Marques da Fonseca, Emmanuel Silva Marinho
e José Cleiton Sousa dos Santos

O conteúdo deste livro, bem como os dados usados e sua fidedignidade, são de responsabilidade exclusiva dos autores. O download e o compartilhamento da obra são autorizados desde que sejam atribuídos créditos aos autores. Além disso, é vedada a alteração de qualquer forma e/ou utilizá-la para fins comerciais.

COORDENAÇÃO EDITORIAL

Cleudene de Oliveira Aragão

PROJETO GRÁFICO E CAPA

Narcelio de Sousa Lopes

REVISÃO VERNACULAR E NORMALIZAÇÃO

O autores

Catálogo da publicação na Fonte

Bibliotecária – Meirilane Santos de Moraes Bastos

B846 Uma breve introdução à simulação computacional aplicada à química [Recurso eletrônico] / Aluisio Marques da Fonseca, Emmanuel Silva Marinho, José Cleiton Sousa dos Santos (Orgs.) – Fortaleza: EdUECE, 2023.

196 p. il.
ISBN: 978-85-7826-874-9

1. Química computacional. 2. Química teórica.
I. Aluisio Marques da Fonseca. II. Emmanuel Silva Marinho. III. José Cleiton Sousa dos Santos.

CDD 545

Todos os direitos reservados

Editora da Universidade Estadual do Ceará – EdUECE

Av. Dr. Silas Munguba, 1700 – Campus do Itaperi – Reitoria – Fortaleza – Ceará

CEP: 60714-903 – Tel: (085) 3101-9893

www.uece.br/eduece – E-mail: eduece@uece.br

Editora filiada à



Dedicamos a todos aqueles que de alguma forma contribuíram direta e indiretamente na construção desse material. Também aos verdadeiros heróis que por séculos passaram a maior parte de sua vida criando possibilidades para que possamos assimilar os conhecimentos básicos aqui inseridos.

Sumário

Prólogo • 7

Capítulo 1 – Métodos Quânticos: introdução para a modelagem molecular • 12

Capítulo 2 – Conceitos básicos sobre *docking* molecular • 36

Capítulo 3 – Estudo de proteínas e avaliação da qualidade pelo gráfico Ramachandran • 60

Capítulo 4 – Propriedades de Adme-T, uma abordagem introdutória • 73

Capítulo 5 – Predição de descritores *in vitro* de adme: modelo de apoio às decisões empíricas. • 94

Capítulo 6 – O uso do software arguslab na modelagem molecular: uma estratégia para o ensino e pesquisa em química por meio remoto • 107

Capítulo 7 – O essencial sobre dinâmica molecular • 126

Capítulo 8 – Um estudo sobre termodinâmica no *docking* molecular • 160

PRÓLOGO

A Química é a área da ciência que se dedica ao estudo da composição, da estrutura, das propriedades e da transformação da matéria, possuindo um caráter experimental. Durante muito tempo, o seu desenvolvimento ocorreu por meio da análise de dados obtidos em experimentos realizados nas bancadas de laboratórios. Essa área da ciência visa promover avanços na tecnologia, descobrir e desenvolver novos medicamentos, compreender de forma minuciosa os processos biológicos, atuar nas indústrias alimentícias e de cosméticos e, ainda, ajudar na compreensão do surgimento da vida, haja vista que a Bioquímica, assim como a Biologia Molecular, depende fundamentalmente da Química. Portanto, compreender atômica e microscopicamente os processos químicos torna-se importante.

Ao longo dos anos, um caráter essencialmente teórico começou a surgir na química. Foi por meio dos sonhos que alguns dos primeiros passos do que viria a ser a Química Teórica começaram a se desenvolver. Na busca pela organização dos elementos químicos, muitos cientistas tentaram encontrar um padrão envolvendo a periodicidade de suas propriedades. Foi em um sonho que o químico russo Dmitri Mendeleev, após muito trabalho, teve a “revelação” de organizar os elementos químicos tendo como base as suas massas atômicas, criando, assim, a tabela periódica e prevendo a descoberta de novos ele-

mentos. Outro exemplo é August Kekulé, outro químico que teve um sonho “profético”, após muito trabalho. Kekulé discute a tetravalência de átomos de carbono para encontrar a estrutura do benzeno e, após um sonho com um Uroboros, uma serpente mordendo a própria cauda, Kekulé surge com ideia da estrutura hexagonal do benzeno, contribuindo para a construção da Teoria Estrutural das Moléculas.

Nos anos seguintes, com o desenvolvimento da Físico-Química, grandes descobertas foram feitas, como os estudos termodinâmicos em equilíbrio químico de J. Willard Gibbs, a proposição das leis da cinética química por Van't Hoff e a introdução da Mecânica Quântica no conceito de valência por Linus Pauling. Ao prever o avanço que a Química Teórica teria no futuro, Svante Arrhenius declarou: “O lado teórico da físico-química é e provavelmente irá permanecer o dominante [...] e por esse motivo a físico-química pode ser considerada uma excelente escola de raciocínio exato para todos os estudantes das ciências naturais.”

Foi com o surgimento dos computadores e com o consequente desenvolvimento de seu poder e da sua capacidade computacional que a Química Teórica ganhou um poderoso aliado: a Química Computacional. Os muitos métodos computacionais de simulações e cálculos teóricos permitiram a criação da terceira área da ciência química: a experimental, a teórica e agora a computacional.

A partir do que foi citado anteriormente, surge a Química Teórica-Computacional (QTC), um ramo da Físico-Química que constrói modelos matemáticos baseados em Física Clássica e Física Moderna para simular reações químicas e ambientes

químicos potencialmente relevantes em diversas pesquisas. Além disso, a QTC utiliza-se de métodos computacionais para realizar simulações e cálculos teóricos, haja vista que soluções aproximadas de equações (como nos métodos quânticos e semiempíricos) são muitas vezes necessárias e integrações numéricas (como nos métodos baseados em mecânica molecular) são comuns, processos que seriam inviáveis de se calcular manualmente.

Existem algumas subáreas da QTC, por exemplo: métodos *Ab initio* (Hartree-Fock), métodos semiempíricos, métodos clássicos (Mecânica Molecular, Dinâmica Molecular e Monte Carlo) e Teoria do Funcional da Densidade (DFT). Cada método tem suas especificações e são mais apropriados para utilização a depender do sistema analisado, bem como do tamanho deste, de quais propriedades se deseja obter, assim como do tempo e da capacidade computacional disponível.

A diversidade de sistemas que podem ser estudados pela QTC faz com que essa área de estudo seja uma ferramenta poderosa para a elucidação de fenômenos físico-químicos, como: sistemas em estado sólido, líquido ou gasoso; sistemas biológicos (lipídios, carboidratos, proteínas, ácidos nucleicos); sistemas em condições extremas de temperatura e pressão; sistemas em ambientes químicos distintos de solvatação; simulação de reações que teriam um alto custo financeiro, caso fosse feito experimentalmente; entre outros.

Em relação às aplicações, a QTC surge como uma alternativa para se estudar e propor mecanismos de reações em diversos meios distintos, detectar possíveis estados de transição e intermediários no decorrer das reações, possibilitar a construção

teórica de gráficos espectroscópicos, permitir cálculos de informações termoquímicas e cinéticas, analisar cargas parciais em ambientes atômicos distintos, determinar o grau de interação entre espécies diferentes, bem como do número médio de espécies que coordenam determinada espécie de referência, além de permitir uma visualização da amostragem espacial das espécies circundantes.

Este livro é uma introdução às diferentes áreas e métodos da Química Teórica e Computacional e contém muitas referências importantes para o aprofundamento do tema. Desse modo, o livro permite ao leitor no fim de sua leitura avançar independentemente para áreas mais específicas da Química. Boa leitura!

Norberto de Kássio Vieira Monteiro

MÉTODOS QUÂNTICOS: INTRODUÇÃO PARA A MODELAGEM MOLECULAR

Leonardo Paes da Silva

Universidade Federal do Ceará

Francisco Wagner de Queiroz Almeida Neto

Universidade Federal do Ceará

Emmanuel Silva Marinho

Universidade Estadual do Ceara, Limoeiro do Norte-CE

Pedro de Lima-Neto

Universidade Federal do Ceará

Resumo:

A Modelagem Molecular tem sido uma ferramenta importante no desenvolvimento de pesquisas na área de química. A Mecânica Quântica vem contribuindo com modelos e descrições do comportamento microscópico de partículas e átomos a partir da resolução da equação de Schrödinger, sendo fundamental em caráter de pesquisa. O presente trabalho tem por objetivo enunciar alguns conceitos dos métodos quânticos como *Ab-initio*, semiempírico e DFT a fim de servir como base para o entendimento de suas aplicações. A evolução dos métodos se deu a partir do método de Hatree-Fock que tentou solucionar os cálculos por aproximação, sendo conhecido como *Ab-initio*. Este método descreve com precisão propriedades em pequenos átomos, mas com custo computacional elevado. A fim de otimizar os cálculos, os métodos Semiempíricos foram desenvolvidos com base em equações parametrizadas com objetivo de ajustar os dados aos valores empíricos, diminuindo o custo computacional para sistemas polieletrônicos e descrevendo bem algumas propriedades químicas. Um dos métodos

mais aplicados nos últimos anos, a Teoria do Funcional da Densidade (DFT), utiliza o funcional da densidade eletrônica dos átomos e moléculas para descrever várias propriedades químicas, sendo uma alternativa precisa e de baixo custo nas operações químico-quânticas trabalhando em sistemas pequenos ou macromoleculares. Dentre as principais aplicações dos métodos quânticos podemos resumir aos procedimentos de energia de ponto único, otimização geométrica, orbitais moleculares, espectros, frequências etc. Por fim, entender os métodos quânticos e suas aplicações ajudam a entender moléculas simples e complexas de modo a permitir previsibilidade de propriedades importantes das moléculas estudadas visando alcançar o desenvolvimento científico e tecnológico nos diversos ramos do conhecimento.

Palavras-chave: *Ab-initio*, DFT, Semiempíricos. Modelagem Molecular, Química Computacional.

1. Introdução

No ano de 1900, o físico alemão Max Planck (1858-1947) publicou suas primeiras hipóteses de que a energia radiante era emitida e absorvida em pequenas quantidades de energia, que ele denominou *quantum*, marcando aí, o nascimento da teoria quântica. Até os dias atuais, o entendimento dos conceitos ainda é abstrato, pois deriva de uma série de postulados e aproximações que atuam formalmente descrevendo de maneira lógica e matemática fenômenos físicos e químicos observáveis¹.

As equações da mecânica clássica (Newtoniana) descrevem muito bem objetos macroscópicos, mas não se aplica adequadamente às partículas muito pequenas como os núcleos atômicos e os elétrons nas camadas mais externas, sendo necessário novas descrições que só a mecânica quântica justifica com certo rigor e precisão.

A Modelagem Molecular tem sido um dos instrumentos mais importantes no desenvolvimento de pesquisas na área de química e correlatas, aplicando ferramentas computacionais. A implementação de modelos matemáticos e físicos e o auxílio de *softwares* de computadores trouxeram grandes contribuições na descrição de propriedades eletrônicas, atômicas e moleculares em sistema de interesse, com obtenção de dados aproximados aos valores experimentais. Os avanços tecnológicos de software e hardware ocorrido nas últimas décadas, somado ao desenvolvimento e/ou aprimoramento dos métodos computacionais, geraram as ferramentas acessórias para a resolução de cálculos ainda mais precisos para o tratamento analítico de sistemas mais complexos⁷.

Cálculos químicos-quânticos são essenciais nos laboratórios acadêmicos, empregados no entendimento de estados de transição, otimização geométrica, orbitais moleculares entre outras¹². Em suma, os métodos quânticos empregados derivam da resolução da equação de Schrödinger, onde a precisão dos resultados é inversamente proporcional ao custo de processamento do cálculo.

O presente trabalho tem por objetivo enunciar alguns conceitos dos métodos quânticos empregados na modelagem molecular, sendo estes os cálculos *Ab-initio*, Semiempíricos e Funcional da Densidade (DFT) a fim de servir como base para o entendimento de suas aplicações. O trabalho não vem a desmerecer os formalismos das equações da mecânica quântica e sim divulgar de uma maneira simples um pouco dos fundamentos, ficando a critério do leitor aprofundar-se na temática.

2. Métodos *Ab-initio*: Hatree-Fock e pós-Hatree-Fock.

Para entender os métodos quânticos, devemos primeiramente conhecer um pouco sobre o que é um Hamiltoniano (\hat{H}) e a função de onda (Ψ). O Hamiltoniano é um operador matemático que apresenta parâmetros que serão trabalhados a nível eletrônico descrito de forma resumida como as energias cinéticas e potencial na equação [1]. Na equação [1], os dois termos iniciais estão relacionados à energia cinética dos elétrons e dos núcleos respectivamente, o terceiro termo é a força eletrostática atrativa entre elétrons e prótons, e o quarto termo representa a repulsão eletrostática entre os elétrons. Em que ∇^2 representa o operador Laplaciano, Z_k é a carga do núcleo, r_{ik} a distância entre o i -ésimo elétron e do k -ésimo núcleo e R_{kl} a distância entre dois núcleos distintos.

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^N \nabla_i^2 - \sum_{k=1}^M \frac{1}{2M_k} \nabla_k^2 - \sum_{i=1}^N \sum_{k=1}^M \frac{Z_k}{r_{ik}} + \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} + \sum_{k=1}^M \sum_{l>k}^M \frac{Z_k Z_l}{R_{kl}} \quad (1)$$

A função de onda é uma função matemática que descreve a resolução da equação de Schrödinger para o sistema de estado¹⁹. Com isso podemos definir como equação de autovalor à aplicação de um operador à uma função que resulta em uma constante real (autovalor) multiplicada pela função inicial denominada de autofunção¹⁹. A equação de Schrödinger independente do tempo e não relativística (equação [2]) é uma equação de autovalor, resolvê-la segundo a mecânica quântica para um sistema que consiste em N elétrons e M núcleos implica obtenção da energia total (E) e demais propriedades²⁸.

(2)

$$\hat{H}\Psi = E\Psi$$

O método *Ab-initio* (do Latim: primeiros princípios) possibilita a obtenção das propriedades eletrônicas e atômicas a partir da resolução da equação de Schrödinger, tentando aproximar ao máximo suas soluções apenas utilizando dados de constantes físicas, operando com alta taxa de precisão. Sua evolução através do tempo faz uso de vários métodos eletrônicos aproximados a fim de soluções exatas²⁹.

A aproximação descrita por Borh-Oppenheimer ¹ é fundamental na mecânica quântica, nela a energia cinética do núcleo atômico pode ser desprezada pois a massa é superior aos elétrons. Assim, calculando uma aproximação da função de onda (Ψ) monoelétrica do núcleo e dos elétrons separadamente, este método descrever, com ótima precisão, o átomo de Hidrogênio. Infelizmente não podemos encontrar uma solução analítica para sistemas polieletrônicos²⁹..

Na busca de resolver tais limitações, o método de Hatree-Fock trabalha com um operador Hamiltoniano (\hat{H}) completo definindo uma função de onda (Ψ) global para o sistema sendo um produto de funções monoelétricas definidas na equação [3], aproximando as interações sem solucioná-las ¹.

(3)

$$\Psi(r_1 r_2 r_3 \dots r_n) = \Psi(r_1) \Psi(r_2) \Psi(r_3) \dots \Psi(r_n)$$

Porém, ao desconsiderar a repulsão elétron-elétron, assumindo que não há interação eletrônica dado que cada um está em suas devidas orbitas (isolado), o método não leva em consideração o Princípio da Exclusão de Pauli¹³. Desta forma, novas formulações matemáticas foram necessárias para descrever a repulsão elétron-elétron, levando ao surgimento do método

pós-Hatree-Fock que é conhecido como teoria do Campo autoconsistente (SCF) que utiliza em sua base determinantes de Slater, como visto na equação [4] ¹, para descrever os elétrons em suas orbitas.

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} X_a(X_1) & X_b(X_1) & \cdots & X_n(X_1) \\ X_b(X_2) & X_b(X_2) & \cdots & X_n(X_2) \\ \cdots & \cdots & \cdots & \cdots \\ X_a(X_N) & X_b(X_N) & \cdots & X_n(X_N) \end{vmatrix} \quad (4)$$

Determinantes de Slater são uma forma de obter funções antissimétricas de um sistema de N elétrons por combinação linear, permitindo que muitos elétrons sejam tratados dessa forma. Em resumo, na aplicação de matrizes na aproximação, o método é descrito pela equação de Roothaan–Hall. (equação 5). [5]

$$Fc = Sc\varepsilon \quad (5)$$

Sendo F a matriz de Fock, c a matriz do coeficiente, ε a matriz energética e S a matriz de sobreposição. O método de campo autoconsistente de Hatree-Fock (SCF-HF) é utilizado como base de softwares químicos-quânticos para cálculos Ab-initio. Entretanto, para sistemas de muitos átomos (proteínas, peptídeos, enzimas etc.), a solução é improvável e o custo computacional muito elevado, fazendo-se necessário a escolha de outros métodos¹⁹.

3. Métodos Semiempíricos.

A função de onda (Ψ) de sistemas diatômicos é fundamentada na distância internuclear como parâmetro. Para moléculas poliatômicas os ângulos, torções e ângulos diedros devem ser levados em consideração no cálculo da função de onda¹³.

Desenvolvidos com o objetivo de entregar resultados mais precisos quando aplicados às moléculas polieletrônicas sem ter um alto custo computacional³, o método quântico semiempírico utiliza um Hamiltoniano simples, com parâmetros ajustados a fim de aproximar os resultados da função de onda eletrônica calculada aos valores experimentais, o que proporciona uma interação quase completa num tratamento teórico¹³

Para entender os métodos semiempíricos, deve-se levar em consideração a divisão dos elétrons mais internos nos átomos, conhecidos como elétrons de caroço (*core* do inglês),¹ dos elétrons da camada de valência, pois é nos elétrons da última camada que ocorrem as reações químicas.

Deste modo os parâmetros utilizados nos softwares de química computacional para métodos semiempíricos, como o MOPAC,³² permitem calcular moléculas grandes do mesmo modo que os métodos de mecânica clássica e apresentar resultados de parâmetros eletrônicos semelhantes aos métodos *Ab-initio* a partir de algumas simplificações que poupam a resolução de algumas integrais nos cálculos.

Existem várias simplificações que evoluíram ao longo do tempo e do formalismo matemático da aproximação ZDO (*Zero Differential Overlap*) que calcula integrais de recobrimento. Aqui, destaca-se o CNDO (*complete neglect of differential overlap*), INDO (*Intermediate Neglect Of Differential*

Overlap), NDDO (*Neglect Of Diatomic Differential Overlap*), MNDO (*Modified Neglect Of Diferencial Overlap*) que utilizam funções de base tipo Slater (STO)²⁴.

Além das aproximações já citadas, houve uma evolução nos modelos matemáticos permitindo a criação dos modelos PM3 (*Parametric Method 3*), PM5 (*Parametric Method 5*) e PM6 (*Parametric Method 6*), derivados do modelo MNDO e que utilizam funções tipo Gaussiana (GTO) como base²⁴. O parâmetro RM1 (*Recife Model 1*) foi desenvolvido no Brasil, reparametrizando o modelo AM1, usando um conjunto mais completo de dados²⁷. A qualidade dos resultados obtidos, a partir dos métodos semiempíricos, depende das aproximações teóricas utilizadas para descrever o sistema. Além disso, deve-se também levar em consideração o modelo de parametrização empregado¹². Todavia, comparado ao método da Teoria do Funcional da Densidade (DFT), o método semiempírico possui um baixo custo computacional, resultando em um menor tempo de cálculo.

O método PM7 foi desenvolvido em 2013, parametrizado utilizando dados base do *ab initio* experimentais, adicionando um novo tipo de dados de referência destinados a definir melhor a estrutura do espaço de parâmetros, sendo testado modelando estruturas cristalinas e calores de formação de sólidos³⁰.

4. Teoria do Funcional da Densidade (DFT).

Ao invés de utilizar a função de onda (Ψ) como variável fundamental nos cálculos, a Teoria do Funcional da Densidade (*Density Functional Theory* do inglês) aplica a densidade de eletrônica (ρ), que sempre será a mesma independente do sistema

estudado baseando-se na função das coordenadas cartesianas (x,y,z) da densidade eletrônica o que difere dos métodos *Ab-initio* que aumentam a complexidade conforme o aumento de elétrons no sistema estudado, já que para um sistema de N elétrons, existem $3N$ variáveis a serem resolvidas, aumentando assim a complexidade do cálculo e custo computacional para a aplicação desse método²⁰.

A Teoria do Funcional de Densidade foi desenvolvida por Pierre Hohenberg (1934 – 2017) e Walter Kohn (1923 – 2016) com a finalidade de simular a interação elétron-elétron de um sistema polieletrônico utilizando um potencial efetivo que é uma função apenas da densidade eletrônica denotada por ρ . Ao se aplicar a teoria desenvolvida por Hohenberg e Kohn, é possível obter a energia do estado fundamental para o sistema e todas as demais propriedades de interesse.¹⁰ Pode-se afirmar que a energia eletrônica é função da densidade eletrônica, definida pelo quadrado do módulo da função de onda descrito na equação [6]^{20;7}.

$$(6)$$
$$\rho(E) = |\Psi(E)|^2$$

Assim, o operador Hamiltoniano utilizado no DFT contempla parâmetros eletrônicos que descrevem a repulsão elétron-elétron e a repulsão núcleo-núcleo como descrito na equação [7]

$$(7)$$
$$\hat{H} = T - U + V_e + E_{rep}$$

Sendo as componentes da equação:

$$T = - \sum_i^n \frac{1}{2} \nabla_i^2 \text{ (Descritor da energia cinética)}$$

$$U = - \sum_A^M \sum_i^N \frac{Z_A}{|R_A - r_i|} \text{ (Descritor do Potencial externo)}$$

$$V_e = \sum_{i < j}^N \sum_j^N \frac{1}{|R_i - R_j|} \text{ (Descritor da repulsão elétron-elétron e repulsão Coulombiana)}$$

$$E_{rep} = \sum_{B < A}^M \sum_A^M \frac{Z_A Z_B}{|Z_A - Z_B|} \text{ (Descritor da repulsão núcleo-núcleo)}$$

Cálculos utilizando o método DFT são bastante utilizados nas pesquisas da área de química pois fornecem dados como eletronegatividade, potencial químico, reatividade e dureza de compostos com precisão e baixo custo computacional, tornando-se uma das metodologias mais empregadas²⁰.

Existem vários funcionais empregados no método DFT, esses funcionais podem ser de troca, correlação ou híbridos (mistura de dois funcionais distintos) uso desses funcionais agrega na descrição das propriedades estudadas. Os funcionais no método DFT podem desenvolver as equações da mecânica quântica semelhante ao método *Ab-initio* empregando dados de constantes físicas ou utilizar de parametrizações como o método semiempírico e agregar dados empíricos⁷.

O funcional híbrido inclui o funcional de energia de troca com os funcionais de troca e correlação de gradiente generalizado. B3LYP é bastante aplicado para a descrição de sistemas de interesse orgânico, sendo composto por um funcional troca de A. D. Becke e 3 funcionais paramétricos de correlação de Lee-Yang-Parr⁴.

É interessante comentar que mesmo os funcionais do método DFT sejam eficientes para uma boa parte das aplicações, não é possível melhorá-los como os baseados na função de onda, sendo necessário fazer um comparativo com outros métodos para estipular o erro nos cálculos²⁴.

5. Algumas das aplicações dos Métodos Quânticos.

Existe uma grande variedade de aplicações para os métodos quânticos frequentemente utilizados nos procedimentos de Modelagem Molecular. Para utilizar essas aplicações deve-se ter um *input* (arquivo de entrada) configurado, nele deve ser escrito as propriedades que se quer analisar e as coordenadas cartesianas (x,y,z) do objeto de estudo e salvo no formato específico que o software de química computacional possa reconhecer para realizar os cálculos e assim gerar um arquivo de saída, conhecido como *output*²⁹.

Para construir esses arquivos de entrada e visualizar os arquivos de saída existem alguns softwares, como por exemplo, o software Avogadro²⁶ que pode ser usado para ambas as finalidades. Nos tópicos a seguir estão representadas algumas das várias aplicações que utilizam cálculos quânticos empregadas na modelagem molecular.

5.1 Cálculos de Ponto Único (*Single Point*).

Os cálculos de ponto único (do inglês: *single point*) são utilizados para prever exatamente como as propriedades de uma estrutura química se comportam sem que ela seja submetida a

uma otimização geométrica ou sofra mudanças na conformação espacial. Esse tipo de cálculo é utilizado para presumir a estabilidade de compostos químicos e seus mecanismos de reação¹⁹.

Um outro uso bastante relevante do cálculo de ponto único é na obtenção das espécies catiônica e aniônica no estudo da reatividade local feito pela função eletrônica de Fukui (f) definida pela equação [8] mostrada a seguir².

$$(8) \\ f = \left(\frac{\partial \rho(r)}{\partial N} \right)_{v(r)}$$

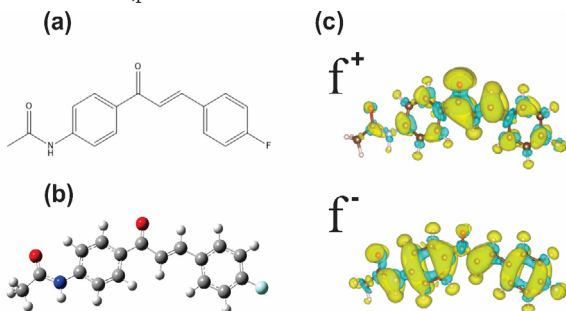
Em que $\rho(r)$ é a densidade eletrônica, N é o número de elétrons e $v(r)$ é o potencial externo que permanece constante na derivação. Se a molécula é suscetível à um ataque nucleofílico, a função eletrônica de Fukui que descreve esse sistema é denotada por f^+ e, se a molécula é suscetível à um ataque eletrofílico, a função de Fukui é denotada por f^- . As formas analíticas dessas funções estão mostradas nas equações [9] e [10] respectivamente. Note que a perspectiva para definição dos ataques (nucleofílico e eletrofílico) é a partir da molécula, isto é, no uso da função f^+ (ataque nucleofílico), a molécula atua como um eletrófilo, recebendo assim densidade eletrônica de outra espécie. No caso do uso da função f^- (ataque eletrofílico), a molécula age como um nucleófilo, doando densidade eletrônica para a outra espécie na qual está reagindo.

$$(9) \\ f^+ \approx \rho_{LUMO}$$

$$(10) \\ f^- \approx \rho_{HOMO}$$

Em que $LUMO$ e $HOMO$ representam as densidades eletrônicas para o Orbital Molecular Desocupado de Mais baixa Energia ($LUMO$ do inglês *Lowest Unoccupied Molecular Orbital*) e para o Orbital Molecular Ocupado de Maior Energia ($HOMO$ do inglês *Highest Occupied Molecular Orbital*). No trabalho de Almeida-Neto *et al.* (2020), uma molécula orgânica denominada N-{4'-[(E)-3-(Fluorofenil)-1-(fenil)-prop-2-en-1-ona]} (Figura 1a) teve sua reatividade estudada utilizando as funções eletrônicas de Fukui (Figura 1b), assim os autores conseguiram determinar quais átomos da molécula são mais suscetíveis a receber densidade eletrônica (ataque nucleófilo) e a doar densidade eletrônica (ataque eletrófilo) a partir de cálculos de ponto único utilizando o método de DFT com funcional B3LYP e conjunto de base 6-311++G(d,p) com metanol como solvente implícito.

Figura 1. (a) Fórmula estrutural da molécula N-{4'-[(E)-3-(Fluorofenil)-1-(fenil)-prop-2-en-1-ona]}, (b) estrutura otimizada da molécula e (c) funções de Fukui calculadas utilizando o método DFT com o nível da teoria B3LYP/6-311++G(d,p).



Fonte: Adaptada do trabalho de Almeida-Neto *et al.* (2020).

5.2 Orbitais Moleculares.

Os orbitais são funções matemáticas que descrevem a probabilidade de se encontrar elétrons em uma determinada região. Orbitais podem ser simétricos, ocupados por elétrons ou desocupados (virtual) servindo ao interesse químico e farmacológico, pois os orbitais HOMO e LUMO são importantes para se entender os conceitos de ligações químicas e reatividade. Os orbitais HOMO e LUMO influenciam diretamente nas substâncias químicas bioativas, deste modo, quanto maior a energia de HOMO maior será seu caráter elétron-doadora em uma reação e quanto menor a energia de LUMO, maior será sua capacidade elétron-aceitadora¹⁴. Os orbitais moleculares podem ser calculados utilizando qualquer um dos métodos quânticos apresentados e visualizados pelo Avogadro²⁶.

No trabalho de Almeida-Neto *et al.* (2020), os autores calcularam usando o método DFT com o nível computacional B3LYP/6-311++G(d,p) os orbitais moleculares da molécula orgânica (Figura 1a) de interesse do estudo e é possível observar o resultado obtido pelos pesquisadores na Figura 2. Segundo o trabalho, a densidade eletrônica disponível para doação se espalha principalmente sobre os anéis aromáticos da molécula e os orbitais moleculares vazios disponíveis para aceitação de carga também se espalham majoritariamente sobre os anéis aromáticos.

A quantificação da reatividade global da molécula pode ser feita pelo uso dos valores de energia dos orbitais HOMO (E_{HOMO}) e LUMO (E_{LUMO}) para computar os descritores quânticos de reatividade: *gap* de energia (ΔE_{gap} , equação 11)²², o potencial de ionização (I, equação 12)¹¹, a afinidade eletrônica (A, equação 13)¹¹, a eletronegatividade (χ , equação

14)^{9,6}, o potencial químico (μ , equação 14)¹⁸ a dureza global (η , equação 15) (JANAK, 1978;²², a maciez global (σ , equação 16)³¹ o índice global de eletrofilicidade (ω , equação 17)¹⁸ e o índice global de nucleofilicidade (ε , equação 18)¹⁸. Utilizando os visualizadores como o Avogadro, podemos obter os valores para as energias do HOMO e do LUMO e então computar os descritores pelas equações (11 – 18).

$$\Delta E_{gap} = E_{LUMO} - E_{HOMO} \quad (11)$$

$$I = - E_{HOMO} \quad (12)$$

$$A = - E_{LUMO} \quad (13)$$

$$\chi = \frac{I + A}{2} \quad (14)$$

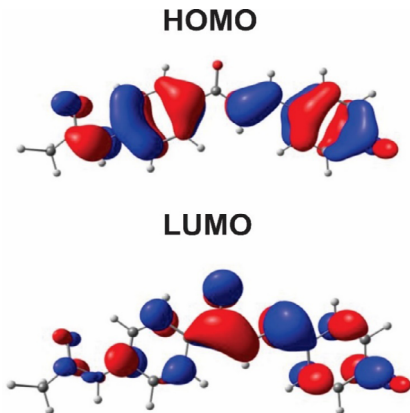
$$\eta = \frac{I - A}{2} \quad (15)$$

$$\sigma = \frac{1}{\eta} \quad (16)$$

$$\omega = \frac{\chi^2}{2\eta} \quad (17)$$

$$\varepsilon = \frac{1}{\omega} \quad (18)$$

Figura 2. Orbitais Moleculares HOMO e LUMO calculados utilizando o nível de teoria B3LYP/6-311++G(d,p) com o método de DFT.



Fonte: Adaptada do trabalho de Almeida-Neto *et al.* (2020).

Podemos também associar os dados de reatividade, com os dados experimentais, para avaliar as correlações de um modelo matemático teórico de previsibilidade. Os descritores de reatividade vistos nas equações [11-18] e com adição de descritores como: eletrodoação (ω^- , equação 19) eletrorecepção (ω^+ , equação 20)⁸ eletrofilicidade líquida ($\Delta\omega$, equação 21) e transferência de fração eletrônica (ΔN , equação 22)⁵, foram correlacionados com os dados de eficiência corrosiva e com o período de indução de antioxidantes (PDA, IONOL e cardanol) em biodiesel de babaçu e óleo de fritura²⁵. O resultado indicou que o descritor (ω^+) obteve excelentes índices, com correlações de no mínimo ($R^2 = 0.9928$) e excelente correlação de ($R^2 = 0.9999$) (Figura 3). Assim, para a eficiência corrosiva das três moléculas, quanto menor a característica da molécula em

receber a carga, maior sua tendência característica de ser um bom anticorrosivo, e sobre correlação com o dado de período de indução, quanto menor a tendência de receber elétrons, melhor sua tendência característica de ser um bom antioxidante.

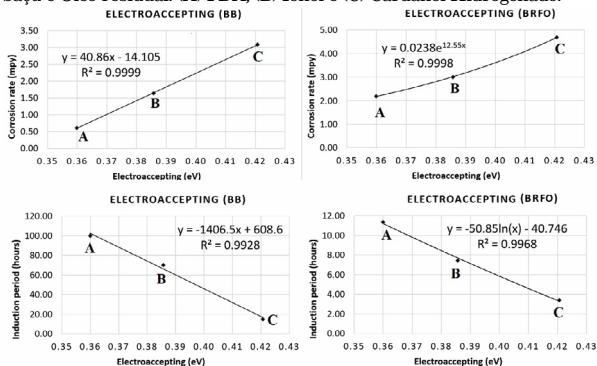
$$(19) \quad \omega^- = \frac{(3I + A)^2}{(16(I - A))}$$

$$(20) \quad \omega^+ = \frac{(I + 3A)^2}{(16(I - A))}$$

$$(21) \quad \Delta\omega = \omega^+ - \omega^-$$

$$(22) \quad \Delta N = \frac{\chi_{Cu} - \chi_{inh}}{2(\eta_{Cu} + \eta_{inh})}$$

Figura 3. Gráfico de correlação matemática obtido do melhor R^2 da taxa de corrosão / eletrorecepção e período de indução/eletrorecepção em Baça e Óleo residual. (A) PDA, (B) Ionol e (C) Cardanol Hidrogenado.



5.3 Otimização Geométrica.

A otimização geométrica é uma das aplicações mais utilizadas na Modelagem Molecular. A técnica modifica a conformação, ajustando os ângulos e os comprimentos de ligação na busca de uma maior estabilidade e menor energia de uma estrutura desenhada (mínimo global) no *input* (arquivo de entrada) ou obtida em um repositório on-line. A otimização de compostos candidatos a ‘ligantes’ é amplamente utilizada nas análises de Química medicinal e no planejamento racional de fármacos.¹⁹ Existem várias maneiras de otimizar geometricamente uma estrutura, softwares como o Orca¹⁷ e MOPAC³² utilizando respectivamente métodos *Ab-initio* e semiempírico para propor conformações estáveis de estruturas, ficando a critério do pesquisador escolher qual utilizar.

A molécula otimizada pelo método de DFT e nível computacional usando o funcional B3LYP/6-311++G (d,p) obtida no trabalho de Almeida-Neto *et al.* (2020) é mostrada na Figura 1a e os autores encontraram como resultado dos cálculos que a molécula não é totalmente planar: o ângulo diedro que liga o anel aromático que contém o grupo acetamida com o restante da estrutura é de 9,9° e o diedro que une o outro anel aromático à estrutura da molécula foi de aproximadamente 0,03°, mostrando que os dois anéis não estão no mesmo plano e esse fato é essencial para a reatividade dessa molécula. Os autores também compararam os comprimentos das ligações da molécula otimizada teoricamente com os resultados de comprimentos de ligações experimentais obtidos por difração de raio-X de moléculas similares na literatura, e eles encontraram que, por

exemplo, para a ligação entre o átomo de carbono e de flúor, o resultado teórico foi de 1,35827 Å e o experimental 1,3624 Å. Portanto, o cálculo de otimização geométrica é crucial para a obtenção de boas correlações com os resultados experimentais.

5.4 Análise conformacional.

Átomos em uma molécula sofrem rotações que causa uma modificação no arranjo espacial. Diferente de uma otimização geométrica, a análise conformacional na Modelagem Molecular permite aos químicos fazerem uma busca das conformações mais estáveis através da superfície energética potencial (SEP), explorando todo o espaço e obtendo dados termodinâmicos teóricos sobre conformações específicas de moléculas, verificando deste modo a estabilidade de composto a partir de seus ângulos de torção (diedros). Dependendo do método utilizado, as análises conformacionais podem ser obtidas em poucos segundos, sendo mais complexa para moléculas com várias ligações passíveis de giro, flexões e torções.³

5.5 Espectroscopia no Infravermelho e Raman

Os métodos quânticos permitem calcular também os modos de vibrações, uma propriedade intrínseca das moléculas. Essas vibrações podem gerar frequências específicas para cada tipo de molécula analisada, tornando-se prática nos estudos teóricos de Infravermelho (IV), assim como espectroscopia de Raman, que é uma técnica analítica fotônica de alta

resolução capaz de prever a estrutura e identificação de vários materiais, sejam orgânicos ou inorgânicos, sendo capaz de confrontar os valores obtidos de maneira experimental ou possibilitar uma melhor análise e obtidas das moléculas em uma faixa específica²⁹.

O método DFT descreve as frequências vibracionais das moléculas de maneira semelhante aos valores reais, sendo possível calcular espectros das técnicas de IV, UV e Raman utilizando simulações computacionais uma alternativa para o cálculo de espectros teóricos¹⁹. Na maioria dos softwares de química computacional, ao se calcular a frequências vibracionais, o software calculará também as propriedades termodinâmicas da molécula.

5.8 Termodinâmica e termoquímica.

A termodinâmica aplica conceitos importantes no entendimento da espontaneidade de reações. Para que moléculas de um reagente colidam e formem produtos, é necessária uma orientação adequada e uma energia cinética capaz de superar a energia do estado de transição. Estudos termodinâmicos das reações orgânicas ou de gases podem ser utilizados em Modelagem Molecular¹⁶.

Assim como termodinâmica, as propriedades termoquímicas são muito importantes, obter os dados entálpicos de uma reação quando se estuda mecanismos de reação e energias de ionização facilita a predição de propriedades a nível experimental, pois os resultados convergem próximos aos valores reais, o que valida os resultados. Programas de química com-

putacional como o MOPAC³² obtém os resultados de através da resolução de um Hamiltoniano simples tanto para termodinâmica como para a termoquímica, aplicando cálculos do método semiempírico.

Os parâmetros termodinâmicos podem ser estimados pela energia obtida dos cálculos para obter energia eletrônica da molécula. Para estimar a energia interna (U) é denotado que $U(X) = E_0 + ZPE + E_{\text{trans}} + E_{\text{rot}} + E_{\text{vib}}$, sendo, E_0 a energia eletrônica total do sistema, ZPE é a energia de ponto zero, E_{trans} , E_{rot} e E_{vib} são as contribuições translacionais, rotacionais e vibracionais para a entalpia, respectivamente. Entretanto, da termodinâmica clássica, entendemos que $H = U + PV$, todavia, para correção térmica da molécula isolada, é utilizado $H = U + K_b T$, onde K_b a constante de Boltzmann. Finalmente, $K_b T$ representa o termo PV (trabalho) e é acrescentado para converter a energia em entalpia. A entropia (S) pode ser avaliada por métodos padrão termodinâmicos estatísticos¹⁵.

6. Considerações finais.

As principais aplicações dos métodos quânticos podem-se resumir aos procedimentos de energia de ponto único, otimização geométrica, orbitais moleculares, espectros, frequências, entre outros. Portanto estes métodos e suas aplicações ajudam a desenvolver mais pesquisas na área de Química Computacional e são ferramentas acessíveis nos laboratórios acadêmicos. Diante dos tópicos apresentados pode-se concluir que existem várias possibilidades nos estudos de Modelagem Molecular que utilizam como base os métodos quânticos. O entendimen-

to dos métodos *Ab-initio*, semiempírico e DFT requer do pesquisador o conhecimento das teorias e a consciência de que cada método tem suas especificidades e limitações, assim faz-se necessário uma busca mais aprofundada na literatura com obras geralmente encontradas na língua inglesa.

Referências

1. ALCÁCER, L. **Introdução à Química Quântica Computacional**. Lisboa: IST Press, 2007.
2. ALMEIDA-NETO, F. W. Q. *et al.* Characterization of the structural, spectroscopic, nonlinear optical, electronic properties and antioxidant activity of the N-{4'-[(E)-3-(Fluorophenyl)-1-(phenyl)-prop-2-en-1-one]}-acetamide. **Journal of Molecular Structure**, v. 120, p. 128765 1-16, 2020.
3. ANDREI, C. C. *et al.* **Da química medicinal à química combinatória e modelagem molecular: um curso prático**. 2 ed. Manole. Barueri, SP. 2012. ISBN 9788520432709.
4. 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.
5. CHATTARAJ, P. K.; CHAKRABORTY, A.; GIRI, S. Net Electrophilicity. **Journal of Physical Chemistry A**, v. 113, n. 37, p. 10068–10074, 2009.
6. CHERMETTE, H. Chemical reactivity indexes in density functional theory. **Journal of Computational Chemistry**, v. 20, n. 1, p. 129–154, 1999.
7. GARCIA, R. D. A. **Modelagem molecular (TD-DFT) aplicada à simulação de espectros UV para cinamatos com perfil de filtros solares**. 2014. 60 f. Dissertação (Mestrado em Fármaco e Medica-

- mentos) - Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, São Paulo, 2014. Disponível em: <<https://www.teses.usp.br/teses/disponiveis/9/9138/tde-11082014-162702/en.php>>. Acesso em: 17 maio 2019.
8. GÁZQUEZ, J. L.; CEDILLO, A.; VELA, A. Electrodonating and electroaccepting powers. **Journal of Physical Chemistry A**, v. 111, n. 10, p. 1966–1970, 2007.
 9. ICZKOWSKI, R. P.; MARGRAVE, J. L. Electronegativity. **Journal of the American Chemical Society**, v. 83, n. 17, p. 3547–3551, 1961.
 10. JONES, R. O., Density functional theory: Its origins, rise to prominence, and future. *Review of Modern Physics*, vol. 87, p. 827 – 923, 2015.
 11. KOOPMANS, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. **Physica**, v. 1, n. 1–6, p. 104–113, 1934.
 12. LEAL, R. C. *et al.* A Química Quântica na compreensão de teorias de Química Orgânica. **Química Nova**, São Paulo, v. 33, n. 5, p. 1211-1215, 2010. Disponível em:<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422010000500037&lng=en&nrm=iso>. Acesso em: 30 set. 2020.
 13. LEVINE, I. N.; **Quantum Chemistry**. Prentice Hall, Upper Saddle River, NJ, EUA, ed. 7, 2013. Disponível em: <[http://www.dll.cheminfinity.com/arks/wp-file/books/Ira-N.Levine-Quantum-Chemistry-Pearson-\(2013\)-cheminfinity.com.pdf](http://www.dll.cheminfinity.com/arks/wp-file/books/Ira-N.Levine-Quantum-Chemistry-Pearson-(2013)-cheminfinity.com.pdf)>. Acesso em: 25 maio 2020.
 14. LOBATO, C. C. et al. **Estudo computacional da reatividade dos produtos isoméricos em reações de halidrificação através da análise dos orbitais de fronteiras – HOMO e LUMO**. In: CONGRESSO BRASILEIRO DE QUÍMICA, ABQ - ASSOCIAÇÃO BRASILEIRA DE QUÍMICA, 52, 2012, Recife, PE. Disponível em: <<http://www.abq.org.br/cbq/2012/trabalhos/13/1657-9612.html>>. Acesso em: 21 out. 2020. Acesso em: 25 maio 2020.
 15. MCQUARRIE, D. A. (Donald A. **Statistical mechanics**. New York: Harper & Row, 1975.

16. MORGON, N. H. Computação em química teórica: informações técnicas. **Química nova**, v. 24, n. 5, p. 676-682, 2001. Disponível em: <https://www.scielo.br/scielo.php?pid=S0100-40422001000500015&script=sci_arttext&tlng=pt>.
17. NEESE, F. Software update: the ORCA program system, version 4.0. **WIRES Computational Molecular Science**, v. 8, n. 1, e1327, 2018. Disponível em: <<https://doi.org/10.1002/wcms.1327>>. Acesso em: 16 maio 2019.
18. OBOT, I. B.; MACDONALD, D. D.; GASEM, Z. M. Density functional theory (DFT) as a powerful tool for designing new organic corrosion inhibitors: Part 1: An overview. **Corrosion Science**, v. 99, p. 1–30, 2015.
19. OLIVEIRA, A. M. **Introdução à Modelagem Molecular para Química, Engenharia e Biomédicas: Fundamentos e exercícios**. 1 ed. Curitiba: Appris, 2018.
20. PAULINO, P. H. S. **Estudo teórico da Fluoxetina**. 2018. 30 f. Trabalho de Conclusão de Curso (Bacharelado em Química) – Coordenadoria de Química, Universidade Federal de São João del-Rei, São João del-Rei, 2018. Disponível em: <https://ufsj.edu.br/portal2-repositorio/File/coqui/TCC/Monografia-TCC-Paulo_H_S_Paulino.pdf>. Acesso em: 16 out. 2019.
21. PAYNE, M. C., TETER, M. P., ALLAN, D.C., ARIAS, T. A., JOANNOPOULOS, J. D., Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Reviews of Modern Physics*, vol. 64, p. 1045 – 1097, 1992.
22. PEARSON, R. G. Hard and Soft Acids and Bases. **Journal of the American Chemical Society**, v. 85, n. 22, p. 3533–3539, 1963.
23. PEARSON, R. G. Recent advances in the concept of hard and soft acids and bases. **Journal of Chemical Education**, v. 64, n. 7, p. 561–567, 1987.
24. PESSÔA, K. F. *et al.* Revisão de alguns principais métodos utilizados em modelagem molecular: Parte II - Métodos quânticos. **Série Tecnologia Mineral**, Rio de Janeiro, 99, 51 f., 2018. ISSN 0103-7382. Disponível em: <<http://mineralis.cetem.gov.br/bitstream/cetem/2171/3/STM-99.pdf>>. Acesso em: 12 maio 2020.

25. RANGEL, N. V. P.; DA SILVA, L. P.; PINHEIRO, V. S.; FIGUEREDEDO, I. M.; CAMPOS, O. S.; COSTA, S. N.; LUNA, F. M. T.; CAVALCANTE, C. L.; MARINHO, E. S.; DE LIMA-NETO, P.; RIOS, M. A. S. Effect of additives on the oxidative stability and corrosivity of biodiesel samples derived from babassu oil and residual frying oil: An experimental and theoretical assessment. **Fuel**, v. 289, p. 119939, 2021.
26. RAYAN, B; RAYAN, A. Avogadro Program for Chemistry Education: To What Extent can Molecular Visualization and Three-dimensional Simulations Enhance Meaningful Chemistry Learning. **World Journal of Chemical Education**, v. 5, n. 4, p. 136-141, 2017. Disponível em: <<http://article.journalofchemicaleducation.com/pdf/wjce-5-4-4.pdf>> Acesso em: 7 mar. 2020.
27. ROCHA, G. B.; FREIRE, R. O.; SIMAS, A. M.; STEWART, J. J. P. RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. **Journal of Computational Chemistry**, v. 27, n. 10, p. 1101–1111, 2006.
28. SILVA, A. L. B. **Introdução a Química Quântica**. 2003. 86 f. 1 ed. Instituto de Física de São Carlos, Universidade de São Paulo. São Paulo, 2003. Disponível em: <<http://www.dominiopublico.gov.br/download/texto/ea000228.pdf>>. Acesso em: 21 out. 2020.
29. SILVA, F. I.; SOUZA, J. A. **Explorando a Química Computacional**. 1 ed. PLUS/ Simplíssimo. Porto Alegre, 2019. [Recurso digital]. ISBN 9788562069871
30. STEWART, J. J. P. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. **Journal of molecular modeling**, v. 19, n. 1, p. 1-32, 2013.
31. YANG, W.; PARR, R. G. Hardness, softness, and the Fukui function in the electronic theory of metals and catalysis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 82, n. 20, p. 6723–6726, 1985.
32. Stewart, J. J. P. MOPAC: A semiempirical molecular orbital program. **Journal of Computer-Aided Molecular Design**, v. 4, n. 1, p. 1–103, 1990.

CONCEITOS BÁSICOS SOBRE *DOCKING* MOLECULAR

Misael Bessa Sales

Universidade da Integração Internacional da
Lusofonia Afro-Brasileira, Acarape-Ceará

Lucas Lima Bezerra

Universidade Federal do Ceará – Fortaleza-Ceará

José Cleiton Sousa dos Santos

Universidade da Integração Internacional da
Lusofonia Afro-Brasileira, Acarape-Ceará

Eduardo Menezes Gaieta

Universidade da Integração Internacional da
Lusofonia Afro-Brasileira, Acarape-Ceará

Aluisio Marques da Fonseca

Universidade da Integração Internacional da
Lusofonia Afro-Brasileira, Acarape-Ceará

Resumo

A necessidade de se propor soluções para a área medicinal com a criação de drogas faz com que o processo de *docking* seja muito necessário e utilizado atualmente. Essa metodologia é crucial na redução do tempo para descobrir novos fármacos, tendo em vista a formulação de métodos que viabilizam esse modelo de triagem virtual. O *docking* molecular é utilizado na previsão de melhor posição de encaixe, orientação e conformação de determinada molécula ligante em proteínas, fazendo assim, a organização de melhor estruturação molecular para a otimização da interação proteína-ligante. Mas não somente isso, nos últimos anos, ele tem sido utilizado, também, para outros processos de interatividade como o de proteína-proteína e ácido nucleico-proteína-ligante. O método de *docking* molecular é periódico e seu funciona-

mento depende do poder de processamento computacional disponível, onde os resultados são disponibilizados após a realização dos cálculos necessários para a verificação da melhor acomodação das moléculas em utilização. Neste trabalho, serão apresentados conceitos básicos sobre o processo de *docking* molecular, visando mostrar na literatura diversos pontos e ferramentas para auxiliar no processo. Além disso, disponibilizar informações que vão contribuir para a manipulação dos resultados obtidos com a realização dos experimentos.

Palavras-chaves: *Docking* Molecular; Fármaco; Proteína; Ligante; Interação; Computacional.

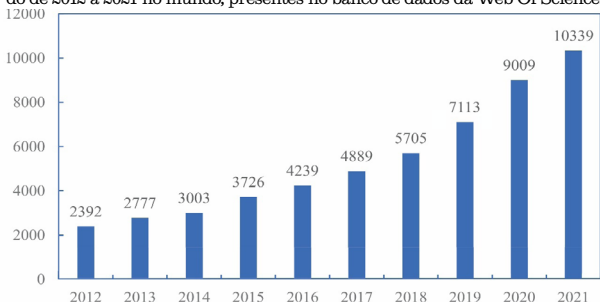
1. Introdução

Os desafios referentes ao descobrimento de novos fármacos são diversos quando os estudos são direcionados ao processo de triagem tradicional das moléculas. Uma das principais dificuldades na obtenção desses possíveis fármacos é a grande quantidade de moléculas a serem analisadas através de testes laboratoriais. Desse modo, é necessário um período maior de execução e, conseqüentemente, elevado custo financeiro para esse tipo de processo de triagem tradicional. Devido a esses desafios, diversos pesquisadores têm procurado alternativas para o desenvolvimento de novos fármacos. As pesquisas *in silico* tem recebido bastante destaque nos últimos anos devido ao baixo custo necessário para as simulações computacionais e a rapidez nessas análises quando comparada a triagem tradicional das moléculas. Dentre essas simulações computacionais, o *docking* molecular é uma técnica proveniente da modelagem molecular que objetiva prever a estrutura do complexo for-

mado entre um possível fármaco e um alvo macromolecular ⁷. Esse alvo pode ser uma proteína, enzima, ácido nucléico, canais iônicos e é associado a uma doença ou processo fisiopatológico no qual se deseja desenvolver um tratamento quimioterápico ⁹. O *docking* molecular também pode ser empregado em outros processos de interatividade como de proteína-proteína e ácido nucleico-proteína-ligante ¹².

Através da ferramenta de pesquisa do banco de dados da Web of Science foi possível visualizar um grande crescimento dos estudos utilizando o *docking* molecular nos anos de 2012 a 2021, totalizando 53192 artigos publicados. Dentre esses artigos publicados, 19348 trabalhos foram referentes aos anos 2020 e 2021, onde grande parte desses trabalhos foram referentes na busca de um fármaco que pudesse ser utilizado no tratamento do novo coronavírus, cuja pandemia se instalou em todos os países. Portanto, esse gráfico confirma a importância do *docking* molecular nos estudos para obtenção de novos fármacos, além disso, mostra o quanto esse método está sendo utilizado pelos pesquisadores no Brasil e no exterior (fig. 1).

Figura 1. Número de publicações com “Molecular Docking” durante o período de 2012 a 2021 no mundo, presentes no banco de dados da Web Of Science.



Fonte: <https://clarivate.com/webofsciencelgroup/solutions/web-of-science/>

Como visto, o surgimento da pandemia causada pelo novo Coronavírus trouxe impactos negativos sociais, econômicos além de grande número de óbitos, em todos os países, consequentemente, os pesquisadores do mundo inteiro estão em uma busca de um medicamento com alta taxa de eficácia no combate desse vírus. Dentre os diversos trabalhos publicados voltados para o vírus do SARS-CoV-2 o artigo de Almeida-Neto *et al*¹ indicou, através de simulações por *docking* molecular, que as quatro acetamidachalconas apresentaram um grande potencial inibitório sobre as proteínas do SARS-CoV-2, sendo uma forte indicação para a viabilidade de teste *in vitro*.

Para entender o método de *docking* molecular, primeiramente é necessário compreender como acontece a interação entre o complexo ligante-proteína. A cinética das interações ligante-proteína apresenta uma certa complexidade devido a grande quantidade de átomos presentes no alvo macromolecular. A especificidade das proteínas é devida à sua natureza

própria do espaço de ligação⁸. O ligante pode interagir em inúmeras posições no receptor, conseqüentemente, exibe uma grande quantidade de modos de ligação. Para definir o modo de ligação mais estável e provável é necessário avaliar as funções de pontuação do complexo.

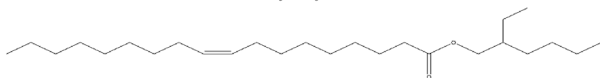
As funções de pontuação, ou de *scoring*, são usadas para selecionar as posições dentro dos grupos de poses gerados pelo sistema de amostragem. Exemplos de seleções, estão as funções que se baseiam nos campos de força, as funções empíricas e as funções baseadas no conhecimento⁵. É baseado nessa pontuação que será selecionado a melhor posição conformacional espacial para determinada interação, sendo proteína-ligante ou outras, sendo que a observação a ser feita é a de qual a menor pontuação, pois é a representação da quantidade de energia gasta no acoplamento das moléculas. Além disso, o Desvio Quadrático Médio da Raiz (RMSD, sigla em inglês) é uma função de pontuação aplicada para eleger os modos de ligação teoricamente mais próximos do modo de ligação “real, dentre os modos de ligação explorados pelo algoritmo de busca, classificando-os de acordo com as energias de interação¹⁶. Para validar as simulações de *docking* molecular, o resultado do RMSD deve ter valor de 1,0 a 3,0 Å entre o alvo macromolecular e o ligante^{3:4}.

Atualmente, diversos *softwares* de *docking* molecular comerciais e gratuitos são disponibilizados, tais como DOCK¹⁵, AutoDock¹⁰, FlexX¹¹, GOLD⁶ AutoDock Vina¹⁴ entre outros. Além dos programas citados acima ou outros que realizam a mesma funcionalidade, existem alguns outros *softwares* que se fazem necessárias para a iniciar as simulações por *docking* molecular.

O ChemDraw é um *software* que auxilia no desenho da estrutura do ligante. Nele poderá ser feita a aplicação da es-

estrutura desenhada, ou seja, o usuário vai manipular algumas funções de desenhos para construir a estrutura da molécula desejada. Além disso, o programa fornece a possibilidade de escrita da nomenclatura da molécula e assim gera o arquivo com a estrutura automaticamente desenhada.

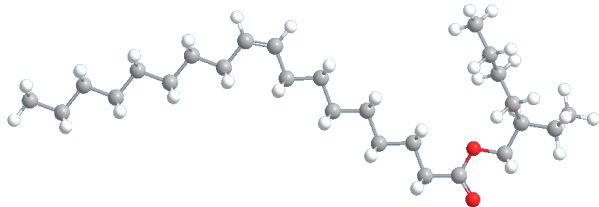
Figura 2. Estrutura do 2 - Ethylhexyl Oleate feita no programa ChemDraw.
2-Ethylhexyl oleate



Fonte: Elaborado pelo autor.

O Chem3D é outro exemplo de programa, onde neste caso, é o processo posterior ao ChemDraw[®], devido que nele será convertido o arquivo para uma extensão que contém a molécula em formato tridimensional. Esse último será usado para realizar o procedimento de otimização do ligante, e posteriormente, aplicado nos programas principais que irão gerar as informações essenciais após o *docking*. Essa parte é de grande importância pois é necessária essa estrutura 3D para o funcionamento das ferramentas de cálculo para a triagem virtual.

Figura 3. Estrutura do 2 - Ethylhexyl Oleate feita no programa Chem3D



Fonte: Elaborado pelo autor.

Antes de iniciar as simulações é necessário realizar uma preparação do alvo macromolecular e do fármaco a ser testado. A estrutura tridimensional do alvo macromolecular pode ser obtida gratuitamente através do repositório Protein Data Bank (PDB) (<https://www.rcsb.org/>). Em seguida, é realizado a preparação desse alvo através da remoção das moléculas de água e de outros resíduos presentes na estrutura cristalizada, seguido da adição de hidrogênios polares no alvo. Posteriormente, a otimização geométrica do ligante é realizada para obter a estrutura no mínimo global de energia. Finalmente, as simulações de *docking* molecular podem ser realizadas.

2. Algoritmo genético lamarkiano e sobre método estocástico

2.1. Algoritmo de busca

Os algoritmos de busca estão presentes em todas as simulações de *docking* molecular. O algoritmo de busca deve criar um conjunto ideal de regras e configurações afim de explorar um amplo cenário do ligante na bolsa de ligação. Os algoritmos realizam modificações incrementais nos parâmetros estruturais dos ligantes, como graus de liberdade de torção (diédrico), translacional e rotacional. As vezes também consideram a flexibilidade das cadeias laterais dos aminoácidos encontrados na bolsa de ligação, com a finalidade de encontrar conformações de ligantes com interações favoráveis.²⁰

De acordo com a metodologia aplicada para explorar a flexibilidade do ligante os algoritmos de busca são divididos em três respectivos grupos: buscas sistemáticas, estocásticas e determinísticas.²¹

2.2. Método Estocástico

O termo estocástico significa aleatório ou acaso, e diferente de um modelo determinístico que prediz apenas um único resultado de um determinado conjunto de circunstâncias, um modelo estocástico prevê um conjunto de resultados possíveis ponderados por suas probabilidades.⁵⁰

Em um método estocástico, uma variável aleatória é definida como um objeto X na qual o seu valor está sujeito a variações devido à aleatoriedade. Isto indica que esta variável pode assumir um conjunto de valores diferentes possíveis, ou seja, ela não tem valor fixo e único, e cada valor está associado a uma probabilidade. Diz-se que uma variável aleatória é discreta $\{X, t \geq 0\}$ se a mesma pode assumir um número de valores contáveis, ou seja se existe um conjunto contável, exemplo: número de elétrons na banda de condução de um semicondutor, o número de moléculas de um determinado componente em uma mistura reagente. Ou a variável aleatória pode ser contínua $\{X(t), t \geq 0\}$ em um dado intervalo: um componente de velocidade da partícula browniana (intervalo $-\infty, +\infty$); a energia cinética daquela partícula $(0, \infty)$; a diferença de potencial entre os pontos finais de uma resistência elétrica ($-\infty, +\infty$), exemplo: a energia de um elétron na presença de um centro de ligação.^{43,51}

Vale ressaltar que as variáveis aleatórias assumem valores no conjunto S , chamado de espaço de estados ou espaço de busca do processo estocástico. E em muitas aplicações o parâmetro “ t ” representa o tempo, mas este parâmetro também pode representar qualquer índice.⁴⁰

Um processo estocástico é uma coleção de variáveis aleatórias: $\{X(t)\}$, considerando um processo em tempo contínuo, ou $\{X_n\}$, considerando um processo em tempo discreto, na qual o t é definido como um parâmetro executado sobre um conjunto de índices adequado $T = \{0, 1, 2, \dots\}$. Em essência, o processo estocástico é uma função com dois argumentos, $\{X(t, \omega), t \in T, \omega \in \Omega\}$, em que Ω indica o espaço amostral ou o conjunto do processo. O processo estocástico assume valores no espaço de estados ou espaço de busca S , o que implica que Ω é o conjunto de todas as séries temporais possíveis, podendo ser geradas pelo processo.^{40,43,51}

Seja $x: T \rightarrow S$ uma função, e $\{X(t), t \in T\}$ um processo estocástico com espaço de estados S , então, pode-se dizer que: $\{x(t), t \in T\}$ é uma possível evolução (trajetória) de $\{X(t), t \in T\}$. As funções x são chamadas de caminhos de amostra do processo estocástico.^{40,45}

Estes conceitos foram introduzidos dentro da química computacional e os métodos estocásticos passaram a ser usados dentro dos procedimentos de docking molecular.³⁶ No docking, na etapa de busca espaço conformacional, alguns softwares fazem o uso do algoritmo de busca estocástico para a modificação incremental dos parâmetros estruturais dos ligantes, como grau de liberdade de torção (diédrico), translacional e rotacional.^{30,42}

De fato, o processo de docking molecular de um composto em um sítio de ligação do receptor é computacionalmente intensivo e exigente, o que pode ser identificado como um problema de minimização de energia. Contudo, existe o fato de que a maioria dos programas de mecânica molecular serem sensíveis aos locais mínimos para encontrar as poses de en-

caixe desejadas. Deste modo, os métodos de busca estocástica podem ser pensados como uma alternativa fácil na exploração de minimização de energia.^{35,57}

De modo geral, os algoritmos estocásticos realizam uma busca nas orientações e conformações de ligação do ligante, mas em invés de sistemática, realizam alterações aleatórias aos parâmetros estruturais do ligante em cada etapa no espaço conformacional e no espaço translacional/rotacional do respectivo ligante. Isto gera um conjunto de conformações moleculares e preenche uma ampla variedade de cenário energético, com o propósito de evitar a contenção da solução final em um mínimo de energia local e aumentar a probabilidade de achar um mínimo global. A mudança aleatória será aceita ou rejeitada de acordo com um critério probabilístico para manter ou não uma determinada conformação de teste. No entanto, novas conformações de teste são selecionadas por alguma alteração aleatória de uma conformação já testada. Um fator limitante deste procedimento é o custo computacional.^{30,38}

A vantagem dos algoritmos de busca estocásticos está na velocidade para encontrar a solução ideal muito rapidamente. E como desvantagem, observa-se que estes algoritmos, não garantem uma busca completa do espaço conformacional, de tal modo que a verdadeira solução pode ser perdida. A falta de convergência é parcialmente resolvida aumentando o número de iterações do algoritmo.⁴⁸

Algoritmos Evolutivos (EA), Algoritmos de Otimização de Enxame (SO), Algoritmos, Algoritmos de Busca Tabu (TS) e Algoritmos de Monte Carlo (MC), são quatro algoritmos típicos que pertencem à classe de métodos.¹⁸

2.3. Os Algoritmos Evolutivos

Os Algoritmos Evolutivos são algoritmos de otimização meta-heurística que são baseados na ideia de evolução biológica, como: reprodução, mutação, recombinação e seleção. Neste método são empregados os conceitos da biologia como: cromossomo, mutação, gene, cruzamento.^{26,48}

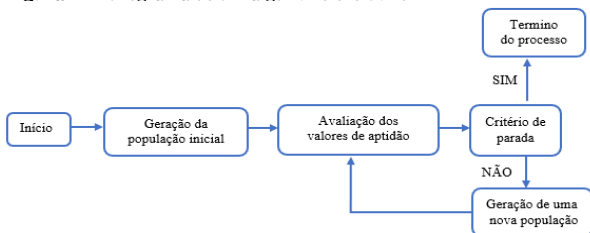
Em geral para a funcionalidade deste algoritmo, um conjunto (população) aleatório de soluções candidatas, identificadas como indivíduos ou cromossomos é selecionado de acordo com alguma representação da solução. Em particular os graus de liberdade são codificados em genes e cada conformação do ligante é descrita por um cromossomo (coleção de genes).^{26,48}

Os algoritmos evolutivos usam a função de pontuação ou a função de aptidão sobre a solução atual para determinar a qualidade e classificar a solução candidata dentro da população para fins de seleção. A solução de cromossomos (pais) é submetida a uma nova descendência, baseada em variações aleatórias de cruzamentos e mutações para evoluírem para indivíduos cada vez melhores, e os cromossomos com maior aptidão sobrevivem e substituem os piores.^{26,43}

O processo de evolução continua até o critério de término especificado pelo usuário for satisfeito, e uma nova população ser gerada. Caso o critério de parada não for satisfeito, uma nova população será gerada novamente e o processo será repetido até que um ou mais critérios de parada forem satisfeitos. O critério de término pode ser estático ou dinâmico.^{26,43}

De forma resumida, a Figura 1 exibe as etapas de funcionamento de um algoritmo evolutivo.

Figura 1. Fluxograma de um algoritmo evolutivo.



Fonte: Elaborado pelo autor.

2.4. O algoritmo de otimização de enxame

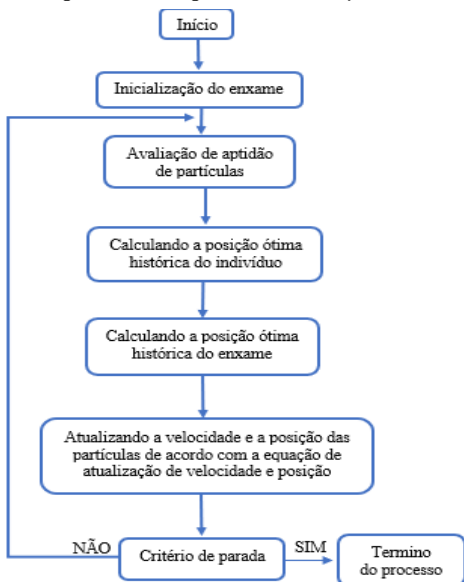
O algoritmo de otimização de enxame é uma técnica de otimização estocástica, baseada no comportamento social de animais. Este algoritmo de otimização de enxame simula o comportamento social de animais como: insetos, rebanhos, pássaros e peixe. Estes exames coadunam uma forma cooperativa de achar comida, e cada membro dos enxames vai mudando o seu padrão com forme suas próprias experiências e aprendizagem e de outros membros. Este mecanismo é usado para guiar as partículas na busca de soluções globais satisfatórias.^{17,54}

No algoritmo de otimização de enxame, as soluções candidatas (por exemplo, peixes) são chamadas de partículas. Com base nisso, primeiramente, este algoritmo começa inicializando a população. Em segundo, são calculados os valores de aptidão de cada partícula. Essas partículas viajam através de um espaço de busca para encontrar uma solução ideal, realizando interações e compartilhando informações com partículas vizinhas, ou seja, a sua melhor solução individual e computando a melhor vizinhança. Posteriormente segue-se de uma atualização dos melhores e posteriormente, a velocidade e a posição das partículas são atualizadas.^{17,24}

Em cada geração de melhor indivíduos, a melhor solução global obtida de todo enxame é atualizada. Se apropriando de todas estas informações, chega-se à convergência é alcançada atraindo as partículas para a localização do espaço de busca onde foi obtido a melhor solução. Uma função de aptidão é usada para avaliar a melhor partícula.^{24, 54} Caso o critério de parada não for satisfeito, o processo é repetido desde a segunda etapa até que um ou mais critérios de parada forem satisfeitos.

O fluxograma deste algoritmo é mostrado na figura abaixo:

Figura 2. fluxograma de um algoritmo de otimização de enxame



Fonte: Elaborado pelo autor.

2.5. O algoritmo de Busca Tabu (TS)

Um problema com algoritmos de busca local é que eles tendem a ficar presos em uma parte limitada do espaço de busca. O algoritmo de Busca Tabu (TS) resolve isso, pois ele baseia-se na premissa da incorporação de memória adaptativa e explorativa responsiva para a resolução de problemas de busca local. Desta forma, o método de busca tabu permite que o algoritmo se mova para soluções vizinhas que são piores do que a solução atual se não houver vizinhos melhores disponíveis.^{31,46}

O algoritmo de Busca Tabu consiste em métodos baseados em seguir um local de busca sempre que um ótimo local é encontrado, permitindo movimentos não melhorados. O algoritmo evita o retorno às soluções já visitadas anteriormente através do uso de memórias, chamadas tabu, na qual registram os recentes históricos de busca.^{32,33} As regras do impedimento destes retornos são fornecidas pelo usuário.⁵⁶

O algoritmo de busca tabu também foi adaptado a várias ferramentas e softwares de docking. No docking, uma posição gerada aleatoriamente é analisada comparando-a com as existentes na lista tabu. Caso a posição aleatoriamente gerada não forneça arranjos com diferenças significativas ou melhoramento drástico na energia mínima comparado com as posições já armazenadas na lista tabu, o algoritmo vai impor restrições que impedem a busca de revisitar áreas conformacionais já exploradas, promovendo assim a análise de novas regiões.^{48,49}

2.6. Algoritmos de Monte Carlo (MC)

O método Monte Carlo (MC) é um método estocástico, na qual faz uso de uma sequência de números aleatórios para

construir a amostra de uma solução, na qual podem ser obtidas variáveis probabilísticas de um determinado parâmetro/problema em interesse.^{39,55}

Devido ao grande número de estados possíveis que um sistema pode ter, torna-se ineficiente calcular o valor médio de todos esses estados, desta forma o método Monte Carlo foca em uma amostra e estima esses valores médios, a partir de distribuições de probabilidade.²²

A simulação de Monte Carlo é agora uma ferramenta científica muito usada, com ela, pesquisadores exploram sistemas complexos, examinam quantidades ocultas em experimentos e repetem ou modificam experimentos facilmente. Uma das vantagens das simulações de Monte Carlo é a exigência de enormes recursos computacionais.³⁴

No contexto da simulação molecular, o algoritmo de Monte Carlo gera conformações aleatórias de graus de liberdade geométricos do ligante em vez de integração das equações de movimento de Newton. E usa um critério de aceitação baseado em energia que promove a convergência pseudo-trajetória para um mínimo energético.^{23,28}

Então, se a pontuação de energia da conformação for melhorada, a mudança é aceita, caso contrário, ela é aceita de acordo com a probabilidade que é calculada usando a seguinte função de probabilidade de Boltzmann:

$$P \sim \exp \left[\frac{-(E_1 - E_0)}{k_B T} \right]$$

Onde E_0 e E_1 representam respetivamente as pontuações de energia do ligante antes e depois da mudança aleatória; kB é a constante de Boltzmann e T é a temperatura absoluta do sistema.^{37,52}

2.7. Algoritmo genético lamarckiano

O algoritmo genético lamarckiano (LGA) é um algoritmo híbrido, resultado da combinação do método de busca local com algoritmo genético. Este algoritmo permite a transmissão das características aprendidas de pais para filhos. No LGA, o cromossomo inviável é substituído por sua versão reparada após a aplicação da heurística de reparo local.^{25,44}

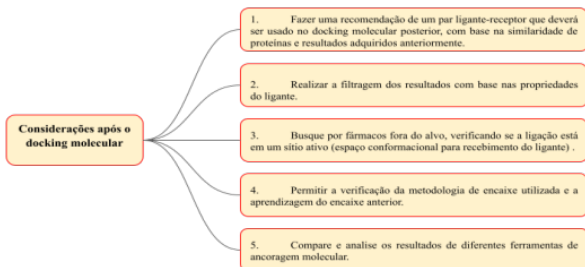
O procedimento do GA começa com uma população (P) de n cromossomos é inicializada aleatoriamente. Esta população pode ser gerada a partir de uma distribuição aleatória gaussiana para aumentar a diversidade. A população inclui múltiplas soluções, que representam cromossomos de indivíduos, onde o usuário define o número de indivíduos na população. A aptidão de cada cromossomo em P é calculada. Dois cromossomos dizem que $C1$ e $C2$ são selecionados da população P de acordo com o valor de aptidão. O operador de cruzamento de ponto único com probabilidade de cruzamento (Cp) é aplicado em $C1$ e $C2$ para produzir um descendente, chamado “ D ”. Depois disso, o operador de mutação uniforme é aplicado no descendente produzido (D) com probabilidade de mutação (Mp) para gerar D' . O novo descendente D' é colocado na nova população. As operações de seleção, cruzamento e mutação serão repetidas na população atual até que a nova população esteja completa.^{19,41,47}

Uma geração consiste em cinco estágios: mapeamento e avaliação de aptidão, seleção, cruzamento, mutação e seleção elitista, nessa ordem. No Algoritmo Genético Lamarckiano, cada geração é seguida por busca local, sendo realizada em uma proporção da população definida pelo usuário. As questões mais importantes que surgem em algoritmos genéticos lamarckiano envolvem o mapeamento de desenvolvimento, que transforma as representações genóticas em fenotípicas.⁴⁴

3. Análise dos resultados de docking molecular

Alguns pontos devem ser considerados no momento de realizar a análise dos resultados após o *docking*, pois é necessário se ater às condições que vão ser de muita importância para a construção do conhecimento científico mundial. Segue um fluxograma na figura 4 para este procedimento.

Figura 4. Fluxograma das considerações que devem ser feitas pelo cientista após a realização do procedimento de *docking*⁴³.



Fonte: Elaborado pelo autor.

1. Os pesquisadores, podem não saber todas as informações necessárias sobre um determinado receptor, então vão buscar recomendações externas de moléculas que obtiveram sucesso no acoplamento.
2. Filtrar os resultados é necessário para as condições de pesquisa do pesquisador, principalmente em relação aos testes laboratoriais.
3. Drogas fora do alvo, foram projetadas para se ligar a um receptor, mas adicionalmente se ligam a outro receptor do mesmo ou de um organismo diferente.
4. Toda informação gerada a partir do *docking* molecular é válida, mesmo que sejam resultados imprecisos e erros de cálculos, pois outros indivíduos poderão se basear nesses dados para a realização de procedimentos futuros.
5. É um passo muito importante, pois analisando os resultados utilizando diferentes algoritmos computacionais que possibilitam o processo de *docking*, é possível ter dados mais precisos sobre o procedimento experimental. Poderá ser observado se há uma diferença significativa no processo e, ainda, mostrar ao pesquisador possibilidades de como deverá ser a abordagem sob a metodologia de experimentação.

4. Considerações finais

O processo de *docking* molecular é um grande avanço científico, pois possibilita a manipulação de uma grande quantidade de dados moleculares e fazer a triagem virtual. O presente trabalho mostrou alguns conceitos básicos de *docking* molecular e como se dá todo o processo computacional, explicando de for-

ma análoga todo o procedimento experimental realizado, com base na literatura científica. Ademais, algumas ferramentas para a realização do *docking* foram explanadas para maior conhecimento sobre o assunto, e destacando que não somente elas existem para tal ação. O mundo conta com diversas metodologias e algoritmos que vão auxiliar os cientistas, e podem ser exploradas com mais detalhe através do referenciamento. Portanto, foi observado que, apesar de todo o processo de *docking* molecular ser complexo, as tecnologias estão cada vez mais ajudando a facilitar, pois surgem a cada dia, maior poder de processamento por causa da evolução tecnológica computacional, que proporciona métodos de cálculos mais precisos e rápidos.

5. Referências

1. ALMEIDA-NETO, F. W. et al. In silico study of the potential interactions of 4'-acetamidechalcones with protein targets in SARS-CoV-2. **Biochemical and Biophysical Research Communications**, v. 537, p. 71–77, 2021.
2. ChemOffice, perkinelmerinformatics.com, 2021, Disponível em: (<https://perkinelmerinformatics.com/products/research/chemdraw/>)
3. ERICKSON, J. A.; JALAIE, M.; ROBERTSON, D. H.; LEWIS, R. A.; VIETH, M. "Lessons in Molecular Recognition: The Effects of Ligand and Protein Flexibility on Molecular *Docking* Accuracy". **J. Med. Chem.**, 47:45, 2004.
4. HAMMES, A. S. O. Modelagem Molecular de inibidores de Aspartil Protease: potenciais novos compostos antimalariais. [Dissertação] Instituto Oswaldo Cruz. Programa de PósGraduação em Biologia Computacional e Sistemas. Rio de Janeiro, 2012.

5. HUANG, S.-Y.; GRINTER, S. Z.; ZOU, X. Scoring functions and their evaluation methods for protein–ligand *docking*: recent advances and future directions. **Physical Chemistry Chemical Physics**, v. 12, n. 40, p. 12899, 2010.
6. JONES G, WILLETT P, GLEN RC, LEACH AR, TAYLOR R. Development and validation of a genetic algorithm for flexible *docking*. **J Mol Biol**. 1997;267:727–748.
7. LEACH, A. R. Molecular modeling - principles and applications. 2ª ed. Hall (England): 2003. 744pp.
8. LIU, S., Chapter 9 - Protein-ligand interactions, Elsevier, 2020, Pages 351-399, ISBN 9780128210123.
9. MORGON, N. H.; COUTINHO, K. **Métodos de Química Teórica E Modelagem Molecular**. Editora Livraria da Física, 2007.
10. ÖSTERBERG F, MORRIS GM, SANNER MF, OLSON AJ, GOODSELL DS. Automated *docking* to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. **Proteins**. 2002; 46:34–40.
11. RAREY M, KRAMER B, LENGAUER T, KLEBE G. A fast flexible *docking* method using an incremental construction algorithm. **J Mol Biol**. 1996; 261:470–489.
12. STANZIONE F; GIANGRECO, I.; COLE, J. ; Chapter Four - Use of molecular *docking* computational tools in drug discovery, Volume 60, 2021, Pages 273-343,
13. TEMELKOVSKI, D.; KISS, T.; TERSTYANSZKY, G.; GREENWELL, P.; Extending molecular *docking* desktop applications with cloud computing support and analysis of results, **Future Generation Computer Systems**, Volume 97, 2019, Pages 814-824,
14. TROTT O, OLSON AJ. AutoDock Vina: improving the speed and accuracy of *docking* with a new scoring function, efficient optimization, and multithreading. **J Comput Chem**. 2010;31:455–461.
15. VENKATACHALAM CM, JIANG X, OLDFIELD T, WALDMAN M. LigandFit: a novel method for the shape-directed rapid *docking* of ligands to protein active sites. **J Mol Graph Model**. 2003;21:289–307.

16. VERDONK, M. L.; COLE, J. C.; HARTSHORN, M. J.; MURRAY, C. W. ET AL. Improved protein-ligand *docking* using GOLD. **Proteins**, v. 52, n. 4, p. 609-23, 2003.
17. AB WAHAB, M. N.; NEFTI-MEZIANI, S.; ATYABI, A. A Comprehensive Review of Swarm Optimization Algorithms. **PLOS ONE**, v. 10, n. 5, p. e0122827, 18 maio 2015.
18. AKHTER, M. Challenges in docking: Mini review. **JSM Chemistry**, v. 4, n. 2, p. 1025, 2016.
19. AUDET, C.; HARE, W. Genetic Algorithms. In: **Springer Series in Operations Research and Financial Engineering**. [s.l: s.n.]. p. 57-73.
20. BITENCOURT-FERREIRA, G.; DE AZEVEDO, W. F. How Docking Programs Work. In: JR, W. F. D. A. (Ed.). **Docking Screens for Drug Discovery**. 1. ed. Porto Alegre, RS, Brazil: Springer Science+Business Media, 2019. p. 35-50.
21. BROOIJMANS, N.; KUNTZ, I. D. Molecular Recognition and Docking Algorithms. **Annual Review of Biophysics and Biomolecular Structure**, v. 32, n. 1, p. 335-373, jun. 2003.
22. CEVALLOS-TORRES, L.; BOTTO-TOBAR, M. Monte Carlo Simulation Method. In: **Studies in Computational Intelligence**. [s.l: s.n.]. v. 824p. 87-96.
23. CIEMNY, M. et al. Modeling of Disordered Protein Structures Using Monte Carlo Simulations and Knowledge-Based Statistical Force Fields. **International Journal of Molecular Sciences**, v. 20, n. 3, p. 606, 31 jan. 2019.
24. COUCEIRO, M.; GHAMISI, P. Particle Swarm Optimization. In: **Fractional Order Darwinian Particle Swarm Optimization: Applications and Evaluation of an Evolutionary Algorithm**. SpringerBriefs in Applied Sciences and Technology. Cham: Springer International Publishing, 2016. p. 1-10.
25. CUADRA, L. et al. A Lamarckian Hybrid Grouping Genetic Algorithm with repair heuristics for resource assignment in WCDMA networks. **Applied Soft Computing**, v. 43, p. 619-632, jun. 2016.

26. DEVI, R. V.; SATHYA, S. S.; COUMAR, M. S. Evolutionary algorithms for de novo drug design - A survey. **Applied Soft Computing Journal**, v. 27, p. 543–552, 2015.
27. DIAS, R.; DE AZEVEDO JR., W. Molecular Docking Algorithms. **Current Drug Targets**, v. 9, n. 12, p. 1040–1047, 1 dez. 2008.
28. EARL, D. J.; DEEM, M. W. Monte Carlo simulations. **Methods in Molecular Biology**, v. 443, p. 25–36, 2008.
29. FERREIRA, L. et al. Molecular Docking and Structure-Based Drug Design Strategies. **Molecules**, v. 20, n. 7, p. 13384–13421, 22 jul. 2015a.
30. FERREIRA, L. et al. Molecular Docking and Structure-Based Drug Design Strategies. **Molecules**, v. 20, n. 7, p. 13384–13421, 22 jul. 2015b.
31. FRED, R. G.; MARTÍ. **Tabu Search**. Boston, MA: Springer US, 1997.
32. GENDREAU, M. Chapter 2: AN INTRODUCTION TO TABU SEARCH. In: **Search Methodologies**, Springer. Montréal: [s.n.]. p. 165–186.
33. GENDREAU, M.; POTVIN, J.-Y. Tabu Search. In: BURKE, EDMUND K.; STIRLING, UNIVERSITY OF BUILDING; COTTRELL SCOTLAND, U. K. (Ed.). **Search Methodologies**. Second Edi ed. Boston, MA: Springer US, 2014. p. 243–263.
34. HARRISON, R. L.; GRANJA, C.; LEROY, C. **Introduction to Monte Carlo Simulation**. AIP Conference Proceedings. **Anais...**2010. Disponível em: <<http://aip.scitation.org/doi/abs/10.1063/1.3295638>>
35. HATMAL, M. M.; TAHA, M. O. Combining Stochastic Deformation/Relaxation and Intermolecular Contacts Analysis for Extracting Pharmacophores from Ligand–Receptor Complexes. **Journal of Chemical Information and Modeling**, v. 58, n. 4, p. 879–893, 23 abr. 2018.
36. HOU, T. et al. Automated docking of peptides and proteins by using a genetic algorithm combined with a tabu search. **Protein Engineering, Design and Selection**, v. 12, n. 8, p. 639–648, ago. 1999.

37. HUANG, S. Y.; ZOU, X. Advances and challenges in Protein-ligand docking. **International Journal of Molecular Sciences**, v. 11, n. 8, p. 3016–3034, 2010a.
38. HUANG, S.-Y.; ZOU, X. Advances and challenges in protein-ligand docking. **International journal of molecular sciences**, v. 11, n. 8, p. 3016–34, 18 ago. 2010b.
39. JOHANSEN, A.; EVERS, L. **MCM_Sequential Monte Carlo**. [s.l.: s.n.].
40. KULKARNI, V. G. **Modeling and Analysis of Stochastic Systems**. Third Edit ed. Boca Raton: Taylor & Francis Group, 2017.
41. LAMBORA, A.; GUPTA, K.; CHOPRA, K. **Genetic Algorithm-A Literature Review**. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMIT-Con). **Anais...IEEE**, fev. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8862255/>>
42. MENG, X.-Y. et al. Molecular docking: a powerful approach for structure-based drug discovery. **Current computer-aided drug design**, v. 7, n. 2, p. 146–57, jun. 2011.
43. MORALES, J. M. et al. **Integrating Renewables in Electricity Markets**. Boston, MA: Springer US, 2014. v. 205
44. 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, 15 nov. 1998.
45. NEUMANN, F.; WITT, C. Stochastic Search Algorithms. In: **Bioinspired Computation in Combinatorial Optimization**. Natural Computing Series. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 12.
46. RENMAN, C.; FRISTEDT, H. A comparative analysis of a Tabu Search and a Genetic Algorithm for solving a University Course Timetabling Problem. 2015.
47. SAEED, G. Structural Optimization for Frequency Constraints. In: **Metaheuristic Applications in Structures and Infrastructures**. First Edit ed. [s.l.] Elsevier, 2013. p. 389–417.

48. SALMASO, V.; MORO, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. **Frontiers in Pharmacology**, v. 9, p. 16, 22 ago. 2018.
49. SOUSA, S. F.; FERNANDES, P. A.; RAMOS, M. J. Protein-ligand docking: Current status and future challenges. **Proteins: Structure, Function, and Bioinformatics**, v. 65, n. 1, p. 15–26, 21 jul. 2006.
50. TAYLOR, HOWARD M.; KARLIN, SAMUEL. **An Introduction to Stochastic Modeling Third Edition**. 3rd ed. ed. California: Academic Press, 1998.
51. VAN KAMPEN, N. G. STOCHASTIC VARIABLES. In: **Stochastic Processes in Physics and Chemistry**. [s.l.] Elsevier, 2007. p. 1–29.
52. VERONICA SALMASO. **Exploring protein flexibility during docking to investigate ligand-target recognition**. [s.l: s.n.].
53. VORATAS KACHITVICHYANUKUL. Comparison of Three Evolutionary Algorithms: **Industrial Engineering & Management Systems**, v. 11, n. 3, p. 215–223, 2012.
54. WANG, D.; TAN, D.; LIU, L. Particle swarm optimization algorithm: an overview. **Soft Computing**, v. 22, n. 2, p. 387–408, 17 jan. 2018.
55. YORIYAZ, H. Monte Carlo Method : principles and applications in Medical Physics. **Revista Brasileira de Física Médica**, v. 3, n. 1, p. 141–149, 2009.
56. ZHOU, K. et al. A Parallel Method with Hybrid Algorithms for Mixed Integer Nonlinear Programming. In: **Computer Aided Chemical Engineering**. [s.l.] Elsevier B.V., 2013. v. 32p. 271–276.
57. ZSOLDOS, Z. et al. eHiTS: A new fast, exhaustive flexible ligand docking system. **Journal of Molecular Graphics and Modelling**, v. 26, n. 1, p. 198–212, 2007.

ESTUDO DE PROTEÍNAS E AVALIAÇÃO DA QUALIDADE PELO GRÁFICO RAMANCHANDRAN

Neidelênio Baltazar Soares

Universidade da Integração Internacional da
Lusofonia Afro-Brasileira, Acarape-Ceará

Hélcio Silva dos Santos

Universidade Estadual Vale do Acaraú – Sobral-Ceará

Emanuelle Machado Marinho

Universidade Estadual do Ceará – Limoeiro do Norte-Ceará

Aluisio Marques da Fonseca

Universidade da Integração Internacional da
Lusofonia Afro-Brasileira, Acarape-Ceará

Resumo

As Proteínas são cadeias de aminoácidos com capacidade de regular a maioria das atividades biológicas nos organismos. A pesquisa de proteínas em bioinformática surge com a necessidade de dar resposta aos problemas de Predição da Estrutura de Proteínas (PPEP), pois conhecendo a estrutura de uma proteína é possível identificar o tipo de atividade que a proteína realiza na célula e criar drogas para combater, pois a maioria delas apresenta diferentes níveis de organização. Este trabalho apresenta a forma como as proteínas devem ser tratadas/preparadas, bem como a forma de analisar a sua qualidade por meio do gráfico Ramachandran. Para o preparo foram utilizados 2 softwares livres: Discovery Studio e AutoDock Tools e para análise da qualidade foi utilizado programa PROCHECK. A proteína utilizada como exemplo nesse estudo, demonstrou valores satisfatórios de qualidade quando submetido a análise por meio do gráfico Ramachandran apresentado

a porcentagem de resíduos favoráveis em 93, 2 % correspondente aos ângulos torsionais, que confere uma qualidade excelente.

Palavras Chaves: Proteínas. Qualidade. Ramachandran.Preparação.

1. Introdução

O Projeto Genoma Humano (PGH) teve como objetivo o sequenciamento dos 3,1 bilhões de bases nitrogenadas do genoma humano. O genoma é o conjunto de DNA de um ser vivo, e o DNA é formado pela ligação sequencial de moléculas denominadas nucleotídeos⁸. Uma riqueza de dados biológicos foram gerados pelo projeto do genoma humano e projetos de sequenciamento em outros organismos¹.

Devido a grande demanda dos dados biológicos se faz necessária a implementação de um mecanismo que possibilita a sua análise, a bioinformática, que tem como foco principal a obtenção e orientação de dados biológicos e a organização destas informações em bases de dados com vista à elaboração de métodos para a integração de informações relacionadas a partir de diferentes fontes. E, para atingir o objetivo, algoritmos são desenvolvidos para acelerar e reforçar a investigação biológica¹⁵.

Quanto às suas aplicações, a bioinformática pode ser dividida em três vertentes principais que correspondem à análise de sequências de macro (DNA e proteínas) e micromoléculas (transcritos e metabólitos), análises funcionais e análises estruturais. As análises de sequências de macromoléculas englobam comparações genômicas, as análises funcionais correspondem à predição de vias metabólicas, e a identificação de sítios ativos.

Análises estruturais compreendem predições de estruturas secundárias, terciárias de macromoléculas, comparação e classificação de proteínas, DNA e RNA²².

Um dos maiores objetivos da bioinformática estrutural é compreender a relação entre a estrutura primária, a sequência de aminoácidos das proteínas e sua estrutura tridimensional, uma vez que as análises estruturais e teóricas de proteínas são centrais para a compreensão de mecanismos moleculares complexos e são fundamentais para o processo de descoberta de drogas²³.

Proteínas são cadeias de aminoácidos que regulam diversas atividades em todos os organismos conhecidos. O Problema da Predição da Estrutura de Proteínas (PPEP) é um tópico importante de pesquisa em Bioinformática, cuja solução ainda está em aberto. Conhecer a estrutura de uma proteína pode ajudar a descobrir que tipo de atividade a proteína realiza na célula e criar drogas para combater doenças¹¹.

As proteínas apresentam diferentes níveis de organização e elas podem ser primárias, secundárias, terciárias e quaternárias¹². A organização das proteínas, permite a predição de estrutura de proteínas (PSP, do inglês *Protein Structure Prediction*), que é atualmente uma das áreas mais pesquisadas na Biologia Molecular Computacional em diferentes arranjos, pois nenhum método computacional simples e preciso para a determinação da estrutura tridimensional é conhecido até o momento⁶.

Os métodos computacionais para Predição das Estrutura das Proteínas (PSP) são divididos basicamente em duas linhas: métodos baseados em conhecimentos sobre o domínio do problema e métodos de predição por primeiros princípios

ou *Ab initio*. Os métodos baseados em conhecimento têm sido amplamente empregados, porém possuem uma série de limitações: (1) dependem de grandes bancos de dados de estruturas previamente determinadas, (2) demandam constante atualização desses bancos, (3) requerem algoritmos eficientes para alinhamento de múltiplas sequências, (4) necessitam de métodos computacionais eficientes que avaliem a similaridade entre estruturas^{6;24}.

As predições *Ab initio* não apresentam as restrições, verificadas nos métodos baseados em conhecimentos sobre o domínio do problema, eles dependem apenas da sequência de aminoácidos¹⁰.

A realização da modelagem das estruturas proteicas requer que as mesmas passem por um processo de validação dos modelos através de análises das suas qualidades, principalmente se diferentes orientações referentes aos resíduos do sítio ligante forem encontradas para os diferentes modelos de proteínas gerados^{18;21}.

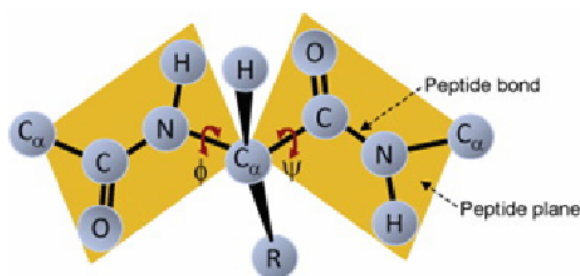
A análise de qualidade das proteínas pode ser feito por meio de diversos softwares que são utilizados com a finalidade de avaliar vários parâmetros estereoquímicos como é o caso do PROCHECK, que também serve para avaliar os comprimentos de ligação, os ângulos planos, a planaridade dos anéis de cadeias laterais, a quiralidade, as conformações das cadeias laterais, a planaridade das ligações peptídicas, os ângulos torcionais da cadeia principal e das cadeias laterais, o impedimento estérico entre pares de átomos não-ligados e a qualidade do gráfico de Ramachandran^{9;14}.

2. Fundamentação teórica

2.1 Aminoácidos e Ligações Peptídicas

A estrutura de um aminoácido é feita por ligações peptídicas, que são ligações amida entre os grupos $-NH_2$ e $-COOH$ de aminoácidos vizinhos. A ligação peptídica ($C\equiv N$) tem um caráter de ligação dupla parcial, isso faz com que ele tenha um caráter rígido no plano, impedindo a rotação para girar. Ligações peptídicas são consideradas ligações *trans* - ou seja, o oxigênio da carbonila e o hidrogênio da amida estão na posição *trans*. O $N-C_\alpha$ e o $C_\alpha-C$ - As ligações C não são rígidas e podem girar livremente, sendo apenas limitadas pelo tamanho e caráter dos grupos-R. O ângulo de rotação (também denominado ângulo de torção ou ângulo diedro) em torno da ligação $N-C_\alpha$ é denominado phi (ϕ) e aquele em torno da ligação $C_\alpha-C$ é denominado psi (ψ). A figura (1) demonstra a estrutura da ligação peptídica.

Figura 1. Estrutura da Ligação peptídica



Fonte: (CHOUDHURI, 2014)⁵

2.2 Gráfico Ramachandran

O gráfico de Ramachandran é particularmente útil porque ele define os resíduos que se encontram nas regiões energeticamente mais favoráveis e desfavoráveis e orienta a avaliação da qualidade de modelos teóricos ou experimentais de proteínas¹⁷.

No mapa Ramachandran representado na Figura 2 apresenta os ângulos torsionais correspondem ao ângulo phi (ângulo entre o carbono- α e o nitrogênio do grupamento amino) e psi (ângulo formado entre o carbono- α e o carbono do grupamento carboxi) de cada aminoácido, que são plotados no gráfico de Ramachandran, o qual fornece informações acerca da qualidade do arranjo estrutural²⁰.

3. Metodologia

3.1. Banco de Dados para Obter as Estruturas das proteínas

Para obter a proteína estrutural para estudos de modelagem existem atualmente vários repertórios e bancos de dados disponíveis onde as mesmas encontram-se depositadas por meio de códigos. O quadro 1 demonstra alguns dos mais importantes repertórios para obter as proteínas.

Quadro 1. Principais Bancos de dados disponíveis para obtenção das proteínas

Nome	Link de Acesso	Referências
PDB	https://www.rcsb.org/	(BERMAN, 2000)
GenBank	https://www.ncbi.nlm.nih.gov/protein/?term=5m5b	(BENSON, 2000)
MODBASE	https://modbase.compbio.ucsf.edu/modbase-cgi/index.cgi	(SANCHEZ, 1999)
CATH	https://www.cathdb.info/	(ORENGO et al., 1999)
SWISS-PROT	https://www.uniprot.org/	(GASTEIGER; JUNG; BAIROCH, 2001)

Fonte: Elaborado pelo autor (2021)

3.2 Preparo de Proteínas para Realização de Docking

- **Utilizando Discovery Studio (1° Passo do preparo de proteína)**
 - a) Retirar a água clicando em “Water” e pressionar a tecla delete;
 - b) Observar se existe metal (retirar);
 - c) Eliminar outras partes da estrutura caso seja um dímero, trímero, tetrâmero, ou polímero, deixar apenas o monômero “A”
 - d) Se for uma estrutura com ligante, deixar onde contém o ligante principal;
 - e) Eliminar os demais íons;
 - f) Realizar a Determinação da região do sítio ativo clicando em Ligands group, e selecionar “Ligand_1”
 - g) Na barra de menu clique em “Receptor-ligands interations”
 - h) Em seguida clique em “Define and edit binding site” e clique em “From current Seletion”;
 - i) Em seguida clique em “SBD_ site _sphere” e expanda a região se necessário;
 - j) No botão direito selecionando clique em “SBD_site_sphere” e clique em “atributes of SBD_site_sphere”;
 - k) Em seguida salva os dados da grade do sítio ativo em um documentp txt ou bloco de notas com as cordenadas x,y,z;
 - l) Com o botão direito do mouse clique em SBD_site_sphere e escolha a opção “HIDE”
 - m) Para finalizar delete o Ligante e salva a molécula em PDB.

- **Utilizando o Autodock Tolls (2º Passo do preparo da proteína)**
 - a) Abra a proteína no formato PDB recém preparada;
 - b) Clique seguindo a sequência “Edit>Misc>Cheking for Missing atoms>select all resides>desmiss
 - c) Em seguida clique em Edit>Misc>Repair missing atoms
 - d) Salve a quantidade corrigida “x” sets>Dismiss
 - e) Clique para discelecionar a proteína clicando no icon amarelo junto a nome da proteína;
 - f) Adicione Hidrogénios polares clicando em Edit>Hydrogen>Add polar only>ok
 - g) Adicione as cargas de kallman clicando em Edit>Charges> Add kallman charges>ok
 - h) Agora faça as correções das cargas Clicando em edit>Charges>checking total on resides >spread charge/ ok

3.3 Aplicativos Utilizado para realização de analise de Modelos proteicos

PSIPRED

O software PSIPRED é um programa que permite a predição da estrutura secundária de proteína baseada em matrizes de pontuação especificas de posição⁴.

PATCHDOCK

O PATCHDOCK é um servidor da web disponíveis gratuitamente para encaixe molecular. Que permite a avaliação da interação entre duas moléculas realiza a previsão da estrutura dos complexos proteína-proteína e proteína-molécula pequena¹⁹.

PROCHEK

O programa PROCHECK fornece uma verificação detalhada da estereoquímica da estrutura de uma determinada proteína. Os resultados que advem dele incluem vários gráficos em formato PostScript e uma lista abrangente de resíduo por resíduo. Estes fornecem uma avaliação da qualidade geral da estrutura em comparação com estruturas bem refinadas da mesma resolução e também destacam regiões que podem precisar de investigação adicional. Os programas PROCHECK são úteis para avaliar a qualidade não só de estruturas de proteínas em processo de solução, mas também de estruturas existentes e modeladas em estruturas conhecidas⁹.

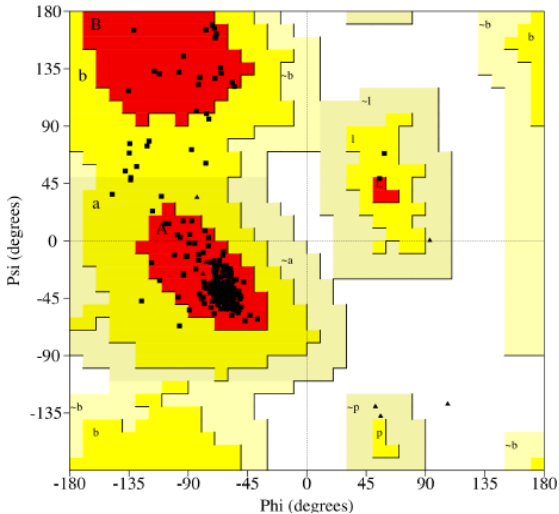
4. Resultados

4.1 Gerando Resultados da Qualidade das proteínas Pelo Ramachandran

Para realização da qualidade de determinada proteína, a mesma é submetida à análise através do gráfico de Ramachandran utilizando-se o software PROCHECK¹ de plataforma on-line. Como exemplo nesse estudo foi utilizado a proteína 4m4b, depositada pelo método de difração de raios X com a resolução de 2,00 Å. A distribuição dos resíduos da proteína pode ser vista no gráfico de Ramachandran (fig. 2), e os valores para cada região do gráfico estão dispostos no quadro 2.

¹ <https://saves.mbi.ucla.eu/>

Figura 2. Gráfico de Ramachandran da proteína 4m4b. Os resíduos localizados na região vermelha são considerado os mais favorável, os resíduos localizados na região marrom são os adicionais e considerado permissíveis, já os localizados na região begees.



Fonte: Elaborado pelo autor.

O quadro 2, que segue, descreve os resultados encontrado para proteína 4m4b.

Quadro 2. Resultados da análise de Qualidade da Proteína 4M4B.

Regiões	Siglas	Nº de resíduos	%
Mais favoráveis	[A,B,L]	229	93,2 %
Adicionais permissíveis	[a,b,l,p]	19	7,7 %
Permissíveis	[~a,~b,~l,~p]	0	0,0 %
Não permissíveis	xx	0	0,0 %

Fonte: Elaborado pelo autor (2021).

5. Considerações Finais

Os resultados das análises do gráfico de Ramachandran realizado nesse estudo e utilizando como exemplo, mostraram que a qualidade da estrutura depositada no PDB (Protein Data Bank), com o código 4m4b foi exatamente 93, 2 %, a porcentagem é relativa aos valores dos ângulos tornionais localizado na região favoravel. E desta forma, o valor do percentual confere que a qualidade da proteína é excelente por se aproximarem a 90%.

Referências

1. BAYAT, A. Science, medicine, and the future: Bioinformatics. **BMJ**, v. 324, n. 7344, p. 1018–1022, 27 abr. 2002.
2. BENSON, D. A. GenBank. **Nucleic Acids Research**, v. 28, n. 1, p. 15–18, 1 jan. 2000.
3. BERMAN, H. M. The Protein Data Bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235–242, 1 jan. 2000.
4. BUCHAN, D. W. A.; JONES, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. **Nucleic Acids Research**, v. 47, n. W1, p. W402–W407, 2 jul. 2019.
5. CHOUDHURI, S. Additional Bioinformatic Analyses Involving Protein Sequences. In: **Bioinformatics for Beginners**. [s.l.] Elsevier, 2014. p. 183–207.
6. GABRIEL, P. H. R.; MELO, V. V. DE; DELBEM, A. C. B. Algoritmos evolutivos e modelo HP para predição de estruturas de proteínas. **Sba: Controle & Automação Sociedade Brasileira de Automatica**, v. 23, n. 1, p. 25–37, fev. 2012.
7. GASTEIGER, E.; JUNG, E.; BAIROCH, A. SWISS-PROT: connecting biomolecular knowledge via a protein database. **Current issues in molecular biology**, v. 3, n. 3, p. 47–55, jul. 2001.

8. GÓES, A. C. DE S.; OLIVEIRA, B. V. X. DE. Projeto Genoma Humano: um retrato da construção do conhecimento científico sob a ótica da revista Ciência Hoje. **Ciência & Educação (Bauru)**, v. 20, n. 3, p. 561–577, set. 2014.
9. LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, n. 2, p. 283–291, 1 abr. 1993.
10. LOPES, H. S. Evolutionary Algorithms for the Protein Folding Problem: A Review and Current Trends. In: **Computational Intelligence in Biomedicine and Bioinformatics**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 297–315.
11. LOPES, J. N.; VENSKE, S. M. **Predição da Estrutura de Proteínas Utilizando Algoritmo Evolutivo Adaptativo**. Anais do 12. Congresso Brasileiro de Inteligência Computacional. **Anais... ABRICOM**, 13 out. 2015Disponível em: <http://abricom.org.br/eventos/cbic_2015/cbic2015_submission_128>
12. NELSON, D. L.; COX, M. M.; LEHNINGER, A. L. **Princípios de bioquímica de Lehninger**. 6ª ed. Porto Alegre: [s.n.].
13. ORENGO, C. A. et al. The CATH Database provides insights into protein structure/function relationships. **Nucleic Acids Research**, v. 27, n. 1, p. 275–279, 1 jan. 1999.
14. RAMACHANDRAN, G. N.; SASISEKHARAN, V. Conformation of Polypeptides and Proteins. In: [s.l: s.n.]. p. 283–437.
15. RISLER, J.-L. Developing Bioinformatics Computer Skills. **Computers & Chemistry**, v. 26, n. 5, p. 553–555, jul. 2002.
16. SANCHEZ, R. ModBase: A database of comparative protein structure models. **Bioinformatics**, v. 15, n. 12, p. 1060–1061, 1 dez. 1999.
17. SANTOS FILHO, O. A.; ALENCASTRO, R. B. DE. Modelagem de proteínas por homologia. **Química Nova**, v. 26, n. 2, p. 253–259, mar. 2003.
18. SCHAFFERHANS, A.; KLEBE, G. Docking ligands onto binding site representations derived from proteins built by homology modelling. **Journal of Molecular Biology**, v. 307, n. 1, p. 407–427, mar. 2001.

19. SCHNEIDMAN-DUHOVNY, D. et al. PatchDock and SymDock: servers for rigid and symmetric docking. **Nucleic Acids Research**, v. 33, n. Web Server, p. W363–W367, 1 jul. 2005.
20. SHEIK, S. S. et al. Ramachandran plot on the web. **Bioinformatics**, v. 18, n. 11, p. 1548–1549, 1 nov. 2002.
21. SILVA, V. B.; SILVA, C. H. T. . Homology Modeling of Target Proteins. **Revista Eletrônica de Farmácia**, v. IV, n. 1, p. 15–26, 2007.
22. TSOKA, S.; OUZOUNIS, C. A. Recent developments and future directions in computational genomics. **FEBS Letters**, v. 480, n. 1, p. 42–48, 25 ago. 2000.
23. VILLOUTREIX, B. O. Structural Bioinformatics: Methods, Concepts and Applications to Blood Coagulation Proteins. **Current Protein and Peptide Science**, v. 3, n. 3, p. 341–364, 1 jun. 2002.
24. XIONG, J. **Essential Bioinformatics**. 1^a ed. New York: Cambridge University Press, 2006.

PROPRIEDADES DE ADME-T, UMA ABORDAGEM INTRODUTÓRIA

Eduardo Gaieta

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Matheus Nunes da Rocha

Universidade Estadual do Ceará – Limoeiro do Norte-Ceará

Francisco Wagner de Queiroz Almeida Neto

Universidade Federal do Ceará

Aluisio Marques da Fonseca

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Resumo

Compreender o comportamento que certo candidato a droga (CD) vai ter em um organismo biológico é um fator crucial no processo de pesquisa e desenvolvimento de novos fármacos, que vai desde reduzir a proporção de falhas, reduzir os gastos orçamentais, ao auxílio na descoberta de novas entidades químicas. O fato de que cada composto químico tem a sua propriedade, significa que cada um terá um impacto diferente em nosso organismo. Deste modo, prever como um CD vai ser administrado e absorvido em nosso organismo, como será Distribuído e Metabolizado em nosso corpo, o tempo que vai precisar para desempenhar a ação terapêutica desejada até ser eliminado, e o não menos importante, se durante este processo o nosso CD vai expor o nosso organismo a efeitos prejudiciais, se tornou uma abordagem

valiosa e indispensável para o sucesso no processo de P&D de novas drogas. Este capítulo foca em fazer uma abordagem conceitual introdutória sobre os parâmetros de absorção, distribuição, metabolismo, excreção e toxicidade (ADME-T) destacando a sua importância no processo P&D de novos fármacos, e fazer uma breve exemplo de como predição dessas propriedades em uma molécula.

Palavras-chave: Absorção; Distribuição; Metabolismo; Excreção; Toxicidade.

1. Introdução

O processo de pesquisa e desenvolvimento (P&D) de novas drogas é um processo complexo, desafiador e dispendioso para a indústria farmacêutica que envolve desde as fases pré-clínica (descoberta ou pesquisa básica), clínica (desenvolvimento) e comercial (quando aprovado para o consumo). Para se ter uma ideia, esse processo pode demorar entre 10 a 15 anos^{19,16}, e ainda assim muitos candidatos a drogas não chegam a se tornar um medicamento. Note-se que cerca 50% das falhas neste processo de P&D de uma nova droga são atribuídas a uma fraca exibição das suas propriedades de absorção, distribuição, metabolismo, excreção e toxicidade (ADME-T)⁹.

Entende-se que o estudo de uma possível droga se torna promissor e seguro ao exibir uma combinação bem ajustada de farmacodinâmica (seletividade, afinidade, potência) e farmacocinética (biodisponibilidade, absorção, toxidez)¹³, o que nos leva a compreender que a determinação e a otimização prévia dessas propriedades em novas entidades químicas (NCEs, do inglês *New Chemical Entities*)³⁷ desempenham papéis vitais em cada estágio da descoberta e desenvolvimento de novos fár-

macos, sendo crucial para eficácia do efeito terapêutico para o qual está sendo desenvolvido. Foi considerando essas propriedades que se evoluíram os conceitos de semelhança com droga e drogabilidade.

2. Revisão literária

O conceito de drogabilidade compreende a capacidade de interação de uma estrutura se ligar à uma molécula semelhante a uma droga que apresenta atividade e afinidade terapêuticamente útil^{30,1}. E a semelhança com droga pode ser definida como aqueles compostos que apresentam propriedades de ADME-T consistentes com a maioria dos medicamentos conhecidos, e que essas propriedades sejam aceitáveis o suficiente para sobreviver aos ensaios clínicos da primeira fase³⁴. Estes conceitos também foram usados por Lipinski e colaboradores para desenvolver critérios decisivos para a identificação de compostos semelhantes a drogas, também conhecida como a regra dos cinco.

A regra de Lipinski, ou Regra dos Cinco, foi elaborada por Lipinski e colaboradores em 1997, e é baseada em critérios físico-químicos que estabelecem as características que um fármaco deve ter para poder ser absorvido (administração oral), distribuído, metabolizado e a sua estabilidade para que possa ser mantido no organismo a tempo suficiente para desempenhar a ação desejada e depois ser excretado^{15,39}.

Esses critérios definem que uma NCE deve ter:

- coeficiente de partição octanol / água ($A \log P$) ≤ 5 ;
- peso molecular (MW) ≤ 500 g/mol;

- número de doadores de ligação de hidrogênio (HBDs) (soma dos grupos OH e NH) ≤ 5 ;
- número de aceitadores de ligação de hidrogênio (HBAs) (soma dos átomos de N e O) ≤ 10 ;

Extensões.

- Número de ligações rotacionáveis ≤ 10 ;
- Área de superfície polar (PSA) $< 140 \text{ \AA}$.

Note que: A regra dos cinco se aplica apenas a compostos administrados por via oral, não se aplicando a produtos naturais ou substratos de transportadores biológicos^{1;18}.

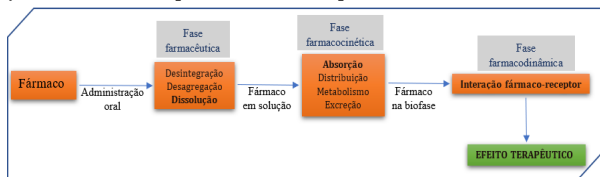
De acordo com Piccirillo e Do Amaral, (2018) e Bennet *et al*, (2016), Lipinski sugeriu que compostos que não obedecessem a dois ou mais desses parâmetros são os mais susceptíveis a apresentar baixa capacidade de solubilidade e/ou permeabilidade.

O percurso ou os aspectos de um fármaco em nosso organismo podem representados por: farmacêutica, farmacocinética e farmacodinâmica (fig. 1).

- A propriedade farmacêutica ou biofarmacêutica corresponde a administração do fármaco na forma farmacêutica liberando-o para absorção, incluindo a libertação e a dissolução^{6;33}.
- A farmacocinética refere-se aos aspetos relacionados a absorção, distribuição, metabolismo, excreção e toxicidade (ADME-T) do fármaco^{6;8}.
- A farmacodinâmica diz respeito a interação do fármaco no local de ação pretendido, para que se tenha o efeito terapêutico desejado, que também está relacionada com as características de potência, afinidade, seletividade da molécula em estudo^{16;33}.

De acordo com os autores^{6,36} as previsões das propriedades de ADMET podem ser feitos em estudos *in silico*, *in vitro* e *in vivo*.

Figura.1 - Etapas do processamento de um fármaco, desde a administração oral, até o desempenho do efeito terapêutico.



Fonte: Elaborado pelo autor.

2.1 ADMET

O efeito químico e terapêutico que um determinado fármaco causa em um organismo é analisado pelo conceito de ADME-T. Este conceito descreve o potencial impacto que um fármaco tem quando após a sua administração é **A**bsorvido, **D**istribuído para o seu local de ação, **M**etabolizado, **E**xcretado após desempenhar a ação terapêutica desejada, e os efeitos **T**oxicológicos gerados após o seu processamento em um sistema vivo no contexto da biologia celular e bioquímica^{26,29}.

2.2 Absorção

O processo em que um fármaco sai do local de administração, atravessa as barreiras biológicas e chega até o sistema circulatório, é caracterizado como absorção^{32,35}. Os locais mais importantes de absorção é o trato gastrointestinal e a pele²².

A absorção depende de movimentos transmembranares, do fluxo sanguíneo no sítio absorptivo, extensão e espessura da superfície de absorção, formas farmacêuticas e vias de administração escolhidas (seja ela, oral; bucal; sublingual; dental; intracanal; retal; parental e etc.)²¹. No entanto, durante o processo de P&D de novos fármacos, a administração via oral é a mais preferida e conveniente²², mas independentemente da via de administração, o fármaco precisa passar pelas membranas intestinais para chegar até a circulação sistêmica e, entre alguns mecanismos responsáveis por este exercício, destacam-se os meios de difusão passiva e os mecanismos de transporte mediados por portadores ativos^{7;20}. É importante ter em conta, que nem todos fármacos passam por essa etapa de absorção no organismo, pois existem aqueles que são administrados por injeção direta na circulação sistêmica, tendo assim biodisponibilidade absoluta²¹. A biodisponibilidade, pode ser entendida como a quantidade da dose administrada que chega até a circulação sistêmica para exercer a ação desejada³³.

Entre vários fatores que afetam a absorção intestinal de um fármaco, destacam-se a solubilidade e a permeabilidade²⁸. Uma das características que influencia diretamente na solubilidade de uma droga é o logaritmo negativo de base 10 da constante de dissociação ácida (pK_a)². O pK_a de uma droga é associado ao fato de que muitas drogas são ácidas ou bases fracas, o que dizer que elas podem estar presentes em solução nas formas ionizadas e não ionizadas. As moléculas não ionizadas geralmente apresentam um maior grau de lipossolubilidade e com isso atravessam as membranas por difusão passiva. As moléculas ionizadas apresentam um baixo grau de solubilidade lipídica portanto, não podem atravessar a membrana celular⁷. O pK_a

de uma droga pode ser expresso em decorrência de sua natureza química e do pH do meio. Um baixo pH no estomago é favorável para a absorção de ácidos fracos, enquanto que as bases fracas apresentam um maior grau de absorção no intestino²².

A lipofilicidade (Log P) é a principal força motriz da permeabilidade²⁸, que pode ser definida como a afinidade que um substrato exibe para a bicamada lipídica. A lipofilicidade pode ser estimada usando o coeficiente de partição (P) em um sistema bifásico (1 octanol / água) ou a partir de membranas artificiais. Em geral, as drogas que apresentam um alto valor de coeficiente de partição, tendem a exibir maior taxa de permeabilidade e maior perfil de biodisponibilidade, no entanto, estudos apontam que drogas altamente lipofílicas (Log P > 5), exibem má absorção^{22;4}. A exibição de valores de Log P negativos podem ser indicativos de hidrofílicidade e como também podem estar associados à má absorção²².

2.3 Distribuição

A simples presença de uma droga no sistema sanguíneo não é um indicativo de ocorrência da ação terapêutica. Para que uma droga desempenhe a ação terapêutica desejada, é necessário que após a administração e absorção seja distribuída de modo a alcançar o seu órgão-alvo. Este processo de distribuição acontece quando a droga reversivelmente deixa o leito vascular e entra no interstício, e então nas células dos tecidos por intermédio do sistema circulatório e sistema linfático^{14;38}. O processo de distribuição depende de fatores como volume de distribuição, permeabilidade capilar, fluxo sanguíneo, e lipofilicidade³⁸.

2.4 Metabolismo

Quando uma droga é distribuída em nosso organismo, ela está sujeita a ter vários destinos como retenção inalterada, transformação química espontânea, eliminação inalterada ou metabolismo enzimático. Mas, dentre eles, o mais predominante é o metabolismo enzimático, também conhecido como biotransformação⁷.

Durante o processo de biotransformação a droga sofre modificações químicas mediadas por enzimas onde são transformadas em um ou mais novos compostos metabólitos (que também podem ser metabolizados) diferentes do original administrado²⁰.

As reações que ocorrem durante o metabolismo podem ser classificadas em reações de oxidação/redução também conhecida como Fase I e reações de conjugação/hidrólise também conhecida como fase II¹⁴. A metabolização sistemática de uma droga pode acontecer em diversos órgãos como os do sistema gastrintestinal, os pulmões, os rins, a pele e o fígado. Entretanto, por conter uma maior diversidade e quantidade de enzimas metabólicas, o fígado é o principal órgão em que ocorre a biotransformação dos fármacos.

É durante a fase I que ocorrem as reações de oxidação, redução ou hidrólise que geralmente originam metabólitos mais hidrofílicos do que as drogas originalmente administradas, isso, com a introdução de um grupo funcional. As isoformas da família do citocromo P450 (CYP ou CYP450) são as principais enzimas para a oxidação na fase I²⁵. Na sequência, as reações da fase II que englobam a adição dos grupos glicuronato, sulfato, glutatona e acetato, conjugam ou hidrolisam uma droga transformando-as em conjugados de alta polaridade de modo a au-

mentar a sua solubilidade, facilitando assim a sua excreção pela urina ou na bile. As enzimas de maior destaque na fase II são estrogênio sulfotransferase (EST), fenol sulfotranferase (PST), glutationa-S-transferase (GST) e glucuronosil transferase dependente de UDP (UGT)^{36,21}. Entre as isoenzimas do CYP450, de maiores destaques estão as CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP2E1 e CYP3A4 como as principais correspondendo cerca de 90% do metabolismo oxidativo dos fármacos disponíveis no mercado, e sendo a CYP3A4 responsável por cerca de 50% do metabolismo oxidativo das drogas³.

O metabolismo da droga é uma chave determinante de várias propriedades importantes como a estabilidade metabólica, interações medicamentosas, toxicidade de drogas²¹.

2.5 Excreção

Os processos reacionários que ocorrem na biotransformação tendem a transformar as drogas e seus metabolitos suficientemente polares para serem eliminados do organismo para o meio externo. A excreção ocorre através dos emunctorios que incluem a secreção biliar, suor, rim, saliva, pulmões, lagrimas, fezes e outros sítios quantitativamente menos importantes. A eliminação das drogas e seus metabolitos ocorrem por sua maioria por meio renal e biliar¹¹.

A excreção renal das drogas envolve os processos de: A) filtração glomerular, na qual não é influenciado pela lipossolubilidade e pelo pH, somente pela velocidade de filtração glomerular (VFG); B) Secreção tubular proximal onde a secreção ocorre primariamente nos túbulos proximais por dois meca-

nismos de transporte ativo que exigem energia: um para ânions (p. Ex., formas desprotonadas de ácidos fracos) e outro para cátions (p. Ex., formas protonadas de bases fracas); C) Reabsorção tubular distal, neste processo enquanto o fármaco se desloca em direção ao túbulo contorcido distal, sua concentração aumenta e excede à do espaço perivascular. Este processo é influenciado pelo pH³⁸.

O aumento da taxa de filtração glomerular, diminuição da ligação e do fluxo sanguíneo às proteínas plasmáticas, causam uma excreção mais rápida dos fármacos¹⁴.

A partir do fígado, as drogas na forma de conjugados são levadas pela bile ao intestino, onde podem ser reabsorvidos pelo circuito êntero-hepático após ficar na forma livre ou excretados através das fezes. Muitas das vezes está reabsorção pode prolongar seus efeitos, deste modo se excretam fármacos de alto peso molecular, os muito polares e aqueles que são ativamente englobados em micelas de sais biliares, colesterol e fosfolípidios.

A excreção depuração corporal total e a meia-vida do fármaco são variáveis importantes para otimizar o tratamento medicamentoso e minimizar a toxicidade. A depuração total (CL_T), do inglês Total clearance é a soma de todos os mecanismos de eliminação potenciais, sejam eles a depuração metabólica hepática (CL_H), depuração renal (CL_R), depuração biliar (CL_B), e depuração metabólica extra-hepática (CL_{EH}), que removem irreversivelmente o composto da circulação sistêmica²⁴.

O metabolismo e a excreção estão relacionados com a meia-vida, uma vez que ao diminuir a concentração da droga ativa no sangue, reduzem o tempo na qual uma droga é capaz de atuar sobre um órgão alvo. A meia-vida de eliminação de

uma droga é definida como o tempo durante o qual a concentração do fármaco no plasma diminui para a metade de seu valor original¹⁴.

2.6 Toxicidade

A toxicologia de uma droga enfoca os efeitos prejudiciais da exposição que um novo fármaco tem em todo o organismo, um órgão ou subestrutura de um sistema biológico (o corpo humano por exemplo) e é responsável pelo fracasso de muitos compostos durante o processo de P&D de uma nova droga e pela retirada de um número expressivo de drogas do mercado após de aprovados^{29,36}.

A toxicidade é indiscutivelmente a propriedade do medicamento mais difícil de prever adequadamente, porque pode ser específica da espécie, do órgão e pode envolver vários fatores, dentre eles citotoxicidade, cardiotoxicidade, toxicidade hepática, e toxicidade genética²², a qual faz referência a carcinógenos, ou agentes provocam câncer ou aumentam o seu grau de incidência. Estes podem ser divididos em carcinógenos mutagênicos e não mutagênicos^{22,17}. É amplamente acordado que a predição das propriedades de absorção, distribuição, metabolismo, excreção e toxicidade, ajudam a prevenir de falhas de NCEs durante o P&D, e melhorar na projeção de uma droga com maior eficácia, por esta razão a abordagem de ADME-T podem ser feitos em modelos *in silico*, *in vitro* e *in vivo*^{12,23}.

Quadro 1. Parâmetros de absorção, distribuição, metabolismo, excreção de drogas e toxicidade (ADME-T) que podem ser determinadas *in silico*, *in vitro* e *in vivo*^{22;10}.

Processo	Determinantes	Parâmetros
Absorção	<ul style="list-style-type: none"> Solubilidade Permeabilidade Extração de primeira passagem Transportadores 	<ul style="list-style-type: none"> Estabilidade química Solubilidade (SO) pKa, HBD and HBA PSA Log P or Log D P_{app} Relação efluxo / influxo (Efflux/influx ratio) Estabilidade metabólica Absorção intestinal humana (HIA) Permeabilidade Caco-2
Distribuição	<ul style="list-style-type: none"> Permeabilidade Transportadores Equilíbrio de partição Ligação a proteínas plasmáticas 	<ul style="list-style-type: none"> Relação efluxo / influxo (Efflux/influx ratio) Estabilidade metabólica Log P or Log D f_u (fração não ligada) Substrato de glicoproteína P, inibidor Penetração da barreira hematoencefálica (BBB) Barreira sangue-placenta (BPB)
Metabolismo	<ul style="list-style-type: none"> Permeabilidade Interação com enzimas Ligação às proteínas plasmáticas 	<ul style="list-style-type: none"> Log P or Log D Liberação intrínseca Identificação de DME e porcentagem contribuição para o metabolismo (fm) Identificação de metabólito Substrato e inibidor do citocromo P450 (CYP450)
Excreção	<ul style="list-style-type: none"> Permeabilidade Interação com transportadores Ligação às proteínas plasmáticas 	<ul style="list-style-type: none"> Log P or Log D pKa Relação efluxo / influxo (Efflux/influx ratio) Meio tempo (t_{1/2}) Depuração renal
Toxicidade	<ul style="list-style-type: none"> Citotoxicidade Cardiotoxicidade Toxicidade hepática Toxicidade genética 	<ul style="list-style-type: none"> Vitotox Ames clastogenicidade esteatose ligação de dofetilide Pg-P P450 Toxicidade aguda Lesões oculares e corrosão ocular

Fonte: Elaborado pelo autor.

HBD - doador de ligação de hidrogênio; HBA - aceitador de ligação de hidrogênio; log P ou D - lipofilicidade; PSA - área de superfície polar; pKa - constante de ionização, DME - enzimas metabolizadoras de drogas, P_{app} - permeabilidade aparente; P450 - citocromo CYP450; Pg-P - Glicoproteína P

Exemplo: Como avaliar o perfil de ADME-T de um composto?

Por razões como velocidade, custo e precisão, hoje em dia é reconhecido que a previsão do perfil ADME-T de uma NCE através de modelos *in silico* tem se tornado um método de escolha no estágio inicial de P&D de novas drogas e é uma mais valia, que em combinação com as previsões *in vitro* e *in vivo*, ajuda na obtenção de melhores resultados²⁷.

As predições *in silico* não substituem, nem desqualificam os testes *in vitro* e *in vivo*, mas os complementam, e devem andar em parceria para a busca do melhor resultado. Como exemplo, a previsão *in silico* de perfil ADME-T de um composto pode ser feita em algumas ferramentas digitais, como admetSAR, pkCSM, swissADME (quadro 2).

Quadro 2. Ferramentas de predição de perfil ADME-T.

Ferramenta	Licença	Endereço
admetSAR	Licença gratuito	http://lmmd.ecust.edu.cn/admetSar2/
pkCSM	Licença paga	http://biosig.unimelb.edu.au/pkCSM/
SwissADME	Licença gratuito	http://www.swissADME.ch/index.php

Fonte: Elaborado pelo autor.

3. Metodologia

3.1 Ferramenta

Foi utilizado o software SwissADME², ferramenta virtual de avaliação de perfil de ADME-T.

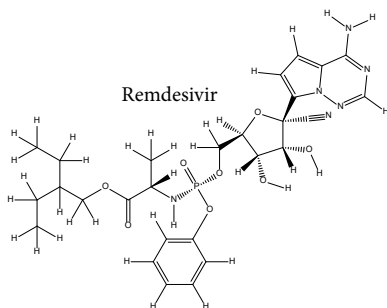
3.2 Escolha do fármaco para avaliação do perfil ADME-T *in silico*

O fármaco escolhido de forma aleatória foi o Remdesivir, a estrutura 2D e o código SMILES. Foram obtidos na plataforma PubChem³, um banco de dados de moléculas.

Código SMILE da estrutura:

```
CCC(CC)COC(=O)C(C)NP(=O)(OCC1C(C(C(O1)  
(C#N)C2=CC=C3N2N=CN=C3N)O)O)OC4=CC=CC=C4
```

Figura 2. Estrutura química do fármaco escolhido para avaliação do perfil ADME-T *in silico*.



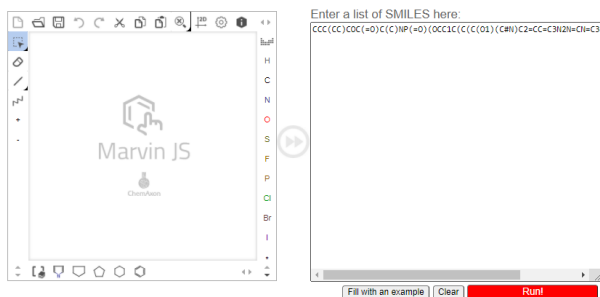
2 SwissADME <http://www.swissadme.ch/index.php>

3 PubChem <http://pubchem.ncbi.nlm.nih.gov/>

3.3 Avaliação ADME-T da estrutura

O uso da plataforma SwissADME é muito intuitivo. O código SMILES foi introduzido no campo próprio da plataforma (figura 3) e em seguida foi dada a ordem para executar a avaliação.

Figura 3. Campo de entrada de dados da plataforma SwissADME.

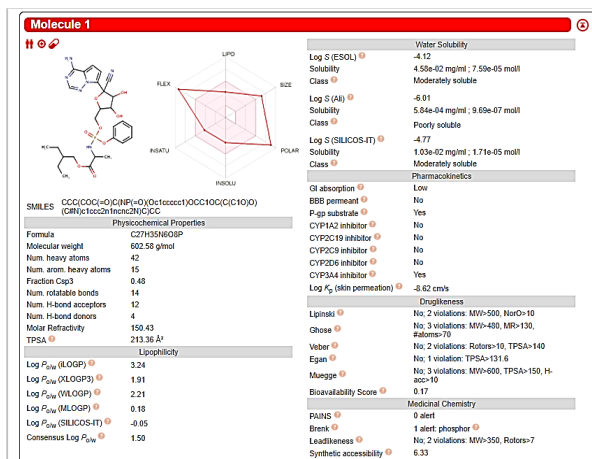


Fonte: Elaborado pelo autor.

4. Resultados

Os resultados da avaliação exibidos podem ser recolhidos da forma que o usuário desejar, como mostram os exemplos da tabela 1 e da figura 4. Obs.: São apenas exemplos, não são métodos padrão.

Figura 4. Resultados da avaliação das propriedades de ADME-T do Remdesivir.



Fonte: Elaborado pelo autor.

Tabela 1. Resultados da avaliação das propriedades de ADME-T do Remdesivir.

Ligantes	Propriedades Físico Químicas					Lipofilicidade		Semelhança com Drogas		Farmacocinéticas						
	mol. wt. g/mol	ROTB (n)	HBA (n)	HBD (n)	ESOL LogS	TPSA (Å)	ClogP (o/w)	Bioavailability Score	Lipinski	GLA	P-gp	CYP1A2	CYP2C19	CYP2C9	CYP2D6	CYP3A4
 REMDESIVIR	602.58	14	12	4	-4.12	213.36	1.50	0.17	No (2)	Low	Yes	No	No	No	No	Yes

Fonte: Elaborado pelo autor.

5. Considerações finais

Pensa-se que nesta abordagem introdutória sobre as propriedades de ADME-T foi apresentada os conceitos bases, para além de relevar seu grau de importância durante o processo de pesquisa e produção de novas drogas, mas isto é só a ponta do iceberg, o estudo das propriedades de ADME-T deve ser aprofundado para que se obtenham excelentes resultados desejados durante a avaliação de uma estrutura química.

Referências

1. AGONI C, OLOTU FA, RAMHARACK P, SOLIMAN ME. Druggability and drug-likeness concepts in drug design: are biomodelling and predictive tools having their say? *J Mol Model.* 2020;26(6):120. doi:10.1007/s00894-020-04385-6
2. ANDRADE EL, BENTO AF, CAVALLI J, et al. Non-clinical studies in the process of new drug development - Part II: Good laboratory practice, metabolism, pharmacokinetics, safety and dose translation to clinical studies. *Brazilian J Med Biol Res.* 2016;49(12):19. doi:10.1590/1414-431x20165646
3. BARREIRO EJ, FRAGA CAM. *Química Medicinal-: As Bases Moleculares Da Ação Dos Fármacos.* 3rd ed. (S.A AE, ed.). Artmed Editora; 2014.
4. _____. *Química Medicinal : As Bases Moleculares Da Ação Dos Fármacos.* 3ª. Artmed; 2015.
5. BENET LZ, HOSEY CM, URSU O, OPREA TI. BDDCS, the Rule of 5 and drugability. *Adv Drug Deliv Rev.* 2016;101:89-98. doi:10.1016/j.addr.2016.05.007
6. BRITO MA de. Pharmacokinetic study with computational tools in the medicinal chemistry course. *Brazilian J Pharm Sci.* 2011;47(4):797-805. doi:10.1590/S1984-82502011000400017.

7. CALDWELL J, GARDNER I, SWALES N. An Introduction to Drug Disposition: The Basic Principles of Absorption, Distribution, Metabolism, and Excretion. *Toxicol Pathol.* 1995;23(2):102-114. doi:10.1177/019262339502300202.
8. CAMPOS DRDE. AVALIAÇÃO IN VITRO – IN VIVO DE DUAS FORMULAÇÕES DE PANTOPRAZOL: estudos farmacocinéticos e perfis de dissolução em meios biorelevantes. Published online 2008. <http://repositorio.unicamp.br/jspui/handle/REPOSIP/310120>.
9. CHENG F, Li W, LIU G, TANG Y. *In silico* ADMET Prediction: Recent Advances, Current Challenges and Future Trends. *Curr Top Med Chem.* 2013;13(11):1273-1289. doi:10.2174/15680266113139990033
10. CHIBALE K, DAVIES-COLEMAN M, MASIMIREMBWA C. Application of *in silico*, in vitro and in vivo ADMET/PK platforms in drug discovery. *Drug Discov Africa Impacts Genomics, Nat Prod Tradit Med Insights into Med Chem Technol Platforms Purs New Drugs.* 2012;9783642281(August):1-438. doi:10.1007/978-3-642-28175-4.
11. DANNI FUCHS, Flávio; WANNMACHER L. *Farmacologia Clínica e Terapêutica.* 5. ed. (GUANABARA KOOGAN LTDA, ed.). GEN; 2017.
12. EKINS S, NIKOLSKY Y, NIKOLSKAYA T. Techniques: Application of systems biology to absorption, distribution, metabolism, excretion and toxicity. *Trends Pharmacol Sci.* 2005;26(4):202-209. doi:10.1016/j.tips.2005.02.006.
13. FERREIRA LLG, ANDRICOPULO AD. ADMET modeling approaches in drug discovery. *Drug Discov Today.* 2019;24(5):1157-1165. doi:10.1016/j.drudis.2019.03.015.
14. GOLAN DE, TASHJIAN JUNIOR AH, ARMSTRONG EJ, ARMSTRONG AW. *Princípios de Farmacologia: A Base Fisiopatológica Da Farmacoterapia.* 2ª. (Koogan G, ed.). Guanabara Koogan; 2009.
15. GUAN L, YANG H, CAI Y, et al. ADMET-score-a comprehensive scoring function for evaluation of chemical drug-likeness. *Medchemcomm.* 2019;10(1):148-157. doi:10.1039/C8MD00472B.

16. GUIDO RVC, Andricopulo AD, Oliva G. Planejamento de fármacos, biotecnologia e química medicinal: Aplicações em doenças infecciosas. *Estud Avancados*. 2010;24(70):81-98. doi:10.1590/S0103-40142010000300006.
17. KHARKAR P. Two-Dimensional (2D) *In silico* Models for Absorption, Distribution, Metabolism, Excretion and Toxicity (ADME/T) in Drug Discovery. *Curr Top Med Chem*. 2010;10(1):116-126. doi:10.2174/156802610790232224
18. KELLER TH, PICHOTA A, YIN Z. A practical view of 'druggability'. *Curr Opin Chem Biol*. 2006;10(4):357-361. doi:10.1016/j.cbpa.2006.06.014
19. LANSDOWNE LE. Explorando o processo de desenvolvimento de medicamentos. Technology Networks. Published 2020. <https://www.technologynetworks.com/drug-discovery/articles/exploring-the-drug-development-process-331894>.
20. LE J. Absorção de Fármacos. MANUAL MSD Versão para Profissionais de Saúde. Published 2019. Accessed March 20, 2021. [https://www.msmanuals.com/pt-br/profissional/farmacologia-clinica/farmacocinetica/absorcao-de-farmacos#:~:text=geral da farmacocinetica.\),Difusão passiva,região de alta concentração \(p.&text=O índice de difusão é,área da superfície de absorção](https://www.msmanuals.com/pt-br/profissional/farmacologia-clinica/farmacocinetica/absorcao-de-farmacos#:~:text=geral da farmacocinetica.),Difusão passiva,região de alta concentração (p.&text=O índice de difusão é,área da superfície de absorção).
21. LI AP. Screening for human ADME/Tox drug properties in drug discovery. *Drug Discov Today*. 2001;6(7):357-366. doi:10.1016/S1359-6446(01)01712-3
22. LIN J, SAHAKIAN D, DE MORAIS S, XU J, POLZER R, WINTER S. The Role of Absorption, Distribution, Metabolism, Excretion and Toxicity in Drug Discovery. *Curr Top Med Chem*. 2003;3(10):1125-1154. doi:10.2174/1568026033452096
23. LOMBARDO F, DESAI P V, ARIMOTO R, et al. *In silico* Absorption, Distribution, Metabolism, Excretion, and Pharmacokinetics (ADME-PK): Utility and Best Practices. An Industry Perspective from the International Consortium for Innovation through Quality in Pharmaceutical Development. *J Med Chem*. 2017;60(22):9097-9113. doi:10.1021/acs.jmedchem.7b00487

24. LUCAS AJ, SPROSTON JL, BARTON P, RILEY RJ. Estimating human ADME properties, pharmacokinetic parameters and likely clinical dose in drug discovery. *Expert Opin Drug Discov.* 2019;14(12):1313-1327. doi:10.1080/17460441.2019.1660642.
25. MARQUES DOS SANTOS E. Farmacogenética: Influências sobre a farmacocinética e a farmacodinâmica. Published online 2007. <https://arquivo.fmu.br/prodisc/farmacia/ems.pdf>
26. MODA TL. Desenvolvimento de Modelos *In silico* de Propriedades de ADME Para a Triagem de Novos Candidatos a Fármacos. Published online 2007.
27. MOROY G, MARTINY VY, VAYER P, VILLOUTREIX BO, MITEVA MA. Toward *in silico* structure-based ADMET prediction in drug discovery. *Drug Discov Today.* 2012;17(1-2):44-55. doi:10.1016/j.drudis.2011.10.023
28. NORINDER U, BERGSTRÖM CAS. Prediction of ADMET properties. *ChemMedChem.* 2006;1(9):920-937. doi:10.1002/cmdc.200600155.
29. PEREIRA AL da C. Síntese, elucidação estrutural e estudos *in silico* de novos compostos 2-amino-tiofênicos imídicos candidatos a fármacos antifúngicos, antileishmanicida e antitumorais. 2019; (May).
30. PEROLA E, HERMAN L, WEISS J. Development of a Rule-Based Method for the Assessment of Protein Druggability. *J Chem Inf Model.* 2012;52(4):1027-1038. doi:10.1021/ci200613b.
31. PICCIRILLO E, Do AMARAL AT. Virtual screening of bioactive compounds: Concepts and applications. *Quim Nova.* 2018;41(6):662-677. doi:10.21577/0100-4042.20170210
32. QUEIROZ, Luz Hirano L. FARMACOCINÉTICA: Conceitos básicos aplicados à medicina veterinária. Published online 2011:36.
33. RODRIGUES AES. Importância do Conhecimento das Interações Fármaco-Nutrientes. Published online 2009. <http://ufpbdigital.ufpb.pt/dspace/handle/10284/989>.
34. ROY N, KADAM R. Recent trends in drug-likeness prediction: A comprehensive review of *In silico* methods. *Indian J Pharm Sci.* 2007;69(5):609. doi:10.4103/0250-474X.38464.

35. SIMÕES ACE. BASES MORFOLÓGICAS E FISIOLÓGICAS DAS VIAS DE ADMINISTRAÇÃO, ABSORÇÃO, METABOLIZAÇÃO E EXCREÇÃO DE FÁRMACOS. Published online 2013.
36. VAN DE WATERBEEMD H, GIFFORD E. ADMET *in silico* modelling: Towards prediction paradise? *Nat Rev Drug Discov.* 2003;2(3):192-204. doi:10.1038/nrd1032.
37. WANG J, URBAN L, BOJANIC D. Maximising use of *in vitro* ADMET tools to predict *in vivo* bioavailability and safety. *Expert Opin Drug Metab Toxicol.* 2007;3(5):641-665. doi:10.1517/17425255.3.5.641.
38. WHALEN K, FINKEL R, PANAVELIL TA. *Farmacologia Ilustrada*. 6ª. (Artmed Editora, ed.). ARTMED EDITORA LTDA; 2016. Accessed March 25, 2021. <https://books.google.com/books?hl=pt-PT&lr=&id=volJDAAAQBAJ&oi=fnd&pg=PP1&dq=Farmacologia+Ilustrada+-+6ª+Edição+Panavelil&ots=aiapiYvC9O&sig=S0ukAxLi5A37haJu7z9ia9ZRFWM>.
39. XAVIER A de L. Design Teórico , Síntese Multicomponente e Comprovação Experimental da Atividade Antinociceptiva de Pirimidinonas em Camundongos através das vias Intraperitoneal e Oral. Published online 2012

PREDIÇÃO DE DESCRITORES IN VITRO DE ADME: MODELO DE APOIO ÀS DECISÕES EMPÍRICAS.

Matheus Nunes da Rocha

Universidade Estadual do Ceará, Limoeiro -CE

Hélcio Silva dos Santos

Universidade Estadual Vale do Acaraú, Sobral-CE

Aluisio Marques da Fonseca

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Francisco Wagner de Queiroz Almeida Neto

Universidade Federal do Ceará

Gabrielle Silva Marinho

Universidade Estadual do Ceará, Iguatu -CE

Resumo

Recentemente novos recursos preditivos de ADME e toxicidade foram aprimorados e disponibilizados online e com estes, a capacidade de se estimarem parâmetros que até então, só seriam tangíveis em laboratório. É crescente o número de compostos disponíveis para testes *in vitro* e *in vivo* que demandam de tempo e custos e que, em muitos casos, levam a resultados indesejáveis. A partir da tão utilizada “regra dos cinco” de Lipinski, vários filtros que combinam propriedades físico-químicas foram aprimoradas ao longo dos anos para otimizar ainda mais o processo de seleção de novos candidatos promissores a medicamento. Este estudo aborda uma proposta preditiva que busca alinhar as propriedades físico-químicas de novos candidatos a fármaco aos atributos fundamentais de ADME: permeabilidade passiva (P_{app}),

volume de distribuição (VD), depuração (CL) e tempo de meia-vida, com o objetivo de alcançar um espaço físico-químico mais confiável para testes preditivos de farmacocinética.

Palavras-chave: Permeabilidade; volume de distribuição; depuração; Pfizer MPO; ADME consensual.

1. Introdução

Um dos maiores conflitos na indústria farmacêutica versam sobre a demanda exacerbada por novos medicamentos com propriedades farmacocinéticas otimizadas, principalmente em questões de segurança, tempo e custo de desenvolvimento destes produtos⁷. Como forma de melhor conduzir a descoberta de novas drogas, as predições de farmacocinética têm dado ênfase no alinhamento entre as propriedades depositadas em bibliotecas *in silico* e parâmetros *in vitro* relatados de ADME¹⁰.

Assim, espera-se que um bom alinhamento entre as propriedades físico-químicas calculadas e as decisões empíricas *in silico* garantem um bom alinhamento entre as propriedades farmacológicas de um candidato promissor a medicamento de via oral e, conseqüentemente, a satisfação de atributos farmacocinéticos e de segurança^{1;16}.

2. Revisão literária

A partir da comumente utilizada “regra dos cinco” (Ro5) de Lipinski⁸, surgiram novos filtros que combinam uma série de propriedades físico-químicas com parâmetros farmacológi-

cos como biodisponibilidade oral, tais como o modelo estatístico de Egan², que utiliza do alinhamento entre lipofilicidade e polaridade (log P e PSA) para estimar a absorção intestinal humana (HIA), bem como as regras de Veber¹⁴ e Ghose⁴ com enfoque em permeabilidade e distribuição nas diversas barreiras biológicas como a barreira hematoencefálica (BBB).

Em um conjunto de dados de parâmetros experimentais avaliados por Martin⁹, cerca de 85% das substâncias com ótima biodisponibilidade oral eram ânions com polaridade topológica (TPSA) inferior a 75 Å¹⁰, enquanto compostos neutros e não-violantes dos critérios da Ro5 compunham cerca de 55% do conjunto de dados, com melhor absorção em substâncias com coeficiente de distribuição (logD) < 4 no pH fisiológico (aproximadamente 7,4).

Recentemente, em um estudo publicado pela empresa farmacêutica Pfizer, Inc., através de Wager e colaboradores, foi elaborado um algoritmo matemático capaz de pontuar o perfil de substâncias semelhantes a medicamentos com base no alinhamento de seis parâmetros físico-químicos, que incluem propriedades intrínsecas e de ordem fisiológica^{17;18}:

- Coeficiente de partição (log P) ≤ 3;
- Coeficiente de distribuição (log D) em pH 7,4 ≤ 2;
- Peso molecular (MW) ≤ 360 g/mol;
- Doadores de ligação de hidrogênio (HBDs) < 1;
- Área de superfície polar topológica (TPSA) entre 40 e 90 Å²; e
- pKa do centro mais básico inferior a 8,0.

A absorção e distribuição são pautas fundamentais na seleção de um novo medicamento, uma vez que estes processos dependem de parâmetros físicos como difusão e volume distribuído (VD) entre o plasma sanguíneo e o tecido adiposo humano, como formas de estimativa de seu local de ação¹².

Neste algoritmo, cada propriedade é convertida em uma pontuação individual que varia de 0 a 1, onde o somatório total das propriedades dentro do espectro ideal do conjunto de dados da Pfizer varia de 0 a 6, conforme a quantidade de parâmetros satisfeitos, denominada pontuação de CNS MPO (do inglês *Central Nervous System Multiparameter Optimization*). A avaliação mostra que compostos com pontuação de CNS MPO superiores a 4,0 apresentam maiores chances de alinhamento entre três atributos fundamentais de ADME: alta permeabilidade passiva (P_{app}), baixo risco de efluxo por glicoproteína-P (P-gp) e baixa taxa de depuração intrínseca da fração livre no sistema microssomal ($CL_{int,u}$)¹⁷.

A abordagem dos autores^{7;5;6} validam a análise consensual entre decisões empíricas de ADME previstas *in silico*, baseadas no alinhamento entre propriedades físico-químicas, e predições de parâmetros farmacocinéticos a partir de conjuntos de dados *in vitro*.

2.1 Permeabilidade passiva

A permeabilidade passiva, ou permeabilidade aparente (P_{app}), é a capacidade de uma droga penetrar membranas biológicas por difusão passiva, em contraste a difusão ativa que está associada à substâncias com decréscimo em seu potencial de

permeabilidade. Com o objetivo de estimar a absorção precoce de drogas em suas etapas de desenvolvimento, dados de testes *in vitro* em linhagens celulares como as de adenocarcinoma color-retal (Caco-2) e células renais caninas Madin-Darby (MDCK) constituem modelos de preditivos de apoio às decisões empíricas *in silico* para seleção de compostos bem absorvidos^{3;15}.

No atual sistema de classificação da empresa Pfizer, Inc., o valor médio para estimativa de permeabilidade passiva está na faixa de $P_{app} > 10$, para a classe de substâncias com alta permeabilidade $2,5 < P_{app} \leq 10$ para substâncias com permeabilidade moderada e $P_{app} \leq 2,5$ para substâncias pouco permeáveis, em unidades na ordem de 10^{-6} cm/s¹⁷.

2.2 Volume de distribuição

O volume de distribuição aparente (VD) mede a fração molecular do fármaco absorvido que é capaz de se distribuir uniformemente entre o fluido sanguíneo, pela ligação às proteínas plasmáticas (PPB), e os tecidos do corpo humano. Como propriedade dependente da estrutura molecular, estruturas mais hidrofílicas e carregadas tendem a um baixo VD aparente devido a alta afinidade com o ambiente plasmático, enquanto as substâncias hidrofóbicas tendem a elevados valores de VD¹¹.

2.3 Depuração (*clearance*)

A depuração (CL) é um dos atributos farmacocinéticos fundamentais na seleção de um novo candidato a medicamento. Esta propriedade está intimamente ligada às propriedades físico-químicas dos fármacos, uma vez que substâncias mais lipofílicas tendem a ser excretadas por intermédio das biotrans-

formações dos mecanismos oxidativos de macromoléculas metabolizadoras (CYPs ou UGTs) presentes no fígado, enquanto compostos mais hidrofílicos tendem a ser excretados via renal com maior facilidade¹³.

No sistema de classificação da Pfizer, compostos com $CL_{int,u} \leq 100$ mL/min/kg possuem baixa taxa de depuração metabólica, o que garante um maior tempo de meia-vida para ação terapêutica de um medicamento¹⁷.

Exemplo: Estimando quantitativamente o perfil “druglikeness” e os descritores farmacocinéticos de um composto.

Em 2021, um novo servidor foi projetado com um conjunto de dados robusto e com mais de 80 propriedades disponíveis para avaliação farmacológica de um composto. A plataforma, chamada ADMETlab 2.0, gera como resultado uma combinação entre valores numéricos e simbologias que viabilizam as decisões empíricas que incluem: propriedades físico-químicas, química medicinal, propriedades de ADME, toxicidade e identificação de toxicóforos¹⁹.

3. Metodologia

3.1 Ferramentas

O software de licença acadêmica MarvinSketch versão 21.20, alimentado pela empresa ChemAxon, foi usado para aplicar o filtro de CNS MPO da Pfizer, apoiado pela predição de descritores in vitro de ADME do servidor ADMETlab 2.0.

3.2 Escolha do fármaco para a estimativa dos descritores

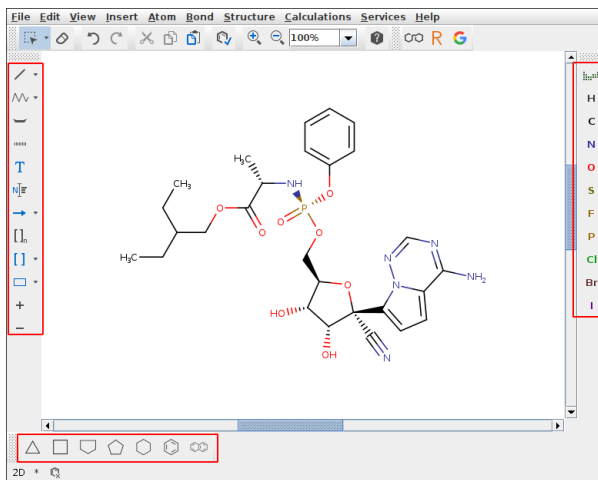
Assim como na secção anterior, foi escolhido aleatoriamente o composto Remdesivir, depositado no conjunto de dados molecular PubChem, sob o código SMILES

```
CCC(CC)COC(=O)C(C)NP(=O)(OCC1C(C(C(O1)(C#N)C2=CC=C3N2N=CN=C3N)O)O)OC4=CC=CC=C4.
```

3.3 Avaliação dos descritores de ADME

De forma intuitiva, a molécula pode ser projetada pelas ferramentas de desenho do programa para desktop MarvinSketch, como mostra a figura 1, assim como o código SMILES pode ser carregado no campo indicado pela ferramenta “ADMET Evaluation” no servidor ADMETlab 2.0, como mostra a figura 2. Então, o alinhamento entre as propriedades físico-químicas e os descritores de ADME pode ser estimado. É importante observar que os resultados da plataforma ADMETlab 2.0 podem variar de vermelho (+++), para as propriedades insatisfatórias ou limitantes, até verde (---), para propriedades dentro do ideal esperado, enquanto os tokens positivos (+) e negativos (-) variam conforme a intensidade do perfil avaliado.

Figura 1. Composto Remdesivir desenhado e corrigido em duas dimensões no software MarvinSketch.



Fonte: Elaborado pelo autor.

Figura 2. Submissão do código SMILES do Remdesivir para predição dos descritores de ADME do servidor ADMETlab 2.0.

Fonte: Elaborado pelo autor.

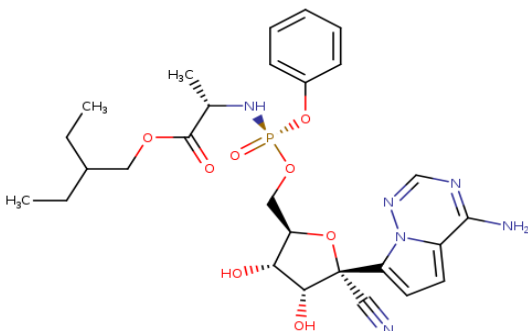
4. Resultados

O teste consensual de ADME podem fornecer centenas de informações sobre um composto, desde propriedades moleculares até respostas tóxicas preditivas, desta forma, o usuário pode organizar os resultados conforme achar coerente com o objetivo. Vale ressaltar que este não é um método padrão e o protocolo pode variar dependendo do objetivo do estudo.

A resposta visual dos resultados do cálculo de CNS MPO e dos descritores de ADME podem ser inspecionados na figura 3 e figura 4 e podem ser organizadas conforme mostra o quadro 1.

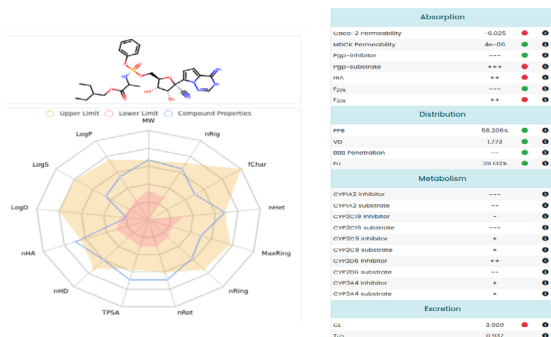
Figura 3. Resultado do cálculo do CNS MPO realizado no software MarvinSketch.

LOGP_SCORE	=	2,01	->	1,00
LOGD_SCORE	=	2,01	->	1,00
MW_SCORE	=	602,59	->	0,00
TPSA_SCORE	=	203,55	->	0,00
HBD_SCORE	=	5,00	->	0,00
PKA_SCORE	=	0,65	->	1,00
CNS_MPO_SCORE	=			3,00



Fonte: Elaborado pelo autor.

Figura 4. Resposta visual da predição dos descritores farmacocinéticos do Remdesivir no servidor online ADMETlab 2.0.



Fonte: Elaborado pelo autor.

Quadro 1. Resultados da predição de CNS MPO combinado a descritores de ADME do Remdesivir.

Parâmetro	Propriedade	Resultado
Propriedades físico-químicas	logP	2,01
	logD	2,01
	MW	602,59 g/mol
	TPSA	203,55 Å ²
	HBD	5
	pKa	0,65
	CNS MPO	3
Absorção	logP _{app} Caco-2	-6,025
	P _{app} MDCK	4·10 ⁻⁶ cm/s
	P-gp (inibidor)	---
	P-gp (substrato)	+++
	HIA	++
	Biodisponibilidade (F _{20%})	---
	Biodisponibilidade (F _{30%})	++
Distribuição	PPB	56,206 %
	VD	1,77
	BBB	--
Metabolismo	CYP2D6 (inibidor)	++
	CYP3A4 (inibidor)	+
	CYP2D6 (substrato)	--
	CYP3A4 (substrato)	+
Excreção	CL	3,91
	T _{1/2}	0,94

Fonte: Elaborado pelo autor.

5. Considerações finais

O teste consensual entre diversas ferramentas - desde programas a recursos disponíveis online - apresentam uma estratégia promissora que versa sobre aumentar o grau de confiabilidade nos testes de ADME e complementar os dados de decisões empíricas. Contudo, estas predições constituem apenas etapas iniciais preditivas para a seleção de novos candidatos promissores a fármaco, sendo necessária, ainda, a avaliação destas substâncias em testes *in vitro* e *in vivo*.

Referências

1. Bickerton GR, Paolini G V., Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. *Nat Chem.* 2012;4(2):90-98. doi:10.1038/nchem.1243
2. Egan WJ, Merz KM, Baldwin JJ. Prediction of drug absorption using multivariate statistics. *J Med Chem.* 2000;43(21):3867-3877.
3. Fredlund L, Winiwarter S, Hilgendorf C. In Vitro Intrinsic Permeability: A Transporter-Independent Measure of Caco-2 Cell Permeability in Drug Design and Development. *Mol Pharm.* 2017;14(5):1601-1609. doi:10.1021/acs.molpharmaceut.6b01059
4. Ghose AK, Herbertz T, Hudkins RL, Dorsey BD, Mallamo JP. Knowledge-based, central nervous system (CNS) lead selection and lead optimization for CNS drug discovery. *ACS Chem Neurosci.* 2012;3(1):50-68. doi:10.1021/cn200100h
5. Hughes JD, Blagg J, Price DA, et al. Physicochemical drug properties associated with *in vivo* toxicological outcomes. *Bioorganic Med Chem Lett.* 2008;18(17):4872-4875. doi:10.1016/j.bmcl.2008.07.071
6. Johnson TW, Dress KR, Edwards M. Using the Golden Triangle to optimize clearance and oral absorption. *Bioorganic Med Chem Lett.* 2009;19(19):5560-5564. doi:10.1016/j.bmcl.2009.08.045

7. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? *Nat Rev Drug Discov.* 2004;3(8):711-716. doi:10.1038/nrd1470
8. Lipinski CA. Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov Today Technol.* 2004;1(4):337-341. doi:10.1016/j.ddtec.2004.11.007
9. Martin YC. A Bioavailability Score. *J Med Chem.* 2005;48(9):3164-3170. doi:10.1021/jm0492002
10. Pathania S, Singh PK. Analyzing FDA-approved drugs for compliance of pharmacokinetic principles: should there be a critical screening parameter in drug designing protocols? *Expert Opin Drug Metab Toxicol.* 2021;17(4):351-354. doi:10.1080/17425255.2021.1865309
11. Pires DE V., Kaminskas LM, Ascher DB. Prediction and Optimization of Pharmacokinetic and Toxicity Properties of the Ligand. In: *Computational Drug Discovery and Design.* Humana Press; 2018:271-284. doi:10.1007/978-1-4939-7756-7_14
12. Rang R, Ritter JM, Flower RJ, Henderson G. Absorption and distribution of drugs. In: *Rang & Dale's Pharmacology.* 8th ed. Elsevier; 2015:261-292.
13. Smith DA, Beaumont K, Maurer TS, Di L. Clearance in Drug Design. *J Med Chem.* 2019;62(5):2245-2255. doi:10.1021/acs.jmedchem.8b01263
14. Veber DF, Johnson SR, Cheng H, Smith BR, Ward KW, Kopple KD. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J Med Chem.* 2002;45(12):2615-2623. doi:10.1021/jm020017n
15. Volpe DA. Variability in Caco-2 and MDCK Cell-Based Intestinal Permeability Assays. *J Pharm Sci.* 2008;97(2):712-725. doi:10.1002/JPS.21010
16. Wager TT, Chandrasekaran RY, Hou X, et al. Defining desirable central nervous system drug space through the alignment of molecular properties, in vitro ADME, and safety attributes. *ACS Chem Neurosci.* 2010;1(6):420-434. doi:10.1021/cn100007x

17. Wager TT, Hou X, Verhoest PR, Villalobos A. Central Nervous System Multiparameter Optimization Desirability: Application in Drug Discovery. *ACS Chem Neurosci*. 2016;7(6):767-775. doi:10.1021/acchemneuro.6b00029
18. Wager TT, Hou X, Verhoest PR, Villalobos A. Moving beyond rules: The development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. *ACS Chem Neurosci*. 2010;1(6):435-449. doi:10.1021/cn100008c
19. Xiong G, Wu Z, Yi J, et al. ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res*. Published online 2021:1-10. doi:10.1093/nar/gkab255

O USO DO SOFTWARE ARGUSLAB NA MODELAGEM MOLECULAR: UMA ESTRATÉGIA PARA O ENSINO E PESQUISA EM QUÍMICA POR MEIO REMOTO

Eduardo Menezes Gaieta

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Regilany Paulo Colares

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Aluísio Marques da Fonseca

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

RESUMO

Como forma de melhorar cada vez mais o aprendizado da Química tem-se feito buscas incessantes de recursos que promovam inovação na relação de aprendizagem. A química computacional emerge como grande possibilidade de fazer essa mediação com, entre outros, a modelagem molecular. Um recurso de investigação estrutural das propriedades moleculares com técnicas de visualização gráfica, fornecendo uma representação tridimensional (3D). Apresenta-se o ARGUSLAB, um software de modelagem molecular, como um recurso alternativo no ensino/ pesquisa, sendo realizada uma descrição do mesmo e de suas principais funcionalidades.

Palavras chaves: Modelagem Molecular; Arguslab; química; ensino - pesquisa.

1. INTRODUÇÃO

Desde o início do estudo da química como ciência, professores, pesquisadores e cientistas da área sempre buscaram criar modelos de estruturas e teorias químicas que auxiliassem na compreensão e construção da ciência química no seu todo. Até então pesquisadores da área concordam que o conhecimento da química passa pelas suas formas, níveis ou dimensões de representação, que são: a macroscópica, microscópica e simbólica, onde:

O nível macroscópico corresponde aos fenômenos e processos químicos observáveis e perceptíveis numa dimensão visível. Já o nível simbólico envolve as fórmulas, equações e estruturas. Por fim, o nível microscópico diz respeito aos movimentos e arranjos de moléculas, átomos e partículas¹⁴

Muitas das vezes como forma de estimular o domínio dos textos científicos e a terminologia científica, recorre-se ao emprego de Textos de Divulgação Científica (TDCs) como uma possibilidade de material didático em que os textos podem ser analisados e articulados conforme a necessidade de contextualizar os conteúdos propostos em sala de aula¹⁷. Um aspecto importante a ser considerado é que as TDCs não substituem os conceitos ou fenômenos de uma pesquisa científicas⁶, o emprego das TDCs vai muito mais além de motivar a curiosidade dos alunos. De algum modo elas promovem a recuperação significados mais amplos e diversificados para o ensino.

Nessa mesma ordem de ideia, modelos moleculares no formato físico têm sido empregados no ensino e pesquisa de química. Estes materiais de fácil manipulação possibilitam uma

visualização que estimulam um raciocínio sobre os fenômenos e processos químicos observáveis, se tornando assim recursos bastante úteis para a construção do conhecimento químico. No entanto, estes modelos são limitados porque não abrangem outros níveis tangíveis aos nossos órgãos sensoriais.

Como a evolução tecnológica é um dos acontecimentos que mais tem impactado a sociedade, por estarem indissociavelmente relacionada a dispositivos ou ferramentas digitais que fazem parte do cotidiano de todo cidadão da atualidade. O advento das Tecnologias de Informação e Comunicação (TICs) trouxe consigo mudanças em todos e diversos aspectos dinâmicos que estão inteiramente interligados com a Química, proporcionando um conjunto de ferramentas e recursos tecnológicos que desempenham papéis cada vez mais importantes contribuindo para o estudo e pesquisa da química a alcançar determinados fins pretendidos.

Um recurso das TDICs são os objetos de aprendizagem (OA), que hoje em dia se podem encontrar aplicados de diversos modos no sistema de ensino-aprendizagem e pesquisa em química trazendo assim:

[...] a possibilidade do aluno fazer inúmeras tentativas para construir hipóteses ou estratégias sobre determinado tema, podendo obter feedback do computador que o auxilia na correção dessa estratégia, tendo o professor como mediador dos conhecimentos embutidos no OA¹

Sendo a química uma ciência com fundamentos abstratos, a compreensão e o estudo do seu universo exigem a criação de várias hipóteses que sustentam a visualização mental dos fenô-

menos químicos que ocorrem em um mundo microscópico, envolvendo, entre outros, a manipulação de partículas, átomos, moléculas e íons.

Para a possível visualização mental e criação de hipóteses são usados os softwares educacionais (SE) considerados como um conjunto de recursos informáticos projetados com a intenção de serem usados no processo de ensino⁸, como também, direcionados às áreas científicas. Uma vez aplicados, no sistema de ensino-aprendizagem e pesquisa em química estes ambientes lógicos se tornam instrumentos facilitadores e ferramentas úteis que nos possibilitam estudar através de simulações de modelos químicos as propriedades de sistemas e substâncias com compreensão microscópicas.

O emprego de SE permite a simulação, a demonstração e as variáveis envolvidas nos fenômenos em que a matéria é transformada, em que são constatadas situações envolvendo a análise de arranjos geométricos, ligações químicas, atomística, processos físico-químicos, química orgânica, entre outros assuntos abordados pela química como ciência investigativa¹⁰

Assim, de forma a aperfeiçoar o entendimento dos já mencionados níveis ou dimensões de representação macroscópico, microscópico e simbólica, encetou-se a aliança entre a química e a tecnologia, resultando no que chamamos hoje de Química Computacional. Um conjunto de técnicas para investigação de problemas químicos no computador⁹. A Química Computacional faz uso de aspectos ou métodos, como o cálculos de mecânica molecular, mecânica quântica - os mais conhecidos: *ab initio*, os métodos semiempíricos e o método da teoria do fun-

cional da densidade - que são aspectos importantes no cálculo de propriedades de moléculas reais ou ainda não sintetizadas, e de sistemas moleculares⁷ e também da

Análise e tratamento da informação química, proveniente de experiências laboratoriais, monitorização instrumental/industrial e simulações, em tempo real ou armazenada em bases de dados (tipicamente com números de entradas da ordem de vários milhões) para, por exemplo: (a) previsão de espectros (RMN, Infravermelho, Massa, etc.) que complementam a comprovação de sínteses químicas⁷.

Com propósito e objetivo semelhante a química computacional, a ciência quimio-informática também se propõe a resolver os problemas de química usando recursos das ciências da computação e informação. A quimio-informática abrange as áreas de

- I. representação, visualização, manipulação e processamento de estruturas químicas;
- II. organização de bases de dados de estruturas químicas; e
- III. estudos das relações quantitativas entre estrutura e atividade/propriedade (QSAR/QSPR, do inglês, quantitative structure-activity/property relationships)²

Deve-se destacar que o conceito da quimio-informática compreender a química computacional, havendo convergência entre estas duas ciências. Os aspectos da química computacional que interligam a química, informática e inteligência artificial associam-se aos da quimio-informática⁷.

Chegamos às simulações de modelos químicos que também podemos chamar de modelagem molecular (MM). Pro põem através de técnicas computacionais uma caracterização profunda de determinados compostos de forma quantitativa e qualitativa podendo-se prever ou planejar a formação de novos compostos por seu intermédio. De acordo com a União Internacional de Química Pura e Aplicada (IUPAC em inglês) Modelagem molecular é a investigação das estruturas e das propriedades moleculares usando a química computacional e as técnicas de visualização gráfica visando fornecer uma representação tridimensional, sob um dado conjunto de circunstâncias.

Um aspecto que nos é fornecido por estes recursos os modelos 3D, podem ser facilmente movimentados em todos os eixos (X, Y e Z) sendo possível construir modelos para exibir as ligações químicas em formato 3D, a fim de compreender como estas moléculas se organizam no espaço, possibilitando ao estudante e o professor desenvolver suas próprias hipóteses e tentar verificar se elas são válidas”

A partir deste contexto, sabendo que através da MM pode se ter a visualização tridimensional, podem-se observar os aspectos estereoquímicos, relacionar a estrutura e as propriedades físicas, fazer o cálculo e a representação de várias propriedades moleculares, incluindo fórmula e massa moleculares e não apenas. O presente trabalho descreve de forma demonstrativa o software ARGUSLAB (<http://www.arguslab.com/arguslab.com/ArgusLab.html>) como um recurso capaz auxiliar no processo ensino-aprendizagem e na pesquisa científica química.

2. Referencial teórico

Além do ARGUSLAB, existem outros softwares de modelagem molecular que apresentam funcionalidades similares que podem ser vistos como alternativas dependendo de cada situação. Como visto no quadro 1.

Quadro 1. Alguns softwares de modelagem e suas características.

Software	Idioma	Licença	Características
Arguslab	Inglês.	Gratuita	Este SE permite a visualização 3D, permite realizar o desenho de estruturas, fazer o docking molecular, visualizar o mapa potencial eletroestático, calcular os estados excitados de uma molécula, e outras.
Chem - Drawn/CS Chem3D	Inglês.	Consultar o Distribuidor	Estas ferramentas permitem aos usuários fazer desenhos de estruturas, e visualizá-las em projeções Fischer, e projeções de Newman, permite também o desenhos de perspectiva, estudar a estereoquímica de uma estrutura, é possível obter o nome da estrutura de acordo com a IUPAC, fórmula química, Simulação de espectro de NMR, Simulação de espectro de massa, limpeza da estrutura, visualização de modelos em 2D e 3D, renderização de alta qualidade e outros.
Avogadro	Inglês, outros para breve.	Gratuita	Desenhar e visualizar estruturas molecular em 3D, montagem, edições de moléculas e outras.

Software	Idioma	Licença	Características
HyperChem	Inglês	Paga	Esta ferramenta permite ao usuário a visualização e manipulação de moléculas 3D, traz consigo a mecânica molecular, dinâmica molecular e métodos orbitais moleculares semi-empíricos e ab-initio, simulações de proteínas Charmm, cálculos de estrutura, espectros, constantes de taxa e muito mais. O HyperChem é aplicável a macromoléculas, bem como a moléculas pequenas, e é programável, ainda contém um grande base de dados de moléculas.
Discovery Studio	Inglês	Gratuita	Esta ferramenta permite visualizar em 3D, permite ter acesso a simulações de motor de campo de força, incluindo ponto único, minimização, Simulações de Dinâmica Molecular e cálculos de Energia, Mecânica quântica completa ab initio, ferramentas de ionização de proteínas baseadas em pH, análise de mutação. revisão confiável de ligação proteína-proteína com ZDOCK, Docking baseado em física original (CHARMm CDOCKER) motor, gere mapas de densidade de elétrons, e outros.

Fonte: Elaborado pelos autores.

3. Procedimento metodológico

Por esta pesquisa ser de caráter qualitativo, a sua realização deu-se através de revisão e análise de bibliografias com relevância neste tema com o propósito de dar aporte e servir de embasamento teórico para esta pesquisa.

A pesquisa bibliográfica busca a resolução de um problema (hipótese) por meio de referenciais teóricos publicados, analisando e discutindo as várias contribuições científicas. Esse tipo de pesquisa trará subsídios para o conhecimento sobre o que foi pesquisado, como e sob que enfoque e/ou perspectivas foi tratado o assunto apresentado na literatura científica⁶.

Também se fez uso da plataforma virtual Protein Data Bank (PDB), um banco de dados em 3D de proteínas e ácidos nucléicos para baixar a proteína usada em um dos exemplos cujo download foi feito de forma gratuita pelo link do site oficial do seu desenvolvedor⁴ e do software ARGUSLAB.

Na perspectiva de PIZZANI et al. (2012)¹⁵ a pesquisa bibliográfica é o marco essencial na elaboração de novos estudos, teses, artigos entre outras que servirão de novo ponto para início de outras pesquisas.

4. Resultados e discussões

Muitos softwares vêm sendo utilizados na química computacional tanto para fins educacionais como para pesquisa. O ARGUSLAB é um software cuja versão se encontra limitada para o sistema operativo Windows até o momento. Quanto a dispositivos compatíveis, até ao momento, apenas é possível instalar e usá-lo em notebooks e PCs.

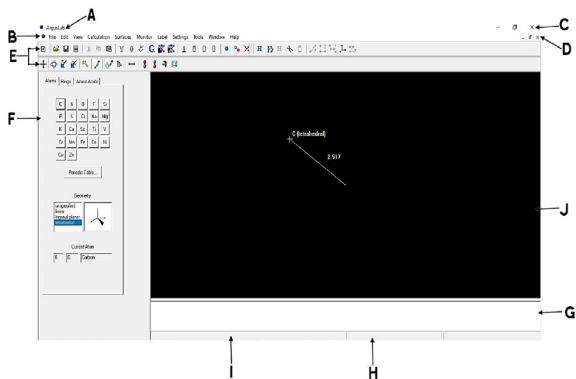
O ArgusLab, como qualquer outro software de MM, possibilita aos seus usuários fazerem docking ou acoplamento mole-

⁴ <https://www.rcsb.org/>

cular, visualizar estrutura moleculares em 3D, realizar cálculos moleculares e análises na construção molecular, calcular uma energia de ponto único de uma molécula, calcular os espectros UV / visível de uma molécula, otimizar a geometria de uma molécula, fazer interpretação das propriedades moleculares, além de outras funcionalidades.

A figura 1 e o quadro 1 ilustram e descrevem a interface do ArgusLab respectivamente.

Figura 1. interface do software ArgusLab.



Fonte : Elaborada pelos autores a partir da interface do software ArgusLab.

Quadro 2. Descrição da interface do software ArgusLab.

A - Barra de título	Esta barra não só mostra o título do projeto que está sendo trabalhado, como também mostra a localização em que o mesmo se encontra salvo.
B - Barras de menu	É nestas barras onde se concentram os comandos do software ArgusLab.
C - Controles do Windows	Esta barra é composta pelos controles maximizar, minimizar e fechar correspondente ao software.
D - Controles janela do projeto	Esta barra contém os mesmos controles maximizar, minimizar e fechar, que correspondem ao projeto aberto no momento.
E - Barra de ferramenta	Aqui se concentram os atalhos dos comandos mais utilizados.
F - Barra de ferramentas de construção	Aqui se concentram as ferramentas usadas na construção de moléculas. Também pode ser possibilitar a visualização, os itens envolvidos no projeto, isso se manualmente for alterado pela barra de menu/tools/molécula tree view.
G - Janela de exibição dos cálculos	Esta janela exibe os cálculos feitos durante a execução de uma configuração aplicada ao projeto
H - Janela de exibição de progresso	Esta janela exibe o estado da execução de uma dada configuração. Ex.: executando (Running); feito (Done)
I - Janela descritiva	Esta janela exibe a atual configuração inserida/aplicada.
J - Janela de desenho do projeto	É nesta janela onde o aluno/pesquisador desenha as estruturas desejada.

Fonte: Elaborada pelos autores

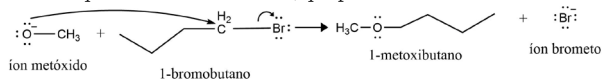
Um das principais funcionalidades do ArgusLab é criar superfícies de mapa de potencial eletrostático (MPE), que apresenta uma escala de cores - que geralmente variam de vermelho à azul - para ilustrar a distribuição de cargas, indicando as zonas ou regiões com maior e menor densidade de elétrons. A cor vermelha indica a região mais negativa, ou seja, com muito elétrons e azul indica a região mais positiva, ou seja, com pouco elétrons¹³.

A partir do princípio de que cargas opostas se atraem e cargas iguais se repelem, esse mapa pode ser usado para identificar regiões onde certas moléculas possivelmente são mais susceptíveis aos ataques nucleofílico e eletrofílico.

Ao estudar química orgânica, uma boa maneira de começar a ligar estruturas a propriedades (como a reatividade química) é encontrar a parte positiva de uma molécula e a parte negativa de outra. Quase sempre, esses serão os locais reativos⁵.

Vejamos o seguinte exemplo (fig. 2). Na reação de substituição nucleofílica entre o 1-bromobutano com o íon metóxido (nucleófilo), o nucleófilo ataca o carbono ligado ao halogênio e com isso a ligação carbono-halogênio é rompida e os elétrons desta ligação ficam com o halogênio (grupo de saída).

Figura 2. Exemplo de uma reação de substituição nucleofílica de um haleto de alquila com um íon alcóxido, que produz um éter.

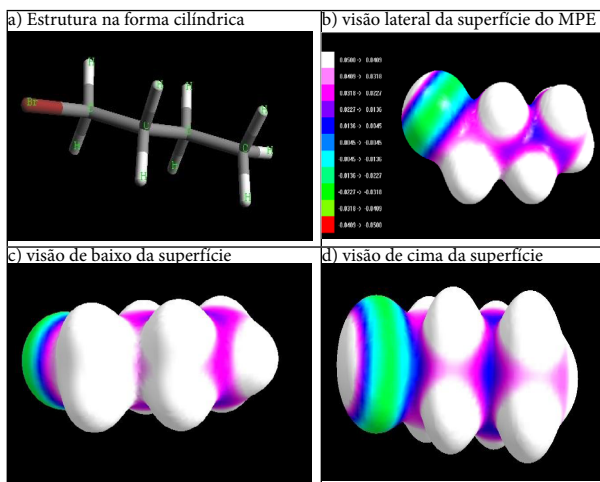


Fonte: Elaborado pelos autores a partir de Word

Desta forma, para complementar os conceitos teóricos podemos usar o ArgusLab para criar um MPE da molécula do 1-bromobutano (fig.3) a fim de identificar as zonas mais susceptível ao ataque do nucleofílico. O ArgusLab exhibe seu MPE com a escala de cor variando de branco para vermelho. Sendo a região vermelha a região mais negativa e a branca a mais positiva.

De acordo com a escala de cor do MPE do ArgusLab, branca na superfície do 1-bromobutano está situando as regiões mais positivas da molécula, para o nosso caso estão indicando as regiões mais favoráveis ao ataque nucleofílico. Já as cores azul e verde estão indicando as regiões intermediárias tendendo para a região negativa da superfície do MPE.

Figura 3. Mapa do potencial eletrostático da molécula do 1-bromobutano.



Fonte: Elaborado pelos autores a partir de imagem gerada no software ArgusLab

Esta funcionalidade se torna bastante útil para o aluno pesquisador, uma vez que facilita a compreensão da interação dos elétrons de duas moléculas, podendo assim prever ou planejar a formação de novos compostos, o que é bastante útil e funda-

mental em áreas como na química farmacêutica. Em química orgânica este MPE é usado como alternativas para descrever moléculas complexas interagem umas com as outras¹⁶.

Como outros softwares o ArgusLab também possui a funcionalidade de criar estruturas de preenchimento no espaço. Assim, podemos usar esta funcionalidade para compreender melhor a reatividade dos haletos de alquila (metila, primário, secundário ou terciário) nas reações de substituição nucleofílica.

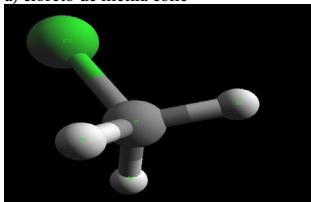
O nucleófilo deve se aproximar do haleto de alquila pelo lado oposto à ligação com o grupo de saída e essa aproximação é impedida pelos substituintes alquila do carbono que está sendo atacado^{5,11}.

No exemplo da Fig. 4 é possível observar que o nucleófilo se aproximaria com maior facilidade ao (C-Cl) do cloreto de metila porque os hidrogênios não estão dando cobertura suficiente ao carbono. Desta forma a reação ocorreria com maior rapidez em relação ao cloreto de terc-butila que é composto por grupos metila, neste caso, a aproximação do nucleófilo em volta do (C-Cl) seria dificultada pelos grupos metilas que estão dando cobertura ou protegendo o carbono que faz a ligação com o haleto.

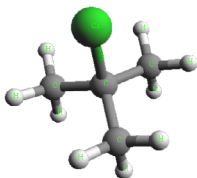
Assim poderíamos dizer que o a substituição de um dos hidrogênios por um grupo metila de determinada forma protege o carbono contra o ataque pelo nucleófilo e faz com que o cloreto de terc-butila seja menos reativo que o cloreto de metila^{5,18}.

Figura 4. Modelos cilindros e bolas (em cima) e preenchimento no espaço (em baixo)

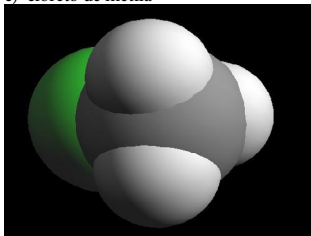
a) cloreto de metila cone



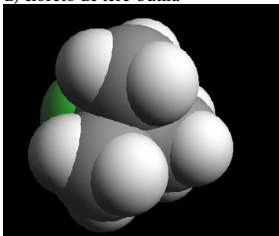
b) cloreto de terc-butila ando



c) cloreto de metila



d) cloreto de terc-butila

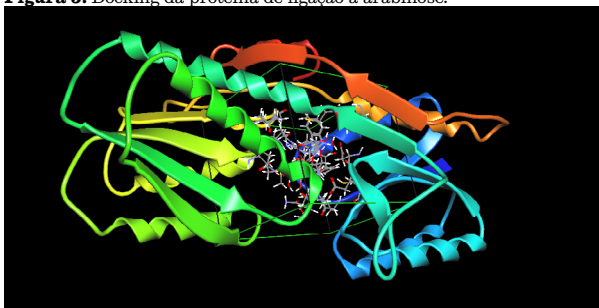


Fonte: Elaborado pelos autores a partir de imagem gerada no software ArgusLab

Numa perspectiva interdisciplinar, outra interessante funcionalidade que este software traz consigo é o docking molecular, que é comumente aplicado no estudo e pesquisada de bioquímica, biologia molecular e farmácia. A característica do docking molecular é a simulação de ligação que ela realiza na forma de pesquisa em uma molécula de modo a achar um potencial local candidato a um ligante e propor hipóteses de novas estruturas. “O Docking molecular pode ser definido como a previsão da estrutura dos complexos receptor-ligante, onde o receptor é geralmente uma proteína ou um oligômero de proteína e o ligante é uma molécula pequena ou outra proteína⁴⁷”.

Como exemplo de demonstração dessa funcionalidade (Fig.5), foram baixados - a partir da plataforma Protein Data Bank (PDB) - os dados proteína de ligação à arabinose (id: 5abp) aleatoriamente, que tem como ligante α -D-galactose.

Figura 5. Docking da proteína de ligação à arabinose.



Fonte: Elaborado pelos autores a partir de imagem gerada no software ArgusLab

5. Considerações finais

É vital a importância das ferramentas tecnológicas e computacionais no mundo da ciência, na química em particular ela é muito abrangente. Acreditamos ter demonstrado, com essa descrição feita do ArgusLab o quão útil está ferramenta pode ser quando usada em determinados fins pretendidos. Num dos exemplos apresentados conhecemos a funcionalidade que nos auxilia a compreender conceitos de interação moleculares, conceitos estes que são bastante abstratos quando explicados só com teorias. A possibilidade de poder ter acesso a imagens em três dimensões, de poder confrontar conceitos teóri-

cos com simulações ao alcance dos nossos órgãos sensoriais e poder criar hipóteses que facilitam a construção de um dado conhecimento são realidades possíveis no uso deste software.

Contudo, acredita-se que o emprego do ArgusLab independentemente da função desejada se torna uma alternativa educacional para o ensino e pesquisa de química e outras áreas como farmácia, sendo bastante útil visto poder realizar trabalhos simples, mas com resultados de qualidade, eficientes e integrativa pois pode ser usado em casa em um computador.

6. Referências

1. AGUIAR, E. V. B.; FLÔRES, M. L. P. Objetos de aprendizagem: conceitos básicos. **Objetos de aprendizagem: teoria e prática. Porto Alegre: Evangraf**, p. 12–28, 2014.
2. ALVES, V. M. et al. *Quim. Nova.*, v. 41, n. 2, p. 202–212, 2018.
3. BOCCATO, V. R. C. Metodologia da Pesquisa Bibliográfica na Área Odontológica e o Artigo Científico como Forma de Comunicação. **Revista de Odontologia da Universidade Cidade de São Paulo**, v. 18, n. 3, p. 265–274, 2006.
4. BROOIJMANS, N.; KUNTZ, I. D. Molecular recognition and docking algorithms. **Annual review of biophysics and biomolecular structure**, v. 32, n. 1, p. 335–373, 2003.
5. CAREY, F. *Química Orgânica-Vol. 1*. v. 1, p. 591, 2008.
6. DE, L. N.; FERREIRA, A.; QUEIROZ, S. L. **Textos de Divulgação Científica no Ensino de Ciências: uma revisão**[dialnet.unirioja.es](http://www.periodicos.capes.gov.br/ez67.periodicos.capes.gov.br/index.php?option=com_phome). [s.l.: s.n.]. Disponível em: <http://www.periodicos.capes.gov.br/ez67.periodicos.capes.gov.br/index.php?option=com_phome>. Acesso em: 27 ago. 2020.
7. FERNANDES, F. M. S. S. Perspectives of Computational Chemistry. n. December 2011, 2011.

8. FIALHO, N. N.; MATOS, E. L. M. A arte de envolver o aluno na aprendizagem de ciências utilizando softwares educacionais. **Educar em Revista**, n. spe2, p. 121–136, 2010.
9. LEWARS, E. G. **Computational chemistry: Introduction to the theory and applications of molecular and quantum mechanics: Third Edition 2016**. [s.l.: s.n.].
10. MACHADO, A. S. Uso de Softwares Educacionais, Objetos de Aprendizagem e Simulações no Ensino de Química. **Química Nova na Escola**, v. 38, n. 2, p. 104–111, 2016.
11. MCMURRY, J. **Química orgânica: combo**. [s.l.] Cengage Learning, 2012.
12. MORELLATO, C. et al. Softwares Educacionais e a Educação Especial: Refletindo sobre Aspectos Pedagógicos. **Renote**, v. 4, n. 1, p. 1–10, 2006.
13. ORTOLAN, A. O.; BRANCO, P. Apostila de práticas de química computacional. 2014.
14. PAULETTI, F.; ROSA, M. P. A.; CATELLI, F. A importância da utilização de estratégias de ensino envolvendo os três níveis de representação da Química. **Revista Brasileira de Ensino de Ciência e Tecnologia**, v. 7, n. 3, 2014.
15. PIZZANI, L. et al. **A arte da pesquisa bibliográfica na busca do conhecimento**
The art of literature in search of knowledge. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 10, n. 1, p. 53, 2012.
16. RIBEIRO, L. Caracterização de estado sólido e análise computacional de uma nova forma cristalina do fármaco antifilariose dietilcarbamazina: um sal de ácido maleico. p. 117, 2011.
17. ROSA, A. P. A utilização de textos de divulgação científica no ensino de Química The use of scientific dissemination texts in Chemistry teaching El uso de textos de divulgación científica en la enseñanza de la Química Recebido: 03 / 04 / 2020 | Revisado : 16 / 04 / . **Research, Society and Development**, v. 2020, p. 1–24, 2020.
18. SOLOMONS, T. W. G.; FRYHLE, C. B. Química Orgânica, 10ª edição. **Rio de Janeiro**, 2012.

19. SANT'ANNA, C. M. R. Glossário de termos usados no planejamento de fármacos (recomendações da IUPAC para 1997). **Química Nova**, v. 25, n. 3, p. 505–512, 2002.
20. FERREIRA, P. F. M.; JUSTI, R. DA S. Modelagem e o “fazer ciência”. **Química nova na escola**, v. 28, p. 32–36, 2008.
21. OLIVEIRA, T. B. **Emprego de ferramentas de quimioinformática no estudo do perfil metabólico de plantas e na desreplicação de matrizes vegetais** Universidade de São Paulo, , [s.d.].o

O ESSENCIAL SOBRE DINÂMICA MOLECULAR

Lucas Pinheiro Coutinho

Universidade Federal do Ceará (UFC), Fortaleza-Ceará

Renato Veríssimo de Oliveira

Universidade Federal do Ceará (UFC), Fortaleza-Ceará

Emmanuel Silva Marinho

Universidade Federal do Ceará (UFC), Fortaleza-Ceará

Norberto de Kássio Vieira Monteiro

Universidade Federal do Ceará (UFC), Fortaleza-Ceará

Resumo

Algumas dificuldades enfrentadas pela indústria química e farmacêutica são o alto gasto com reagentes e tempo de experimentos. Portanto, se faz necessário encontrar alternativas teóricas que nos permitam reduzir essas dificuldades empregando cálculos e simulações computacionais. A química computacional nos possibilita compreender em detalhes o comportamento de sistemas atômicos e moleculares de interesse, reproduzindo condições experimentais e permitindo o estudo de suas propriedades. A dinâmica molecular (DM), um método empírico de química computacional, baseia-se na mecânica clássica e estuda a evolução temporal de sistemas de N -partículas, de maneira a construir as trajetórias dessas partículas e, com base na mecânica estatística, realizar o cálculo de propriedades para análise do comportamento do sistema. As simulações DM permitem o planejamento de fármacos, investigação de diversos sistemas biológicos, como membranas e proteínas e o estudo de sistemas macromoleculares. Atualmente, as simu-

lações DM são amplamente utilizadas em diferentes áreas de pesquisa, tornando-a uma das técnicas teóricas mais citadas. Este capítulo tem por finalidade proporcionar ao leitor uma breve fundamentação sobre a DM, discutindo sobre sua origem, importância e aplicações, conceitos e pressupostos, etapas de construção, tipos de análises, bem como os principais softwares utilizados em DM.

Palavras-chaves: Dinâmica Molecular; Cálculos e Simulações Computacionais; Campo de Força; Software; Algoritmo; Química Computacional.

1. INTRODUÇÃO

1.1. Contexto-histórico

Os antecedentes da DM podem ser traçados desde o início do desenvolvimento do método Monte-Carlo (MC) para a pesquisa científica. Nos anos 1950, um crescente interesse na descrição temporal de sistemas de N-partículas impulsionou o desenvolvimento de métodos como a DM e o MC²⁶. Em 1953, o algoritmo Metropolis-Rosenbluth foi criado para obter amostragens aleatórias de distribuições de probabilidade de difícil amostragem direta. Esse algoritmo foi aplicado na descrição de equações físicas no método MC. Em 1955, Fermi e colaboradores no Laboratório Nacional de Los Alamos (LANL) executaram simulações no computador MANIAC I de uma cadeia unidimensional de 64 partículas de interações não-lineares. Os resultados exibiram pouca tendência para a equipartição da energia do sistema. Apesar de ambos os métodos MD e MC fornecerem descrições atômicas de sistemas em equilíbrio, eles não podem ser confundidos. A DM, um método deter-

minístico, estuda a evolução temporal de partículas clássicas gerando uma trajetória que é obtida pelas equações de Newton. Em contraponto, o método MC, um método estocástico, não fornece informações sobre a evolução temporal. No caso, um conjunto de configurações representativas são geradas a partir de perturbações aleatórias, sendo essas configurações utilizadas para cálculos de propriedades termodinâmicas^{34; 2}.

As primeiras simulações de DM envolviam o estudo de um fluido de esferas duras e foram feitas em 1957 por Alder e Wainwright, no então Laboratório de Radiação de Livermore da Universidade da Califórnia²⁶. Evidências foram identificadas de uma transição de fase sólido-fluido não encontrada em simulações de MC. Em 1964, A. Rahman, conhecido como o “pai da DM”, realizou simulações com argônio líquido afim de estudar as correlações de dependência espacial e temporal de dois corpos utilizando o potencial de Lennard-Jones^{26;38}. Nesse estudo A. Rahman demonstrou como calcular algumas propriedades comuns em DM como as funções de distribuição radial e coeficientes de autodifusão. Em outro estudo, A. Rahman e Stillinger, em 1971, simularam água, até então algo difícil por conta das interações de Van Der Waals, Coulomb e ligações de hidrogênio. As simulações auxiliaram no estudo do mecanismo de difusão da água^{26;38}. Nos anos posteriores, simulações DM começaram a ser utilizadas para sistemas como proteínas. Em 1975, Levitt e Warshel estudaram o enovelamento de proteínas através de simulações DM e em 1977, McCammon e colaboradores realizaram simulações com proteínas dobradas. Ambos os estudos envolviam inibidores pancreáticos de tripsina bovina (PTI). Esses estudos consolidaram o uso de simulações DM para sistemas macromoleculares³⁸.

Nos anos 80, o desenvolvimento de algoritmos mais robustos como o de Andersen e de Parrinello-Rahman, utilizados em *ensembles* de etapas de equilíbrio em simulações DM, auxiliaram a manter constantes temperatura e pressão durante as etapas de produção^{3;26;35}.

Com o salto nos últimos anos no desenvolvimento de poder e capacidade computacional, o potencial de simulações DM aumentou significativamente³⁸. Em 1998, Duan e Kollman realizaram uma simulação DM de uma proteína dobrando em água realizada em um tempo de um microssegundo. Isso foi possível devido ao uso de computadores paralelos que aumentaram a eficiência das simulações¹¹. Em 1999, Sugita e Okamoto utilizaram o método DM *Replica-Exchange* no estudo do enovelamento de proteínas. Esse método permitiu, a partir de uma mesma simulação, obter amostragens para uma grande faixa de temperaturas³⁹. Atualmente, as simulações DM são amplamente utilizadas em diferentes áreas de pesquisa, dando-se destaque para simulações de sistemas de biomoléculas, tornando-a uma das técnicas teóricas mais citadas³⁸.

1.2. Conceituação

Conforme sabemos da Física, a terminologia *dinâmica* refere-se ao ramo da mecânica clássica que estuda as *causas* do movimento. Portanto, a DM faz referência ao estudo da movimentação de espécies (moléculas, átomos, íons) ao longo do tempo com o objetivo de se construir trajetórias que nos possibilitem o cálculo de propriedades físico-químicas relevantes. Nesse sentido, constatamos que as simulações computacionais por DM geram um conjunto de informações microscópi-

cas sobre posições e velocidades atômicas e moleculares que permitem o cálculo de propriedades macroscópicas (pressão, energia interna etc) relevantes para uma análise mais acurada do sistema analisado, em que estas propriedades são calculadas baseando-se na mecânica estatística (ME)².

O ramo da física responsável pelo movimento dos corpos é a mecânica analítica ou dinâmica ou apenas mecânica³⁷. Esse ramo do conhecimento físico foi objeto de algumas das primeiras pesquisas realizadas pelos pioneiros da física, desde os gregos, passando por Galileu, Newton, Lagrange e Hamilton.

De acordo com N. Morgon e K. Coutinho:

A Dinâmica Molecular é uma técnica computacional em que se determinam os movimentos das partículas de qualquer sistema, do qual se conhece o potencial de interação entre estas partículas e as equações que regem seu movimento. A Dinâmica Molecular permite, portanto, estudar a evolução temporal das configurações dos constituintes do sistema e, a partir das sequências de posições geradas, determinar as propriedades macroscópicas do sistema, conforme os princípios fundamentais da Mecânica Estatística (p. 414, 2007).

A DM pode ser compreendida como o ramo da química teórica que realiza o estudo da evolução temporal de um sistema fundamentando-se, para isto, na mecânica newtoniana, lagrangiana e hamiltoniana, assim como na ME desenvolvida por J. C. Maxwell, L. Boltzmann, J. W. Gibbs, entre outros. Contudo, formulações contemporâneas da mecânica, como a mecânica relativística proposta por Einstein, assim como a mecânica

quântica (seja o tratamento ondulatório proposto por Schrödinger ou o tratamento matricial desenvolvido por Heisenberg) não são consideradas em Dinâmica Molecular Clássica.

1.3. A importância e aplicações da Dinâmica Molecular

A DM (assim como os demais métodos de química teórica) é importante para o entendimento e para a fundamentação da química, pois nos permite: compreender a nível atômico-molecular o comportamento de determinado sistema; simular computacionalmente propriedades físico-químicas de interesse e, ademais, possibilita-nos evitar a utilização de reagentes em laboratório sem termos analisado previamente se determinado sistema se comporta conforme o esperado.

Segundo N. Morgon e K. Coutinho, a DM é uma área que possui boa correlação com resultados experimentais, além de que nos permite tratar uma considerável diversidade de sistemas, a saber:

[...] sistemas homogêneos como gases, fluidos supercríticos, líquidos, soluções e misturas, a sistemas pouco (ou nada) homogêneos como interfaces, filmes de Langmuir-Blodgett, biomembranas, polímeros orgânicos e inorgânicos, polissacarídeos, lipídios, proteínas, ácidos nucleicos e seus complexos, passando por sistemas como zeólitos, argilas, sólidos cristalinos e vítreos, e nanomateriais, dentre outros (p. 415, 2007).

Além disso, a DM tem se mostrado uma ferramenta altamente favorável para estudos em Biologia Molecular, contribuindo para o desenvolvimento de novos medicamentos,

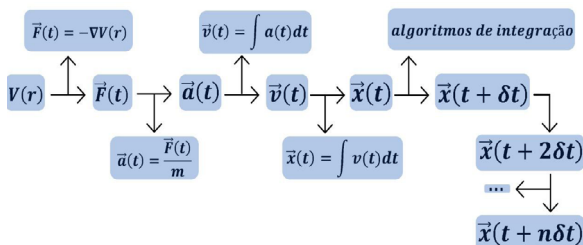
assim como para um estudo sistemático do comportamento de proteínas e outras biomoléculas de relevância, permitindo assim uma resolução temporal acurada.²⁰

2. Fundamentação

2.1. Pressupostos

Conforme podemos observar na Figura 1, uma DM começa, basicamente, conhecendo-se o potencial de interação (também conhecido como campo de força em áreas como química e biologia) que descreve cada espécie j do sistema e termina com o cálculo da trajetória das espécies. Conhecida a trajetória, pode-se realizar o cálculo das propriedades físico-químicas de interesse com base na ME.

Figura 1. Fluxograma de uma DM genérica.



Fonte: autores.

Conhecido o potencial, $V(\mathbf{r})$, escalar dependente da posição, é possível calcular a força, haja vista que esta é dada pelo

gradiente do potencial e, em seguida, a aceleração é encontrada aplicando-se a segunda lei de Newton¹³. Consecutivamente, a velocidade e a posição poderão ser calculadas utilizando-se conceitos de cálculo diferencial e integral, pois é sabido que a aceleração é a taxa de variação temporal da velocidade e a velocidade é a integral da aceleração no domínio do tempo. Semelhantemente, a velocidade pode ser compreendida como a taxa de variação temporal da posição e, por conseguinte, a posição é dada pela integral da velocidade no domínio do tempo¹⁴.

Finalmente, a trajetória das espécies constituintes do sistema poderá ser obtida por meio de algoritmos específicos através do método das diferenças finitas que, utilizando incrementos no tempo ($t + \delta t$) leva a construção da trajetória 25. Em decorrência da trajetória construída, propriedades relevantes poderão ser obtidas estatisticamente dentro do *ensemble* considerado na DM.

Em geral, podemos afirmar que uma DM começa com a resolução do Princípio do momento de Newton ou das equações canônicas de Hamilton, como visto nas Equações (1) e (2):

$$\begin{aligned}
 (1) \\
 F = ma &= \frac{dp}{dt} = \frac{d(mv)}{dt} = m \frac{dv}{dt} = m \frac{d^2q}{dt^2} = - \frac{\partial V}{\partial q} \\
 (2) \\
 \dot{p} &= \frac{dp}{dt} = - \frac{\partial \mathcal{H}}{\partial q} \quad e \quad \dot{q} = \frac{dq}{dt} = \frac{\partial \mathcal{H}}{\partial p}
 \end{aligned}$$

Em que P e Q , representam, respectivamente, o momento generalizado e a coordenada generalizada estabelecidos pela mecânica lagrangeana. \mathcal{H} representa o operador hamiltonia-

no e, para a maioria dos sistemas, corresponde a energia total ($\widehat{\mathcal{H}} = \widehat{T} + \widehat{V}$). Portanto, uma DM consiste em responder três principais perguntas, a saber: qual o tipo de potencial que representa o sistema? Qual o tipo de integração numérica utilizada para solucionar a equação ordinária vista acima? Quais são as condições iniciais para cada trajetória? Tais perguntas serão discutidas nos subtópicos 2.2, 2.3 e 2.4.^{15; 28}

2.2. Campo de Força

Em um sistema de N partículas, uma das formas de expressarmos a energia potencial total (E ou V) dessas partículas constituintes do sistema é através de uma soma referente as interações ligadas (comprimentos de ligação, ângulos de ligação e ângulos diedros) e não-ligadas (interações de van der Waals e de Coulomb), conforme nos mostra a Equação (3):

$$E = V(r) = V_{\text{ligadas}} + V_{\text{não-ligadas}} \quad (3)$$

Em DM a Equação 3 é chamada de campo de força, uma expressão genérica para potencial do sistema³². Desdobrando a Equação 3 em função dos termos moleculares temos a Equação (4) e expandindo os cada termo temos a Equação (5)³⁶.

$$U_T = U_l + U_\theta + U_\phi + U_{vdw} + U_{coul} \quad (4)$$

$$U(r_1, \dots, r_n) = \sum_{\text{ligações}} \left(\frac{a_i}{2}\right) (l_i - l_0)^2 + \sum_{\text{ângulos}} \left(\frac{b_i}{2}\right) (\theta_i - \theta_0)^2 + \sum_{\text{ângulos}} \left(\frac{c_i}{2}\right) (1 + \cos(n\omega_i - n\omega_0)) + \sum_{vdw} (4\varepsilon_{ij}) \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6 \right] + \sum_{coulomb} (k) \left(\frac{q_i q_j}{r_{ij}}\right) \quad (5)$$

Os dois primeiros termos lidam com as deformações do comprimento de ligação l_i e do ângulo de ligação θ_i a partir de seus valores de equilíbrio l_o e θ_o , respectivamente. A parte harmônica destes termos, a_i e b_i , garantem a estrutura das moléculas, mas com a consequência de impedir a quebra de ligações químicas. O terceiro termo lida com as rotações das ligações químicas onde n é o número de máximos e mínimos em torção completa, ω_i é o ângulo diedro e γ_o é o ângulo de fase. Estes são os termos que descrevem as interações ligantes. Os dois últimos termos lidam com as interações entre pares de átomos i e j a uma distância r_{ij} não ligados, as interações de van der Waals expressadas aqui pelo potencial de Lennard-Jones e as interações eletrostáticas. No quarto termo, ϵ_{ij} é a profundidade potencial entre as barreiras repulsiva e atrativa e σ_{ij} é a distância finita na qual o potencial entre as duas partículas é zero. No quinto termo, q_i e q_j são as cargas pontuais de cada átomo e k é a constante na forma de $1/4\pi\epsilon_o$.

Existem vários campos de força desenvolvidos que podem ser colocados em diferentes classes: padrão, classe I, II e III¹². O campo de força padrão utiliza parâmetros atômicos genéricos para a maioria dos átomos na tabela periódica; a classe I usa funções harmônicas para alongamento e curvatura sem termos cruzados e interações de Lennard-Jones; a classe II é muito precisa, pois adiciona termos cruzados e potenciais exponenciais de interações de dispersão; e a classe III adiciona hiperconjugação e polarização às moléculas do sistema. Alguns exemplos de campos de força são mostrados na tabela 1, assim como a classe e aplicabilidade.

Quadro 1. Alguns exemplos de campos de força na literatura com suas siglas, classes e aplicabilidades diferentes.

Software	Interface	Licença	Referência
DREIDING	Padrão	Propósito geral	MAYO <i>et al.</i> , 1990
CHARMM	I	Proteínas, ácidos nucleicos, íons	BROOKS <i>et al.</i> , 1983
GROMOS	I	Proteínas, lipídios, açúcares	VAN GUNSTEREN; FAND; BERENDSEN, 1987
AMBER	I	Proteínas, ácidos nucleicos	WEINER <i>et al.</i> , 1984
OPLS	I	Líquidos orgânicos e iônicos	JORGENSEN, 1988
INTERFACE-PCFF	II	Nanoestruturas	HEINZ <i>et al.</i> , 2013
PIPF	III	Amidas, alcanos líquidos	XIE <i>et al.</i> , 2007

Fonte: Dubbeldam *et al.*(2019), adaptado.

2.3. Equações de Movimento

A DM é um método empírico, baseado na descrição da estrutura de átomos e moléculas através da mecânica molecular (MM), onde o comportamento dinâmico de um sistema é simulado através da integração numérica das Equações de movimento (6) e (7)³².

$$F = m \frac{d^2 r(t)}{dt^2} \quad (6) \qquad F = - \frac{\partial V(r)}{\partial r} \quad (7)$$

A integração numérica das equações de movimento pode ser feita através do método das diferenças finitas, que para um

conjunto de partículas, integra as equações de movimento de uma configuração inicial, repetidas vezes, em novas posições e velocidades gerando a trajetória do sistema, que é a mudança das coordenadas das partículas em função do tempo. Esse processo é realizado por algoritmos que dividem a integração em pequenos intervalos de tempo (Δt), permitindo desenvolver uma expressão que define as posições e as velocidades em pequenos incrementos de tempo ($t + \Delta t$)^{32,31}.

Um dos algoritmos mais utilizados é o *Verlet* devido a sua simplicidade⁴⁵. O início do algoritmo se dispõe da expansão em série de Taylor de duas quantidades: $r(t + \Delta t)$ (8) e $r(t - \Delta t)$ (9), que são as posições dos passos posteriores e anteriores, respectivamente.

(8)

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{F \Delta t^2}{m \cdot 2} + \dots$$

(9)

$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{F \Delta t^2}{m \cdot 2} + \dots$$

Somando as Equações (8) e 9 e isolando $r(t - \Delta t)$ obtém-se a Equação (10):

(10)

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + \alpha\Delta t^2$$

A partir da Equação 10 é possível proceder com a integração numérica e obter as velocidades das partículas do sistema. Integrando mais uma vez se obtém as posições das partículas do sistema, desse modo, também obtendo as energias cinética

e potencial. Com a nova energia potencial substitui-se novamente na Equação de movimento (7) e o algoritmo é repetido por diversas vezes, gerando uma trajetória do sistema, que são todas as velocidades e posições das partículas ao longo do tempo ³². Outros algoritmos de integração válidos são o *Beeman*, *leap-frog* e *Velocity Verlet* ^{32;31}..

2.4. Etapas de Simulação de Dinâmica Molecular

Uma simulação por DM é, normalmente, subdividida em três etapas (minimização, equilíbrio e produção), conforme veremos a seguir.

2.4.1. Etapa de Minimização

A etapa de minimização de energia ou otimização de geometria é a primeira etapa em uma DM e tem por finalidade encontrar uma conformação com a menor energia potencial possível para o sistema, o que pode ser expresso por um mínimo local em uma curva de superfície de energia potencial (SEP) para um conjunto de partículas, assim eliminando maus contatos das estruturas do sistema que será simulado ^{33; 32}.

Para esta etapa, os principais algoritmos utilizados para realização da minimização de energia são: *Steepest Descent* (SD)⁴, *Newton-Raphson* (NR)²² e *Conjugate gradient* (CG)¹⁹. Todos eles têm como objetivo ajustar as posições do conjunto atômico do sistema relaxando ligações, ângulos e interações entre as partículas. Desse modo, o sistema está pronto para a DM³².

O método SD é robusto e garante uma boa aproximação do mínimo, contudo, a taxa de convergência diminui próximo deste. Portanto, normalmente se faz uma primeira mi-

nimização com o método SD e, em seguida, realiza-se outra minimização com o CG, pois este possui capacidade de convergência maior²¹. Entretanto, o método SD é o mais fácil de se implementar computacionalmente¹. Nele um determinado vetor r com coordenadas $3N$ sofre um deslocamento inicial h_0 . A partir disso, as novas posições são calculadas através da Equação (11), onde r_{n+1} e r_n são as posições final e inicial, respectivamente. F_n é a força expressa como a energia potencial U na Equação 5, $\max(|F_n|)$ é a maior força possível em cada partícula e h_n é o deslocamento máximo executado. Assim F_n e U são calculados continuamente para as novas posições. Se $U_{n+1} < U_n$ as posições ou geometria do sistema são aceitas e se $U_{n+1} \geq U_n$ elas são rejeitadas. Esse processo de contínua avaliação se repete até a especificação do operador ou quando os cálculos não convergirem mais, atingindo o valor máximo de F_n menor que o especificado¹.

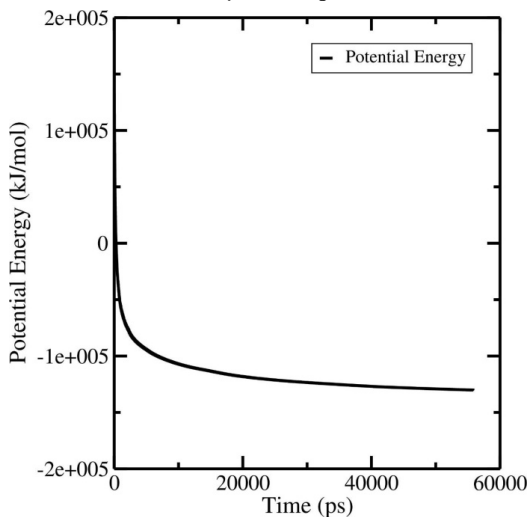
$$(11) \quad r_{n+1} = r_n + \frac{F_n}{\max(|F_n|)} h_n$$

Conforme podemos observar na Figura 2, a etapa de minimização de energia ocorre por uma diminuição na energia potencial do sistema à medida que o tempo aumenta, de acordo com a obtenção da estrutura mais otimizada obtida pelo algoritmo de minimização. Ao atingir o mínimo, será garantido a seguinte condição da Equação (12), pois neste ponto a inclinação da reta tangente é nula e a concavidade gráfica é positiva:

$$(12) \quad \left(\frac{dE}{dR}\right) = 0 \text{ e } \left(\frac{d^2E}{dR^2}\right) > 0 \quad (12)$$

Onde E , neste caso, refere-se a energia e, R , denota uma distância que otimiza a estrutura da espécie. O ponto ao qual as condições acima são vigentes é conhecido como ponto estacionário e, podem existir mais de um ponto estacionário.

Figura 2. Gráfico de minimização de energia



Fonte: autores.

A Figura 3 mostra uma minimização de energia obtida pela software GROMACS utilizando o algoritmo SD. Conforme podemos observar, há uma diminuição na energia potencial a cada etapa, até que haja uma estabilização na energia potencial (para este sistema ocorre em $E_{\text{pot}} = -1.05271e+05$) e o sistema estará otimizado.

Figura 3. Minimização de energia obtida pelo software GROMACS utilizando o algoritmo SD.

```
Step=43705, Dmax= 9.9e-04 nm, Epot= -1.05261e+05 Fmax= 7.19511e+02, atom= 3332
Step=43706, Dmax= 1.2e-03 nm, Epot= -1.05261e+05 Fmax= 7.29110e+02, atom= 3332
Step=43708, Dmax= 7.1e-04 nm, Epot= -1.05262e+05 Fmax= 1.41111e+02, atom= 3332
Step=43709, Dmax= 8.5e-04 nm, Epot= -1.05262e+05 Fmax= 9.85192e+02, atom= 3332
Step=43710, Dmax= 1.0e-03 nm, Epot= -1.05263e+05 Fmax= 3.49684e+02, atom= 3332
Step=43712, Dmax= 6.2e-04 nm, Epot= -1.05263e+05 Fmax= 4.81308e+02, atom= 3332
Step=43713, Dmax= 7.4e-04 nm, Epot= -1.05263e+05 Fmax= 4.99394e+02, atom= 3332
Step=43714, Dmax= 8.9e-04 nm, Epot= -1.05264e+05 Fmax= 5.82557e+02, atom= 3332
Step=43715, Dmax= 1.1e-03 nm, Epot= -1.05264e+05 Fmax= 7.13150e+02, atom= 3332
Step=43716, Dmax= 1.3e-03 nm, Epot= -1.05264e+05 Fmax= 8.45324e+02, atom= 3332
Step=43718, Dmax= 7.7e-04 nm, Epot= -1.05265e+05 Fmax= 9.09935e+01, atom= 3332
Step=43719, Dmax= 9.2e-04 nm, Epot= -1.05265e+05 Fmax= 1.83315e+03, atom= 3332
Step=43720, Dmax= 1.1e-03 nm, Epot= -1.05266e+05 Fmax= 3.18099e+02, atom= 3332
Step=43722, Dmax= 6.6e-04 nm, Epot= -1.05266e+05 Fmax= 4.89844e+02, atom= 3332
Step=43723, Dmax= 7.9e-04 nm, Epot= -1.05266e+05 Fmax= 4.78600e+02, atom= 3332
Step=43725, Dmax= 4.8e-04 nm, Epot= -1.05267e+05 Fmax= 1.01976e+02, atom= 3332
Step=43726, Dmax= 5.7e-04 nm, Epot= -1.05267e+05 Fmax= 5.93568e+02, atom= 3332
Step=43727, Dmax= 6.9e-04 nm, Epot= -1.05268e+05 Fmax= 2.41484e+02, atom= 3332
Step=43728, Dmax= 8.2e-04 nm, Epot= -1.05268e+05 Fmax= 7.59358e+02, atom= 3332
Step=43729, Dmax= 9.9e-04 nm, Epot= -1.05268e+05 Fmax= 4.44152e+02, atom= 3332
Step=43731, Dmax= 5.9e-04 nm, Epot= -1.05269e+05 Fmax= 2.78327e+02, atom= 3332
Step=43732, Dmax= 7.1e-04 nm, Epot= -1.05269e+05 Fmax= 5.90140e+02, atom= 3332
Step=43733, Dmax= 8.5e-04 nm, Epot= -1.05269e+05 Fmax= 4.50882e+02, atom= 3332
Step=43735, Dmax= 5.1e-04 nm, Epot= -1.05269e+05 Fmax= 1.73169e+02, atom= 3332
Step=43736, Dmax= 6.1e-04 nm, Epot= -1.05270e+05 Fmax= 5.75168e+02, atom= 3332
Step=43737, Dmax= 7.4e-04 nm, Epot= -1.05270e+05 Fmax= 3.23124e+02, atom= 3332
Step=43738, Dmax= 8.8e-04 nm, Epot= -1.05270e+05 Fmax= 7.53353e+02, atom= 3332
Step=43739, Dmax= 1.1e-03 nm, Epot= -1.05271e+05 Fmax= 5.41025e+02, atom= 3332
Step=43741, Dmax= 6.4e-04 nm, Epot= -1.05271e+05 Fmax= 2.36336e+02, atom= 3332
Step=43751, Dmax= 1.5e-06 nm, Epot= -1.05271e+05 Fmax= 2.34588e+02, atom= 3332
Step=43752, Dmax= 1.8e-06 nm, Epot= -1.05271e+05 Fmax= 2.32556e+02, atom= 3332
Energy minimization has stopped, but the forces have not converged to the
requested precision Fmax < 10 (which may not be possible for your system). It
stopped because the algorithm tried to make a new step whose size was too
small, or there was no change in the energy since last step. Either way, we
regard the minimization as converged to within the available machine
precision, given your starting configuration and EM parameters.

Double precision normally gives you higher accuracy, but this is often not
needed for preparing to run molecular dynamics.

writing lowest energy coordinates.

Steepest Descents converged to machine precision in 43753 steps,
but did not reach the requested Fmax < 10.
Potential Energy = -1.0527109e+05
Maximum force = 2.3458801e+02 on atom 3332
Norm of force = 4.0326851e+00

GROMACS reminds you: "Fresh Air, Green Hair" (Frank Black)
```

Fonte: autores.

2.4.2. Etapa de Equilíbrio

A segunda etapa em uma DM consiste na busca de tornar as condições do sistema o mais próximo possível das condições experimentais do sistema real. Inicialmente, o sistema encontra-se em repouso, sem vibração (em teoria) e, por conseguinte, a temperatura pode ser considerada nula. Portanto, equilibrar propriedades como a temperatura e a pressão são cruciais em uma simulação por DM.

O *ensemble* tradicional utilizado em DM é o microcanônico (N, V e E constantes), enquanto que o *ensemble* mais comum para o método de MC é o canônico (N, V e T constantes). No entanto, no limite termodinâmico (quando o número de partículas e o volume são infinitos, mas a densidade é constante), pode-se realizar uma equivalência entre *ensembles* e, dessa maneira, adaptar um *ensemble* a outro. Ademais, segundo a hipótese ergódica, que é um dos axiomas fundamentais da ME, a média temporal da propriedade macroscópica ζ equivale a média desta no *ensemble*^{25;38}. Portanto, uma DM poderá ser realizada em outros *ensembles*, além do microcanônico, por exemplo: isobárico-isotérmico (N, P e T constantes), grand-canônico (μ, V e T constantes) ou, no *ensemble* canônico (N, V e T constantes)²⁵.

Como a maioria dos experimentos químicos ocorrem com número de partículas N , pressão P e temperatura T fixos, o *ensemble* propício para este caso mostra-se ser o isobárico-isotérmico. Contudo, por ser mais simples controlar o volume do que a pressão (convém lembrar que o volume da caixa de simulação é fixo), normalmente realiza-se a equilibração em duas etapas, sendo a primeira no *ensemble* NVT e, consecutivamente, a outra no *ensemble* NPT²⁵. Para controlar a tem-

peratura e a pressão durante a evolução temporal do sistema, torna-se extremamente indispensável a utilização de termostatos e barostatos, respectivamente, de tal maneira que a T e a P sejam mantidas fixas. No *ensemble* NVT a T deverá ser mantida constante pelo uso de termostatos e a energia livre de Helmholtz é um mínimo. Não obstante, tanto a T como a P deverão ser constantes no *ensemble* NPT, sendo necessário, para isto, o uso do termostato e barostato e, neste caso a energia livre de Gibbs é um mínimo.

Existem diversos termostatos descritos na literatura para controle da T , como: Andersen³, Berendsen⁶, v-rescale⁹, Nosé-Hoover²⁶, assim como, também existem, diversos barostatos para controle da P , entre estes: Berendsen⁶, Parrinello-Rahman³⁵ e Nosé-Hoover²⁶. As Equações (13) e (14) abaixo referem-se aos termostato e barostato de Berendsen, respectivamente²⁶:

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_0 - T(t)) \quad (13)$$

$$\frac{dP(t)}{dt} = \frac{1}{\tau_P} (P_0 - P(t)) \quad (14)$$

Em que T_0 e $T(t)$ são as temperaturas do banho térmico e do sistema, respectivamente, e τ é uma constante de amortecimento para o termostato. Semelhantemente, P_0 e $P(t)$ simbolizam as pressões do barostato e do sistema, respectivamente, enquanto que τ_P denota a constante de tempo do barostato.

Considerando-se os *ensembles* canônico (N, V, T) e isobárico-isotérmico (N, P, T), a propriedade macroscópica ζ (energia interna, pressão, entropia etc) a qual se está interes-

sado poderá ser calculada com base na função de partição ($Q(N, V, T)$) e ($\Delta(N, P, T)$) do respectivo *ensemble* ao qual foi feita referência na última etapa da simulação computacional (etapa de produção).

A Tabela 2 mostrada em seguida refere-se aos *ensembles* supramencionados e, para o *ensemble* NVT, o potencial termodinâmico mais apropriado é a energia livre de Helmholtz ($A = U - TS$), pois esta é válida para sistemas isovolumétricos e, *ad hoc*, destina-se para condições de N, V e T constantes. No entanto, para o *ensemble* NPT, o potencial termodinâmico mais adequado é a energia livre de Gibbs ($G = H - TS$), pois é aceitável à processos isobáricos e destina-se para condições de N, P e T constantes^{16; 22}.

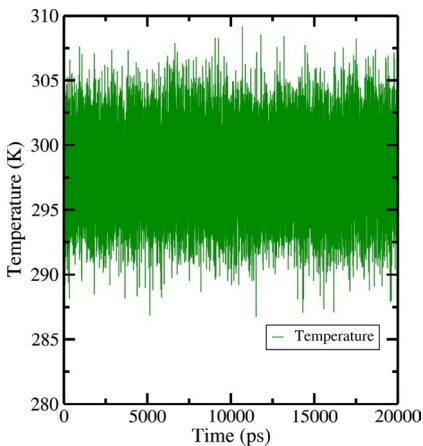
De acordo com a Figura 4, podemos observar a equilibração da temperatura realizada no *ensemble* canônico e, na Figura 5, é possível analisar a equilibração da pressão no *ensemble* isobárico-isotérmico e, por conseguinte, a equilibração da densidade (lembre que $\rho = N/V$ e $P = F/A$, sendo que $V = A \cdot h$, logo ρ é função de P) também ocorre, pois a densidade é função da pressão. Conforme discutido anteriormente, a equilibração da T e da P ocorre em duas etapas consecutivas e são mostradas nas Figuras 4 e 5.

Quadro 2. Ensembles canônico e isobárico-isotérmico para cálculo de propriedades físicas de interesse baseando-se na função de partição.

Ensemble	Função de partição	Propriedade ζ	Equação
Canônico	$Q(N, V, T)$	A	$A = -k_B T \ln Q$
		U	$U = k_B T^2 \left(\frac{\partial \ln \ln Q}{\partial T} \right)_{N,V}$
		P	$P = k_B T \left(\frac{\partial \ln \ln Q}{\partial V} \right)_{N,T}$
		μ	$\mu = -k_B T \left(\frac{\partial \ln \ln Q}{\partial N} \right)_{V,T}$
		C_V	$C_V = \left(\frac{\partial U}{\partial T} \right)_V$
		S	$S = k_B \ln \ln Q + k_B T \left(\frac{\partial \ln \ln Q}{\partial T} \right)_{V,N}$
Isobárico-Isotérmico	$\Delta(N, P, T)$	G	$G = -k_B T \ln \Delta$
		V	$V = -k_B T \left(\frac{\partial \ln \ln \Delta}{\partial P} \right)_{N,T}$
		S	$S = k_B \ln \ln \Delta + k_B T \left(\frac{\partial \ln \ln \Delta}{\partial T} \right)_{N,P}$
		μ	$\mu = -k_B T \left(\frac{\partial \ln \ln \Delta}{\partial N} \right)_{P,T}$

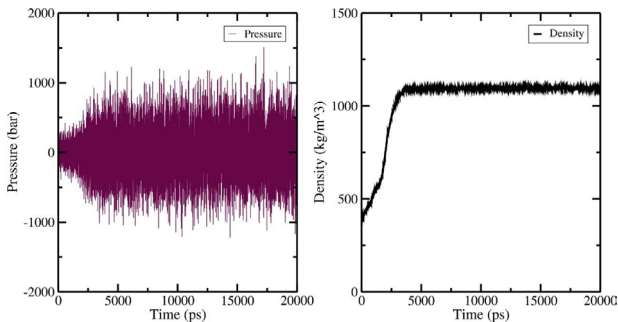
Fonte: Os autores.

Figura 4. Gráfico de estabilização da temperatura no *ensemble* NVT.



Fonte: autores.

Figura 5. Gráfico de estabilização da pressão e densidade no *ensemble* NPT.



Fonte: autores.

2.4.3. Etapa de Produção

Realizadas a minimização de energia do sistema através da otimização geométrica, bem como a equilibração deste deixando as condições de T e P o mais próximo possível do modelo experimental, então agora é necessário movimentar as partículas do sistema para se construir as trajetórias destas e, finalmente, realizar as análises possíveis e necessárias para fins de estudo.

Conforme supracitado no subtópico 2.1, conhecido $x(t)$ pela segunda lei de Newton ou pelas equações de Hamilton, então algoritmos específicos são necessários para se realizar iterativamente o cálculo de $x(t + i\delta t)$, com i variando de 1 até n , onde n é o número de passos da simulação. Entre os principais algoritmos utilizados para a construção da trajetória das espécies, temos: *Verlet*, *Leap-Frog* e *Beeman*, como visto no subtópico 2.3.

3. Análise dos resultados

3.1. Energia Livre de Interação

A variação da energia livre de Gibbs ΔG de um sistema pode ser utilizada como um descritor para a avaliação de interações e reações químicas³³. Métodos de cálculo de ΔG como a Perturbação de Energia Livre (FEP) e a Integração Termodinâmica (TI) foram desenvolvidos por Kirkwood em 1935 e formalizados por Zwanzig em 1954. Desde então, os métodos de cálculo de ΔG foram utilizados para obter valores pK_a , cálculo de propriedades termodinâmicas e obtenção de superfícies de energia para moléculas orgânicas e reações bioquímicas²⁴.

Um dos métodos de cálculo de ΔG mais comuns é a FEP. Consideram-se dois sistemas associados A e B, em que λ é magnitude de pequenos intervalos de perturbação que convertem A ($\lambda = 0$), o sistema inicial e não perturbado, em B ($\lambda = 1$), o sistema final e já totalmente perturbado^{1,33}. A diferença entre A e B é dada pela equação (15).

$$G_A - G_B = \Delta G = - \sum_{\lambda=0}^1 RT \ln \langle \exp(-\Delta H/RT) \rangle_{\lambda} \quad (15)$$

Sendo o logaritmo natural o valor médio da população de moléculas de um sistema e $\Delta H = H_A - H_B$. A partir da variação de λ pode-se aplicar Equação (15) e obter o ΔG ³³.

Outro método de cálculo da ΔG é a TI, que utiliza integração numérica, semelhante ao caso discutido no subtópico de equações de movimento 2.3, substituindo o somatório da Equação (15) por uma integral como mostrado na Equação (16)³³.

$$\Delta G = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H}{\partial \lambda} \right\rangle \quad (16)$$

3.2. RMSD

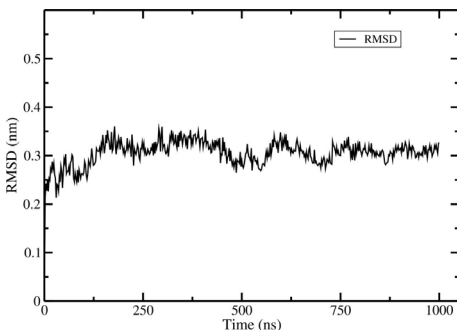
O *RMSD* (do inglês, *Root Square Mean Deviation*) analisa a mobilidade estrutural de um sistema de átomos e moléculas durante um tempo específico de simulação³¹. Frequentemente utilizado para sistemas biomoleculares, o conjunto de átomos escolhido para a análise do sistema são geralmente carbonos α da estrutura de proteínas. O cálculo do *RMSD* produz resultados analisados em um gráfico de $\langle RMSD \rangle$ por tempo que

permite visualizar mudanças na mobilidade ao longo da simulação. Pequenas variações em torno do ponto médio indicam pouca mobilidade, e grandes variações indicam grande mobilidade. O valor $\langle RMSD \rangle$ é calculado através da Equação (17), onde N é o número de partículas do sistema, $r_j(t_i)$ é a distância da partícula j em determinado momento da simulação i , e r_o é uma distância tomada como referência³¹.

$$\langle RMSD \rangle(t) = \left\{ \frac{1}{N} \sum_{j=1}^N [r_j(t_i) - r_o] \right\}^{1/2} \quad (17)$$

O gráfico $RMSD$ da Figura 6, produzido pelo *software Grace* (GRACE, 2005), é de uma proteína oxidoreductase de código *PDB 3LFM* (HAN et al., 2010). A série temporal mostrada varia até $\sim 0,35$ nm ($3,5 \text{ \AA}$), exibindo um bom equilíbrio estrutural durante o tempo de simulação escolhido.

Figura 5. Gráfico do $RMSD$ (em nm) em função do tempo (em ns) da proteína oxidoreductase (PDB ID: *3LFM*).



Fonte: autores.

3.3. RDF [$g(r)$]

O acrônimo RDF vem do inglês (*Radial Distribution Function*) e traduzido significa Função de Distribuição Radial, que representa uma medida probabilística de se encontrar uma partícula α entre as distâncias r e $r + dr$ de uma partícula β .

De acordo com N. Morgon e K. Coutinho:

A função de distribuição de pares, $g_{\alpha\beta}(r)$, fornece a probabilidade de se encontrar uma partícula β em uma camada esférica de espessura dr , a uma distância r da partícula α , cuja posição é considerada a origem, relativamente à probabilidade esperada em uma distribuição completamente aleatória na mesma densidade (p. 443, 2007).

Assim sendo, pode-se observar que a função de distribuição radial de pares é uma ferramenta bastante poderosa para se analisar estruturas de líquidos, amorfos e cristais³⁷.

Para duas espécies α e β , com α centrada na origem ($r = 0$), temos que a $g_{\alpha\beta}(r)dr$ na Equação (18) denota a probabilidade que a partícula β tem de ser encontrada no elemento de volume dr da partícula α .

$$g_{\alpha\beta}(r) = \frac{V N_{\alpha\beta}(r, r + \Delta r)}{N 4\pi r^2 \Delta r} = \frac{N_{\alpha\beta}(r, r + \Delta r)}{4\pi r^2 \rho_{\beta} \Delta r} \quad (18)$$

$$\rho_{\beta} = \frac{N}{V} \quad (19)$$

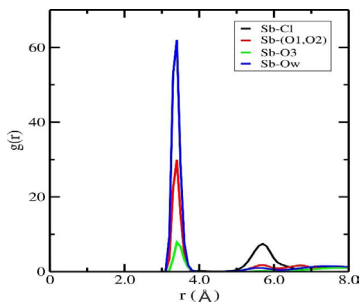
Em que $N_{\alpha\beta}(r, r + \Delta r)$ representa o número médio de espécies β no interior do elemento de volume

dr (entre r e $r + \Delta r$); e $\rho\beta$ na Equação (19) denota a densidade numérica média de espécies β para um fluido isotrópico em solução e $4\pi r^2 \Delta r$ é o elemento de volume¹⁷.

A Figura 7 mostra os valores de $g(r)$ para um sistema constituído por antimônio (Sb) eletricamente neutro em solvente eutético (constituído por cloreto de colina e etilenoglicol, 1ChCl:2EG) com água. Para este sistema, os átomos de oxigênio da molécula de etilenoglicol foram simbolizados por O1 e O2, enquanto o átomo de oxigênio da colina está representado por O3 e, ainda, o átomo de oxigênio da água mostra-se como Ow.

Analisando a Figura 7, é possível constatar que a principal interação ocorre entre o antimônio com o oxigênio da água (Ow), que possui maior valor de $g(r)$. Em seguida, temos a segunda principal interação entre o Sb com os átomos de oxigênio do etilenoglicol (O1,O2). As demais interações mostraram baixos valores de $g(r)$ se comparadas as interações Sb-Cl e Sb-(O1,O2) e, por conseguinte, não faz muito sentido discuti-las.

Figura 7. Gráfico de RDF [$g(r)$].



Fonte: autores.

3.3.1. $g(r)\rho$

Embora o $g(r)$ seja uma boa opção para a análise de interações em líquidos, já foi constatado que esta função de distribuição não permite uma boa análise de interações quando as espécies constituintes do sistema possuem densidades numéricas de átomos muito diferentes, levando a erros na representatividade do comportamento do sistema. Desse modo, faz-se imprescindível uma correção neste erro. Com a finalidade de se corrigir o problema discutido, multiplica-se $g(r)$ pela densidade numérica média de átomos observados (ρ), podendo-se garantir, assim, certa fidedignidade ao comportamento analisado³⁰; ³³.

Segundo Chandler, $g(r)\rho$ representa a densidade de probabilidade condicional de que uma partícula β seja encontrada em r , dado que a outra está na origem ($r = 0$)¹⁰. Logo, esta quantidade nos dá uma estimativa da densidade média de partículas em r .

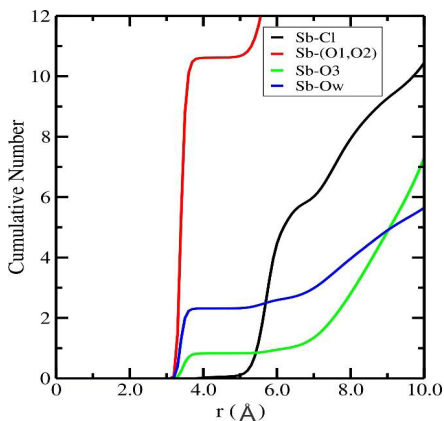
3.4. CCN [$N(r)$]

O número de coordenação cumulativo (*Cumulative Coordination Number*, em inglês) pode ser compreendido como o número de espécies β em torno do centro de referência α e pode ser obtido pela integração da função de distribuição radial de pares, conforme podemos ver na Equação (20) abaixo:

$$(20)$$
$$N(r) = N_{\alpha\beta}(r) = \int_{r_1}^{r_2} 4\pi r^2 \rho_{\beta} g_{\alpha\beta}(r) dr = 4\pi\rho_{\beta} \int_{r_1}^{r_2} r^2 g_{\alpha\beta}(r) dr$$

Nesse caso, r_1 e r_2 denotam os pontos de mínimo que delimitam o máximo no gráfico de $g(r)$. Assim sendo, $N(r)$ pode ser entendido como o número de coordenação³¹. Na Figura 8, é possível observar o gráfico de CCN para o sistema descrito no subtópico 3.3. Embora a principal interação seja entre Sb-Ow, conforme concluído pelo $g(r)$, é possível observar que um maior número de espécies etilenoglicol coordenam o antimônio. Nesse sentido, o CCN permite-nos inferir estatisticamente (lembre que o $N(r)$ é obtido pela integração do $g(r)$, que é uma grandeza probabilística) sobre a quantidade de espécies que coordenam um determinado centro de referência.

Figura 8. Gráfico de CCN [$N(r)$].



Fonte: autores.

4. Softwares utilizados

O quadro 2 mostra alguns dos principais softwares utilizados em DM, sendo alguns gratuitos e outros licenciados. Conforme podemos ver, alguns softwares levam o mesmo nome do campo de força ao qual utilizam, entre eles, temos o CHARMM⁸ e o AMBER⁴⁵. Além disso, alguns softwares em DM possuem grande versatilidade em poder utilizar diversos campos de força distintos, como o GROMACS (*GROningen Machine for Chemical Simulation*)⁴¹, além de poder analisar uma diversidade de sistemas. O AMBER (*Assisted Model Building with Energy Refinement*), *software* e campo de força, tem grande aplicação no estudo de biomoléculas, como proteínas e ácidos nucleicos²⁶.

Quadro 2. Softwares mais utilizados em DM.

Software	Interface	Versão	Licença	Referência
AMBER	CLI	21.0	Proprietário	WEINER; KOLLMAN, 1981
CHARMM	CLI/GUI	47.0	Proprietário	BROOKS et al., 1983
DESMOND	GUI	-	Acadêmico	BOWERS et al., 2006
GROMACS	CLI	2022	Open Source	VAN DER SPOEL et al., 2005
NAMD	GUI	3.0	Acadêmico	PHILLIPS et al., 2005

Fonte: Selvaraj et al. (2018), adaptado.

5. Considerações finais

A DM é um método de grande alcance, sendo uma das técnicas teóricas mais citadas na literatura. Abrangendo sistemas macromoleculares como biomembranas e proteínas à interfaces e filmes de polímeros, as simulações DM apresentam uma

extensa variedade de aplicações. Neste capítulo, discutimos sobre o contexto-histórico da DM (1.1), mostrando sua conceituação (1.2) e importância e aplicações (1.3) na ciência. A fundamentação das simulações DM foi apresentada nos tópicos de pressupostos (2.1), campo de força (2.2) e equações de movimento (2.3), tomando a cautela de se referenciar trabalhos de prestígio na literatura. As etapas de simulação (2.4), minimização, equilíbrio e produção foram exemplificadas com gráficos e dados reais afim de demonstrar os resultados obtidos na DM. Por fim, a análise dos resultados mostradas no capítulo como energia livre de interação, RMSD, RDF, $g(r)\rho$ e CCN foram discutidas através de gráficos reais e com o intuito de incentivar o leitor na execução de tais análises. Agora, o leitor, capacitado por esta introdução a DM, pode utilizar o conhecimento aprendido e a bibliografia dada no capítulo para realizar as próprias simulações e conduzir sua própria pesquisa.

Referências

1. ABRAHAM, M. et al. GROMACS User Manual version 2018.4. www.Gromacs.org, 2018.
2. ALLEN M. P., TILDESLEY D. J. **Computer Simulation of Liquids**. New York: Oxford University Press, 1989.
3. ANDERSEN, H. C. Molecular dynamics simulations at constant pressure and/or temperature. **The Journal of Chemical Physics**, v. 72, n. 4, p. 2384–2393, 15 fev. 1980.
4. ARFKEN, G.B.; WEBER H.J.; HARRIS F.E.; **Mathematical Methods for Physicists**. Waltham: Elsevier, 2013.
5. BACHRACH, S. M. **Computational Organic Chemistry**. 2. ed. New Jersey: John Wiley & Sons, 2014.

6. BERENDSEN, H. J. C. et al. Molecular dynamics with coupling to an external bath. **The Journal of Chemical Physics**, v. 81, n. 8, p. 3684–3690, 15 out. 1984.
7. BOWERS, K. J.; DROR, R. O.; SHAW, D. E. The midpoint method for parallelization of particle simulations. **Journal of Chemical Physics**, v. 124, n. 18, p. 184109, 2006.
8. BROOKS, B. R. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. **Journal of Computational Chemistry**, v. 4, n. 2, p. 187–217, 1983.
9. BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. **The Journal of Chemical Physics**, v. 126, n. 1, p. 014101, 7 jan. 2007.
10. CHANDLER, D; **Introduction to Modern Statistical Mechanics**. New York: Oxford University Press, 1987.
11. DUAN, Y.; KOLLMAN, P. A. Pathways to a Protein Folding Intermediate Observed in a 1-Microsecond Simulation in Aqueous Solution. **Science**, v. 282, n. 5389, p. 740–744, 23 out. 1998.
12. DUBBELDAM, D. et al. Design, Parameterization, and Implementation of Atomic Force Fields for Adsorption in Nanoporous Materials. **Advanced Theory and Simulations**, v. 2, n. 11, p. 1900135, 23 nov. 2019.
13. GOLDSTEIN, H. **Classical mechanics**. 1. ed. Nova York: Addison-Wesley, 1980.
14. GUIDORIZZI, H. L. **Um curso de cálculo**, v. 1. 5. ed. Rio de Janeiro: LTC, 2013.
15. HAN, Z. et al. Crystal structure of the FTO protein reveals basis for its substrate specificity. **Nature**, v. 464, n. 7292, p. 1205–1209, 7 abr. 2010.
16. HANSEN, J. P.; MCDONALD, I. R.; **Theory of Simple Liquids**, London: Academic Press, 1986.
17. HEERMANN D.W. **Computer Simulation Methods in Theoretical Physics**. 2. ed. Berlin: Springer, 1990.
18. HEINZ, H. et al. Thermodynamically Consistent Force Fields for the Assembly of Inorganic, Organic, and Biological Nanos-

- tructures: The INTERFACE Force Field. **Langmuir**, v. 29, n. 6, p. 1754–1765, 12 fev. 2013.
19. HESTENES, M. R.; STIEFEL, E.; Methods of Conjugate Gradients for Solving Linear Systems. **Journal of Research of the National Bureau of Standards**, v. 49, n. 6, p. 409–436, 1952.
 20. HOLLINGSWORTH, S. A.; DROR, R. O. Molecular Dynamics Simulation for All. **Neuron**, v. 99, n. 6, p. 1129–1143, set. 2018.
 21. JENSEN, F.; **Introduction to Computational Chemistry**. 2. ed. Chichester: John Wiley & Sons, 2007.
 22. JENSEN, J. H. **Molecular Modeling Basics**. Boca Raton: CRC Press, 2010.
 23. JORGENSEN, W. L.; TIRADO-RIVES, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. **Journal of the American Chemical Society**, v. 110, n. 6, p. 1657–1666, 1 mar. 1988.
 24. JORGENSEN, W. L.; THOMAS, L. L. Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. **Journal of Chemical Theory and Computation**, v. 4, n. 6, p. 869–876, 1 jun. 2008.
 25. LEACH, A. R. **Molecular Modelling Principles and Applications**. 2. ed. Harlow: Prentice Hall, 2001.
 26. LESZCZYNSKI, Jerzy. **Handbook of computational chemistry**. Dordrecht: Springer, 2012.
 27. MAGINN, E. J.; ELLIOTT, J. R. Historical Perspective and Current Outlook for Molecular Dynamics As a Chemical Engineering Tool. **Industrial & Engineering Chemistry Research**, v. 49, n. 7, p. 3059–3078, 7 abr. 2010.
 28. MAIA, N. M. M. **Introdução à Dinâmica Analítica**. Lisboa: IST Press, 2000.
 29. MAYO, S. L.; OLAFSON, B. D.; GODDARD, W. A. DREIDING: a generic force field for molecular simulations. **The Journal of Physical Chemistry**, v. 94, n. 26, p. 8897–8909, 1 dez. 1990.

30. MIGLIORATI, V.; SESSA, F.; D'ANGELO, P. Deep eutectic solvents: A structural point of view on the role of the cation. **Chemical Physics Letters**, v. 737, p. 100001, 2019.
31. MORGON, N. H; COUTINHO, K. R. **Métodos de química teórica e modelagem molecular**. [S.l: s.n.], 2007.
32. NAMBA, A. M.; DA SILVA, V. B.; DA SILVA, C. H. T. P. Dinâmica molecular: teoria e aplicações em planejamento de fármacos. **Eclética Química**, v. 33, n. 4, p. 13–24, dez. 2008.
33. OLIVEIRA A. M., **Introdução a Modelagem Molecular para Química, Engenharia e Biomédicas: Fundamentos e Exercícios**. Curitiba: Appris Editora, 2018.
34. PAQUET, E.; VIKTOR, H. L. Molecular Dynamics, Monte Carlo Simulations, and Langevin Dynamics: A Computational Review. **BioMed Research International**, v. 2015, p. 1–18, 2015.
35. PARRINELLO, M.; RAHMAN, A. Polymorphic transitions in single crystals: A new molecular dynamics method. **Journal of Applied Physics**, v. 52, n. 12, p. 7182–7190, 1 dez. 1981.
36. PETRENKO, R.; MELLER, J. Molecular Dynamics. In: **Encyclopedia of Life Sciences**. Major Reference Works. Chichester, UK: John Wiley & Sons, Ltd, 2010. PHILLIPS, J. C. et al. Scalable molecular dynamics with NAMD. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1781–1802, 2005.
37. RINO, J. P.; COSTA, B. V. **ABC da Simulação Computacional**. Brasil: Livraria da Física, 2013.
38. SELVARAJ, C. et al. Molecular dynamics simulations and applications in computational toxicology and nanotoxicology. **Food and Chemical Toxicology**, v. 112, p. 495–506, fev. 2018.
39. SUGITA, Y.; OKAMOTO, Y. Replica-exchange molecular dynamics method for protein folding. **Chemical Physics Letters**, v. 314, n. 1–2, p. 141–151, nov. 1999.
40. TURNER, P. J. **Grace**. Portland: 2005. Disponível: <https://plasma-gate.weizmann.ac.il/Grace/> em: 3 mar. 2022.

41. VAN DER SPOEL, D. et al. GROMACS: Fast, flexible, and free. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1701–1718, 2005.
42. VAN GUNSTEREN, W.; FAND, H. J.; BERENDSEN, C. (1987). Groningen molecular simulation (GROMOS) library manual, Biomos, Groningen. www.gromos.net, p. 1-221. 1987.
43. VERÍSSIMO DE OLIVEIRA, R. et al. Analysis of the behavior of Sn²⁺ and In³⁺ ions in DES and in water: A theoretical approach. **Journal of Molecular Liquids**, v. 353, p. 118774, 2022.
44. VERLET, L. Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. **Physical Review**, v. 159, n. 1, p. 98–103, 5 jul. 1967.
45. WEINER, P. K.; KOLLMAN, P. A. Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. **Journal of Computational Chemistry**, v. 2, n. 3, p. 287–303, 1981.
46. WEINER, S. J. et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. **Journal of the American Chemical Society**, v. 106, n. 3, p. 765–784, 1 fev. 1984.
47. XIE, W. et al. Development of a Polarizable Intermolecular Potential Function (PIPF) for Liquid Amides and Alkanes. **Journal of Chemical Theory and Computation**, v. 3, n. 6, p. 1878–1889, 23 nov. 2007.

UM ESTUDO SOBRE TERMODINÂMICA NO DOCKING MOLECULAR

Oseas Gomes

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Pedro João Samuel

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Bernardino Joaquim Caluaco

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Junilson Martinho Canjanja Madureira

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Aluisio Marques da Fonseca

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Regilany Paulo colares

Universidade da Integração Internacional
da Lusofonia Afro-Brasileira, Acarape-Ceará

Resumo:

Como a energia livre é função de estado, o ciclo termodinâmico tem sua relevância quando se estuda a simulação computacional, principalmente quando se refere a pontuações energéticas obtidas em ancoragem molecular. Portanto, o ciclo termodinâmico pode ser montado para se calcular a energia dos complexos envolvidos no sistema e auxiliar na contabilização real de todas as transformações e assim obter boa estimativa para a energia de interação.

Palavras-chave: ancoragem molecular, pontuação, estimativa de energia

1. A HISTÓRIA DA TERMODINÂMICA

A termodinâmica é a ciência que descreve as transformações termodinâmicas, faz parte da física clássica e, desde os tempos antigos vem estudando, com inúmeras tentativas de medir o calor, o trabalho, os processos que envolvem as variáveis de temperatura e o estado de energia, através da “*Termometria*.” A história da física e da química, na medida de sua evolução, foi ao longo do tempo, surgindo questionamentos e perguntas que precisavam ser respondidas de forma a completar as duas ciências, então surge a termodinâmica que vem contribuir grandemente com a mecânica clássica, quântica e estatística, o magnetismo e a velocidade química. Em termos práticos os estudos das máquinas a combustão, máquinas a vapores; (O que é uma raridade na ciência, a teoria vir em segundo plano e ser desenvolvida através da prática já existente, como é o caso dos motores a vapor que na época já existiam em pleno funcionamento. Por curiosidade e paixão só em 1824 é que se começaram as investigações através do engenheiro francês Sadi Carnot para introduzir o “conceito de eficiência termodinâmica”). E no campo de geração de eletricidade e em diversos campos como os de tecnologias da atualidade. Devido a sua relevância a termodinâmica tem contribuído de forma relevante não só para a física, química, mas para a ciência de forma geral.

Na verdade, a termodinâmica já era conhecida no antigo Egito, não como a conhecemos hoje. O calor estava ligado, a Mitologia. Já na filosofia ocidental, se formulou “a teoria dos quatro elementos;” água, terra, ar e fogo. E em seguida se postulou “a teoria do calórico” logo perto do fim da idade moderna e o começo da idade contemporânea, que consiste em o calor como um fluido imponderável. Essa tese foi formulada pelos

seguintes cientistas da época; Pierre Gassendi (astrónomo e filósofo francês), Lavoisier e Bertholet (químicos franceses), William Thomson (Lord Kelvin), William Cleghorn (Edimburgh, 1779). O que não demorou muito e em 1798 foi contestado pelo Conde Rumford o Benjamin Thompson em suas observações na perfuração em canos de canhões através de fricção e o calor gerado em suas observações forneceram argumentos contra a teoria do calórico. Mesmo assim, anos mais a frente Sadi Carnot vai usar essa teoria para introduzir o conceito de eficiência.

A história da termodinâmica surge no período muito conturbado no mundo de uma forma geral. Com o apoio da França às colônias americanas da América do Norte, que lutavam pela sua independência, fazendo com que a França contraísse uma enorme dívida no envio de soldados franceses a fim de ajudar as colônias inglesas, indo de encontro com os interesses da Grã-Bretanha, seu grande rival político e econômico naquele período. Além disso uma seca severa que se passava na França trouxe um desabastecimento de trigo e outros alimentos fazendo com que os franceses sofressem enormemente de fome causando enorme insatisfação com o governo francês, mesmo com o apoio da população a independência das colônias americanas, contudo porém, a insatisfação se voltou contra o governo e em 1789, deu-se início a Revolução Francesa. Além dos fatores econômicas e políticos, os franceses viram na independência das colônias inglesas a oportunidade de direitos mais igualitários entre as classes sociais e o fim dos privilégios feudais e a declaração do direito do cidadão. Neste período, além dos problemas internos, a França ainda tinha a Grã-Bretanha como sua rival econômica e política, passando por uma transformação industrial. Já a Grã-Bretanha vivia um outro momento mesmo com

as guerras. Lá, na metade do século XVII, iniciava a Revolução Industrial, atrelada ao modelo capitalismo.

A revolução industrial iniciou-se na Inglaterra muito por conta dos investimentos foi pela burguesia inglesa, que no primeiro momento era para atender a industrial têxtil e depois esses investimentos migraram para o desenvolvimento das máquinas a vapor, isso já no final do século XVII, por Thomas Newcomem e aprimorada na década de 1760 James Watt. A revolução industrial teve uma contribuição relevante, para o desenvolvimento da termodinâmica, porque neste período na Inglaterra acontecia a substituição da produção artesanal pela de maquinofatura e a troca de energia produzida pelo homem pela energia produzida a vapor, eólica e hidráulica e consequentemente o desenvolvimento de máquinas a vapor, da locomotiva e o telegrafo, um dos primeiros meios de comunicação instantâneo. Neste momento a França e a Inglaterra são duas potências em disputadas econômicas. A França estava muito atrasada no que se refere aos designs ingleses de máquinas a vapor, embora todas as máquinas a vapor fossem bastante rudimentares e com baixa eficiências, o que logo chamou atenção de Sadi Carnot.

Sadi Carnot, filho mais velho de Lazare Carnot, um revolucionário francês e ministro da guerra de Napoleão. Sadi vai integrar a elite da École Polytechnique após atingir seus dezesseis anos, o mínimo para a admissão. Seguindo alguns meses de preparação na Lycée Charlemagne. Seus estudos desde 1812 a 1814 se intensificaram em análise, mecânica, geometria descritiva e química, ministrados por um eminente corpo docente, incluindo alguns cientistas da época, Poisson, Gay-Lussac, Ampère e Arago. A partir de 1819, Carnot passou a se preocupar como projetar um motor a vapor, que já estava em uso

a muito tempo porem, de forma ineficientes. O que chamou a atenção de Carnot foram os designs franceses que tinham ficado obsoletos em relação aos avançados motores ingleses que eram importados pela França, isso depois da guerra. Os trabalhos científicos de Sadi Carnot incluem três: O Manuscrito Recherche, O Réflexions e as Notas Manuscritas. Fragmentos de matemática, traduções de dois dos artigos de Watt e notas de leitura de cursos de Matemática e Física. Esses foram os seus trabalhos que sobreviveram depois da sua morte. Embora seus artigos foram quase todos queimados em virtude de sua doença contagiosa, cólera. Mesmo assim, após sua morte, jornais da época citaram sua obra como: “marcantes por suas visões originais”. Pouco se deu importância aos trabalhos de Carnot que passaram oito anos esquecidos desde a sua publicação, até que então, após dois anos de sua morte (1834) – Clapeyron fornece o formalismo matemático e geométrico da teoria das máquinas térmicas de Carnot. Ele foi pioneiro na formulação do primeiro modelo de máquinas térmicas e apresentou o que hoje consideramos os fundamentos da segunda lei da termodinâmica. Apesar de suas experiências terem sido fundamentadas na teoria do calórico, vigente na época, muitos de seus resultados foram corretos. E, sendo assim ele ficou reconhecido até hoje pelos ciclos de Carnot, teorema de Carnot, e a máquina de Carnot. O pai da termodinâmica.

Rudolf Julius Emanuel Clausius; físico, alemão nascido na Prússia, era filho de um pastor protestante e professor. Estudou na universidade de Berlin e Halle (Doutorado). Fundador moderno da ciência da termodinâmica, que reafirmou o princípio de Carnot, aproximando-o ao trabalho de Joule (1850-1865). Conhecido como ciclo de Carnot, onde ele apresentou seu ar-

tigo *A teoria mecânica do calor*. Que formula a relação entre o fluxo de calor e o trabalho mecânico incluindo as discrepâncias que acreditavam existir entre o princípio de Carnot e o conceito de conservação da energia. Introduz o conceito e as ideias básicas da segunda lei da termodinâmica e o conceito de entropia (do grego, trope, transformação), sendo enunciado a segunda Lei da Termodinâmica, Teoria moderna do calor. Nos trabalhos de Clausius e de Kelvin tornou-se claro que o calor não era nenhum misterioso fluido, mas sim uma forma de energia, contribuindo para a formulação do segundo princípio da termodinâmica e contribuindo para o desenvolvimento da teoria cinética dos gases (1850). Em 1858, Clausius apresenta estudos de choques moleculares e o conceito de percurso livre médio. Já em 1865, a partir do teorema de Carnot, definiu uma nova grandeza: a entropia, com suas explicações matemáticas, e um dos seus princípios que a entropia nunca diminuiria em um processo físico real. Ele também apresentou uma síntese importante para a física, com duas leis importantes, que a energia do universo é constante e a entropia do universo tende ao máximo. Clausius também fez estudos sobre as leis da eletricidade e de magnetismo. Em 1877 fez uma nova abordagem para equações do eletromagnetismo e derivou uma fórmula para a força que atua entre duas cargas que se movem uma em relação a outra. No final de sua vida Clausius também foi nomeado reitor da universidade de Bonn, mesmo doente em virtude de sua participação na guerra, ele teve uma vida agitada e é um dos cientistas que não recebeu em vida os devidos créditos dos seus trabalhos e de sua genialidade tão importante para resolver os problemas mais desafiadores da física, como estabelecimento da lei da entropia.

Hermann Ludwig Ferdinand von Helmholtz, foi um físico e médico alemão que contribuiu para vários campos da ciência. Nasceu em Potsdam, Província de Brandemburgo, e quando jovem ele se interessou pelas ciências naturais, mas seu pai queria que ele estuasse medicina e Helmholtz obteve seu doutorado em medicina. Em fisiologia e psicologia, ele é conhecido por sua matemática do olho, teorias da visão, ideias sobre a percepção visual do espaço, pesquisa da visão de cores e sobre a sensação do tom, percepção do som e empirismo na fisiologia da percepção. Como filósofo, ele é conhecido por sua filosofia da ciência, ideias sobre a relação entre as leis da percepção e as leis da natureza, a ciência da estética e ideias sobre o poder civilizador da ciência. Na física, ele é conhecido por suas teorias sobre a conservação de energia, trabalho em eletrodinâmica, termodinâmica química e em uma base mecânica da termodinâmica. Em 1842 ele serviu no internado do hospital Charité por um ano. (Porque havia apoio financeiro aos estudantes de medicina). Com formação principalmente em fisiologia, Helmholtz escreveu sobre outros assuntos, que vão deste a física teórica, idade da terra e a origem do sistema solar. Ele também foi professor de anatomia na academia de artes em Berlin em 1848, professor de fisiologia na universidade prussiana, 1849 e professor titular de fisiologia e anatomia na universidade de Bonn. Helmholtz foi professor em outras universidades e no ano de 1871 ele aceitou seu último cargo em uma universidade em Berlin como professor de física.

Sua primeira conquista veio no ano de 1847 sobre a conservação de energia, seu trabalho foi escrito no meio do contexto de seus estudos médicos e filosóficos como já era de se esperar de Helmholtz. Seu trabalho de conservação de energia, surgiu quando ele estudava o metabolismo muscular. Ele ob-

servou que nenhuma energia é perdida no movimento muscular, motivado pela implicação de que não havia forças vitais necessárias para mover os músculos. Isso foi uma rejeição especulativa vigente na época, um paradigma na filosofia existente na fisiologia alemão. Baseando-se em trabalhos anteriores como de Sadi Carnot, Benoît Paul Émile Clapeyron e James Prescott Joule, ele postulou uma relação entre mecânica, luz, calor, eletricidade e magnetismo, tratando-os como uma única manifestação de força ou energia na sua terminologia atual, em 1847 ele publicou seus estudos em um livro. *Über die Erhaltung der Kraft* (On the Conservation of Force). Nas décadas de 1850 e 1860, com base nas publicações de William Thomson, Helmholtz e William Rankine, popularizou-se a ideia da morte térmica do universo. Na dinâmica dos fluidos, Helmholtz fez várias contribuições, incluindo os “teoremas de Helmholtz” para a dinâmica de vórtices em fluidos invíscidos. Além deste trabalho ele postulou muitos outros em outras áreas das ciências: fisiologia sensorial, óptica oftálmica, fisiologia nervosa, acústica e estética, eletromagnetismo embora, ele não tenha contribuindo tanto neste último trabalho, mas, um aluno seu, Heinrich Rudolf Hertz ficou muito famoso por demonstrar a radiação eletromagnética e em homenagem foi batizada como “equação de Helmholtz”.

Josian Willard Gibbs, seu nome é o mesmo do seu pai, (aboliconista, linguístico e teólogo), ele pertencia a uma família de ianques com ilustres clérigos e acadêmicos americanos. Seu nome é uma tradição herdada de um ancestral que foi secretário de uma Província da Baía de Massachusetts, Josian Willard. E o Gibbs de sua avó paterna, Mercy (Prescott) Gibbs. Foi um cientista americano que contribuiu para o desenvolvimento da

física, química e matemática. Desenvolveu trabalhos importantes para a termodinâmica e sua contribuição foi de extrema importância para transformar a físico-química em uma ciência indutiva rigorosa. Juntamente com os cientistas James Clerk Maxwell e Ludwig Boltzmann, eles postularam as explicações das leis da termodinâmica explicando como as consequências das propriedades estatísticas dos conjuntos dos possíveis estados de um sistema físico compostos por várias partículas. Com isso ele mesmo, criou a “mecânica estatística”. Gibbs também desenvolveu um trabalho sobre as “equações de Maxwell”, para solucionar problemas na óptica física. Como matemático ele criou o “cálculo vetorial moderno”. É importante salientar que neste período um outro cientista britânico Oliver Heaviside apresentou, um trabalho sobre cálculo vetorial semelhante. Ele morou por três anos na Europa e retornou aos EUA em Yale, onde recebeu o primeiro doutorado em engenharia americano, concedido pela A Yale University uma das três mais antiga e prestigiada do mundo. Ele foi professor de física matemática desta universidade de 1871 até sua morte em 1903. Ele também foi o primeiro cientista teórico americano a galgar o reconhecimento internacional, e foi muito elogiado por Albert Einstein, pelos seus trabalhos e foi considerado pelo próprio, uma das mentes, mas brilhantes dos EUA da América. Em 1901 ele recebeu da comunidade científica internacional em Londres a Medalha Copley Royal Society of London, por sua contribuição na física matemática. Embora seu trabalho mesmo sendo teórico trouxe um grande impacto internacional com suas ideias, mesmo assim Gibbs vivia uma vida solitária e tranquila. Ele nunca se casou, teve uma vida discreta e um caráter impecável segundo seus amigos e colegas de universidade.

Logo após sua formatura Gibbs se tornou tutor na universidade, por três anos onde ensinou latim e filosofia natural, (física) e em 1866 patenteou um projeto de “freios ferroviários”. Depois disso viajou para França onde assistiu palestras na Sorbonne e no Collège de France, ministradas por cientistas matemáticos ilustres como, Joseph Liouville e Michel Chasles, muitos já conhecidos na época. Tendo empreendido um regime de estudo punitivo, Gibbs pegou um resfriado grave e um médico, temendo a tuberculose, o aconselhou a descansar na Riviera, onde ele e suas irmãs passaram vários meses e onde ele se recuperou totalmente. Logo depois Gibbs foi para Berlin onde participou de palestras dos matemáticos Karl Weierstrass e Leopold Kronecker, e do químico Heinrich Gustav Magnus. Gibbs também teve acesso aos trabalhos dos físicos Gustav Kirchhoff e Hermann von Helmholtz, e do químico Robert Bunsen. Neste período, os acadêmicos alemães eram as principais autoridades nas ciências naturais, especialmente química e termodinâmica. Em seu retorno aos Yale Gibbs ensinou francês aos estudantes de engenharia e é possível que 1869 que ele tenha estudado um projeto a vapor, esse foi sua última investigação significativa em engenharia mecânica. Ele também publicou seu primeiro trabalho em 1873. Seus artigos sobre a representação geométrica de quantidades termodinâmicas; *Transactions of the Connecticut Academy*. Seus artigos colocou o uso de diagramas de fase de diferentes tipos, que eram seus auxiliares favoritos para o processo de como ele imaginava ao fazer pesquisas, em vez dos modelos mecânicos, como os que Maxwell, ele usou para construir sua teoria eletromagnética, que podem não representar completamente seus correspondentes fenômenos.

Já em 1875 a 1878 Gibbs escreveu uma monografia intitulada “Sobre o Equilíbrio de Substâncias Heterôgenas”. Onde ele não estendeu sua análise termodinâmica para sistemas químicos multifásicos. (Ou seja, sistemas compostos para mais de uma forma da matéria). Esse trabalho tinha setecentas equações matemáticas e citações de Rudolf Clausius, que mais tarde seria a primeira e segunda leis da termodinâmica. “A energia do mundo é constante. A entropia do mundo tende ao máximo”. Seu trabalho também foi descrito como o princípio da termodinâmica e como tinha um escopo ilimitado serviu como base para a física química. Gibbs lançou o termo “mecânica estatística”. É importante salientar que a Entropia; é o grau de aleatoriedade de um sistema, então quando dizemos que um determinado sistema tem uma maior variação de entropia logo sabemos que esse sistema tem uma maior desordem, logo menos energia para ser utilizada. A entropia está ligada diretamente a alguns processos espontâneos de física e química logo, que isso acontece o sistema ficara menos organizado o que fica impossível dizer que a entropia diminuiria em um determinado sistema e voltasse acontecesse um processo não espontâneo. Gibbs, introduziu conceitos-chave na descrição matemática correspondente de sistema físicos e noção de potencial químico (1876), e conjuntos estatísticos (1902). “A derivação de Gibbs das leis da termodinâmica” a partir das propriedades estatísticas de sistemas que consistem em muitas partículas foi apresentada em seu influente livro *Elementary Principles in Statistical Mechanics*, publicado em 1902, um ano antes de sua morte. Gibbs foi um membro assíduo da igreja congregacional foi um leitor fervoroso do partido republicano, mas em 1884 em virtude de muita corrupção associada a política mecâni-

ca nos EUA ele optou por apoiar um candidato democrata o Grover Cleveland, Gibbs morreu aos 64 anos de uma infecção intestinal aguda, ele com certeza é um desse caras maravilhosos que muito contribuiu para a termodinâmica e a ciência de um modo geral.^{a-e}

2. METODOLOGIA

Para a realização deste trabalho foi feita, inicialmente, uma pesquisa do referencial teórico sobre aplicação da termodinâmica na docagem molecular, os critérios de espontaneidade das reações químicas, a história da termodinâmica e suas relações com a revolução industrial. Nos livros didáticos, nos repertórios das plataformas de buscas acadêmica/periódicos científicos, o que nos permitiu fazermos um delineamento da pesquisa por meio das informações coletadas com relação ao assunto, isto é, concepções de diferentes autores que discorreram sobre o tema em questão, e as suas formas de interagir os pontos importantes deste trabalho, termodinâmica e sua aplicação em docagem molecular e critérios de espontaneidade de uma reação Química. Neste sentido, quanto ao procedimento técnico adotado à pesquisa caracterizou-se como bibliográfico.

Segundo Gil (2002, p.44), a pesquisa bibliográfica é desenvolvida com base em material já elaborado, constituído principalmente de livros e artigos científicos. Embora em quase todos os estudos seja exigido algum tipo de trabalho dessa natureza, há pesquisas desenvolvidas exclusivamente a partir de fontes bibliográficas. Boa parte dos estudos exploratórios pode ser definida como pesquisas bibliográficas.¹⁵

Logo, a escolha de tal procedimento técnico justifica-se, por ser mais adequado com o objetivo do presente estudo, pois a pesquisa bibliográfica de acordo com Marconi & Lakatos (2003),²⁰ é desenvolvida com o objetivo de colocar o pesquisador em contato direto com todo o acervo já escrito, sobre um determinado assunto. O principal objetivo desta discussão é de descrever a importância da termodinâmica na docagem molecular, visando obter informações pertinentes à aplicação da termodinâmica nesse método computacional de cálculo, visto que a docagem molecular é um método muito promissor atualmente, no que tange as interações entre as moléculas. Assim, trata-se de um trabalho com o objetivo exploratório, para Gil (2002, p. 41)¹⁵, considera que, “Estas pesquisas têm como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses”.

Neste viés, por natureza investigativo, este é um trabalho que se baseou na abordagem qualitativa e descritiva, na qual segundo Prodanov e Freitas (2013), corroboram que, na abordagem qualitativa, a “[...] interpretação dos fenômenos e a atribuição de significados são básicas no processo de pesquisa qualitativa. Esta não requer o uso de métodos e técnicas estatísticas [...]. Tal pesquisa é descritiva. Os pesquisadores tendem a analisar seus dados indutivamente”²⁷

3. FUNDAMENTAÇÃO TEÓRICA

A Termodinâmica, do grego “*therme*”, que significa calor e “*dynamis*”, que significa força, é uma área da Física, que estuda as trocas de calor entre um sistema e sua vizinhança, analisando o seu comportamento nas variações de volume, pressão e temperatura.

Em termos moleculares, o calor é a transferência de energia que faz uso do movimento caótico (aleatório) das moléculas. O movimento desordenado das moléculas é denominado movimento térmico. O movimento térmico das moléculas nas vizinhanças quentes de um sistema frio estimula a movimentação mais vigorosa das moléculas do sistema e, em virtude disso, a energia do sistema aumenta. Quando o sistema aquece as suas vizinhanças, são as moléculas do sistema que estimulam o movimento térmico das moléculas nas vizinhanças³.

O pensamento termodinâmico é predominantemente dedutivo e desenvolve-se a partir de certas premissas e conceitos fundamentais de natureza empírica, pois resultam de medidas e operações experimentais²⁵. A termodinâmica é uma ciência experimental, pois a partir da observação de alguns fenômenos físico-químicos foram elaboradas leis básicas, conhecidas como a Lei “Zero”, a Primeira, a Segunda e a Terceira Leis da termodinâmica. Os problemas que a termodinâmica se propõe a resolver normalmente envolvem a determinação do valor do calor e/ou trabalho (formas de energia) necessários ou liberados num processo ou então as mudanças de estado de uma substância ou mistura provocadas pela transferência de calor ou pela realização de trabalho¹⁰.

A Lei “Zero” da Termodinâmica afirma que, se dois corpos apresentam igualdade de temperatura com um terceiro corpo, então eles também apresentam igualdade de temperatura entre si¹⁰. A primeira lei preocupa-se em acompanhar as variações de energia e permite o cálculo da quantidade de calor que uma reação produz. A segunda lei, explica porque algumas reações ocorrem e outras não³. A Terceira Lei permite definir as entropias absolutas das substâncias e a entropia-padrão de uma reação⁴.

A capacidade de prever se uma reação pode ou não ocorrer, é dada pelos critérios de equilíbrio e espontaneidade termodinâmica, através do estudo de suas grandezas, que analisam a espontaneidade das reações químicas. Essas grandezas são denominadas como: Entalpia (ΔH), Entropia (ΔS), Energia Livre de Helmholtz (ΔA) e Energia Livre de Gibbs (ΔG), que serão descritas logo abaixo.

A função termodinâmica chamada entalpia, (da palavra grega *entalpein*, que significa “aquecer”) responde pelo fluxo de calor nas mudanças químicas que ocorrem à pressão constante quando nenhuma forma de trabalho é realizada a não ser PV. A entalpia que representada pelo símbolo H é igual à energia interna mais o produto da pressão pelo volume do sistema (BROWN et al, 2005):

$$(1) \\ H = U + PV$$

Onde H é a entalpia, U é a energia interna do sistema, PV produto pressão e temperatura, ou seja, é o trabalho realizado a pressão constante. A entalpia (H) é uma função do estado, ou seja, não depende do percurso (caminho) realizado pelo sistema durante a mudança de estado, dependendo somente do estado inicial e final ($\Delta H = H_f - H_i$).

A entropia, S, é uma grandeza termodinâmica de importância fundamental, uma vez que a ela está atrelado o sentido das transformações, ou, em outras palavras, a ordem natural com que os eventos ocorrem⁶. A variação da entropia do universo (ΔS_{univ}) é determinada pela soma da variação da entropia do sistema (ΔS_{sis}) e da variação da entropia de sua vizinhança (ΔS_{viz}):

$$\Delta S_{\text{univ.}} = \Delta S_{\text{sis.}} + \Delta S_{\text{viz.}} \quad (2)$$

Nas investigações em termodinâmica, o universo se divide em duas partes: o sistema e as vizinhanças do sistema. O sistema é a parte do universo em que estamos interessados. As vizinhanças são a parte externa do sistema onde fazemos as medidas. A entropia de um sistema, a pressão constante deve aumentar se a entropia do sistema se mantém constante, a entalpia deve diminuir, pois em qualquer processo espontâneo é essencial haver aumento da entropia das vizinhanças⁴.

Muitas reações biológicas são acompanhadas pela diminuição de entropia do sistema e, portanto, devem ser forçadas por uma fonte externa de energia. Essa energia vem da luz do sol e dos alimentos que armazenaram energia solar. A energia para forçar uma reação biológica não-espontânea é fornecida por um processo bioquímico diferente, que gera muita entropia quando ocorre na sua direção espontânea. Esse processo empurra outras reações na direção não-espontânea; por exemplo, para formar uma molécula de proteína ou para contribuir para a construção de uma molécula de DNA. Em outras palavras, os processos bioquímicos estão acoplados: uma reação pode ser levada para energias livres crescentes por outra reação que leva a energias livres decrescentes³

Um processo espontâneo a temperatura e volume constantes é acompanhado por uma diminuição na energia de Helmholtz. A variação na energia de Helmholtz é igual ao trabalho máximo que acompanha um processo a temperatura constante⁴. Ela é definida como:

$$A = U - TS \quad (3)$$

Onde A é a energia livre de Helmholtz, U é a energia interna do sistema, T é a temperatura absoluta e S a entropia do sistema.

Um processo espontâneo, a temperatura e pressão constantes é acompanhado por uma diminuição na energia de Gibbs. A variação na energia de Gibbs é igual ao trabalho máximo diferente do que de expansão que acompanha um processo a temperatura e pressão constantes⁴. Acreditava-se que a espontaneidade de uma reação química só seria determinada pela variação da entalpia (ΔH) e pela variação de entropia (ΔS). Gibbs propôs uma nova função de estado, agora chamada de energia livre de Gibbs (ou apenas energia livre). A energia livre de Gibbs, G , de um estado é definida como (BROWN et al, 2005):

$$(4) \\ G = H - TS$$

Onde T é a temperatura absoluta. Para um processo que ocorre a temperatura constante, a variação na energia livre do sistema, ΔG , é dada pela expressão:

$$(5) \\ \Delta G = \Delta H - T\Delta S$$

Critérios de Equilíbrio e Espontaneidade Termodinâmica

Todos os processos reais ocorrem naturalmente em um sentido, entretanto o processo inverso é considerado não natural. Por exemplo, uma fruta cai de uma árvore, contudo não vemos um fruto caído ao chão voltar espontaneamente a uma árvore. Deste modo os processos espontâneos são também considerados processos irreversíveis. Assim, o estudo deste

conteúdo, especificamente dos fatores que determinam o sentido espontâneo de um processo, se constitui como tema importante do ensino de Química (OLIVEIRA et al, 2019).

Entropia (ΔS)

Para um sistema em equilíbrio térmico com suas vizinhanças na temperatura T:

$$(6) \quad \Delta S \geq \frac{dq}{T}$$

Essa relação é conhecida como “*Desigualdade de Clausius*”.

Quando a energia sai do reservatório quente como calor, a entropia do reservatório diminui. Quando a mesma quantidade entra no reservatório frio, a entropia aumenta mais do que a diminuição que ocorreu no reservatório quente. Há então um aumento global de entropia e o processo é espontâneo.

Energia Livre de Helmholtz (ΔA)

Imaginemos um sistema em equilíbrio térmico com as suas vizinhanças, na temperatura T. Quando ocorre uma mudança no estado do sistema e há troca de calor entre o sistema e as suas vizinhanças, a desigualdade de Clausius diz que:

$$(6) \quad \Delta S \geq \frac{dq}{T} \Rightarrow T\Delta S \geq dq \quad (7)$$

Da 1ª Lei da Termodinâmica, temos que:

$$(8) \quad (9)$$

$$du = dq + dW \Rightarrow dq = du - dW$$

Substituindo a equação (9) na Segunda Lei, equação (7), temos:

$$(10) \quad (10.1)$$

$$TdS \geq dU - dW \Rightarrow TdS - dU + dW \geq 0$$

Sendo $dW = dW_{exp.} + dW_{útil}$; considerando $dW_{útil} \geq 0$,
 $dW_{exp.} = -pdV$

Substituindo $dW_{exp.}$ na equação 10, temos:

$$(10.2)$$

$$TdS \geq dU + pdV$$

Para um sistema isolado

$$dU = 0; dW = 0; dq = 0$$

Para essa condição, podemos rescrever a equação 10 como segue abaixo:

$$(10.3)$$

$$TdS \geq 0 \Rightarrow dS = 0$$

Assim, de acordo com a equação 10.3, tem - se o seguinte enunciado: “a entropia pode apenas crescer e atingir um valor máximo no equilíbrio”.

Para a equação 10.2, considerando T e V constantes, temos:

$$TdS \geq dU + pdV$$

$$d(TS) \geq dU + pdV$$

$$d(TS) - dU \geq 0$$

$$(10.4)$$

$$-d(U - TS) - du \geq 0$$

Assim, a partir da equação 10.4, a Energia Livre de Helmholtz é definida segundo a equação 10.5 a seguir:

$$(10.5)$$

$$A = U - TS$$

Essa equação pode ser escrita como:

$$(10.6)$$

$$dA = dU - TdS \text{ ou } \Delta A = \Delta U - T\Delta S$$

$$(8)$$

$$\text{Sendo } du = dq + dW \text{ e } TdS = dTS + dST$$

$$\text{Então: } dA = dq + dW - dTS - TdS$$

Para processos reversíveis:

$$dA = TdS + dW - dTS - dST$$

$$dA = dW - dTS \text{ ou } dA = dW - SdT$$

Para um processo irreversível e a T constante (isotérmico):

$$dA = dW \text{ (10.7)}$$

Para um sistema isolado

$$\Delta U = 0$$

$$dA = dU - TdS$$

$$(10.8)$$

$$dA = -TdS$$

Para um processo a T constante e que nenhum trabalho é produzido:

Rescrevendo a equação (10.7), temos:

$$dA = dW; dA = 0$$

$$dA = 0$$

Quando a Energia livre de Helmholtz diminui em um processo a volume e temperatura constantes, a entropia do universo aumenta e o processo é espontâneo.

$$dA < 0 \text{ (espontâneo)}$$

$$dA = 0 \text{ (equilíbrio)}$$

$$dA > 0 \text{ (não espontâneo)}$$

Energia Livre de Gibbs (dG) (8)

Da 1ª Lei da Termodinâmica, temos que: $du = dq + dW$
(9)

$$\Rightarrow dq = du - dW$$

Substituindo a equação (9) na Segunda Lei, equação (7), temos:

(10)

$$TdS \geq dU - dW$$

Sendo $dW = dW_{exp.} + dW_{\acute{u}til}$, temos:

$$TdS \geq dU - dW_{exp.} - dW_{\acute{u}til}$$

Sendo $dw_{exp.} = -pdV$, teremos:

$$TdS \geq dU - (-pdV) - dW_{\acute{u}til}$$

$$TdS \geq dU + pdV - dW_{\acute{u}til}$$

$$TdS - dU - pdV \geq -dW_{\acute{u}til}$$

(10.9)

$$-d(U + pV - TS) \geq -dW_{\acute{u}til}$$

Pela equação 10.9, definiremos então a Energia Livre de Gibbs (G), como sendo:

(11)

$$G = U + pV - TS$$

Sabendo que:

(1)

$$H = U + pV$$

(3)

$$A = U - TS$$

Assim, a equação 11 também pode ser escrita como:

$$G = H - TS \text{ ou } G = A + pV$$

Substituindo a eq. 11 em 10.9, temos que:

$$-dG \geq -dW_{\text{útil}}$$

Para uma transformação que nenhum trabalho é produzido:

$$-dG \geq 0$$

$$dG \leq 0$$

Isso significa que para sabermos se uma reação é ou não espontânea, basta verificarmos a variação da energia livre de Gibbs (ΔG ou dG), se diminui, o processo é espontâneo e reversível, se aumenta, o processo é não espontâneo e irreversível.

$$dG < 0 \text{ (espontâneo)}$$

$$dG = 0 \text{ (equilíbrio)}$$

$$dG > 0 \text{ (não espontâneo)}$$

A energia livre de Gibbs é mais comum que química do que a do Helmholtz. O critério $dG \leq 0$ no âmbito da química diz que, a uma temperatura e pressão constantes, as reações químicas são espontâneas no sentido da diminuição da energia de Gibbs⁴. Ela é comumente empregada em química porque, na natureza, as reações ocorrem mais a pressões constantes do que a volume constantes. Esse critério $dG \leq 0$ nos permite prever se uma reação química é termodinamicamente possível, ou seja, se teremos a conversão total dos reagentes em produtos.

4. APLICAÇÃO DA TERMODINÂMICA NO DOCKING MOLECULAR

4.1. INTRODUÇÃO

Segundo De Oliveira (2018)²³, o docking molecular, ancoramento molecular ou ainda modelagem molecular é uma área da química computacional que se destina a simular sistemas físicos e químicos, tais como moléculas, átomos, sólidos, líquidos, gases, soluções e desenvolvimento de novos fármacos fornecendo informações importantes para o processo de planejamento. De acordo com Barreiro (1996)⁵, o estudo de docking é uma das técnicas da modelagem molecular mais importantes no desenvolvimento de fármacos. Trata-se de um estudo de interação energética entre o ligante e o receptor utilizando campo de força, onde a interação do fármaco ao receptor é um problema complexo, visto que inúmeras forças são envolvidas, tais como: eletrostáticas, eletrodinâmicas e estéricas, assim como a interação com o solvente e os graus de flexibilidade das moléculas são também fatores importantes a serem considerados.

O desenvolvimento destes métodos computacionais que permitem simular sistemas químicos a nível atômico-molecular tem acompanhado as principais tendências no desenvolvimento de novos materiais, aplicados às mais diversas demandas tecnológicas e a evolução dos programas computacionais levará a obtenção de resultados cada vez mais seguros e futuros experimentos poderão ser alcançados.

O docking molecular é amplamente usado nos estudos de síntese de fármacos, por ser um recurso eficiente para analisar e determinar os aspectos que regem as interações dos fárma-

cos (ligantes), com seus alvos (proteínas), antes mesmo que sejam sintetizados.

A termodinâmica surgiu com o advento da revolução industrial, onde o interesse principal era a possibilidade de transformação de calor em trabalho, também pode fornecer informações úteis sobre propriedades materiais com base em algumas medidas experimentais sem especificação de detalhes moleculares. A sua aplicação no docking molecular, tem emergido significativamente nos últimos anos, sendo uma nova área de pesquisa emergente nas fronteiras das várias disciplinas.

Sá (2020)³² destaca que a aplicação da termodinâmica no docking, pode melhorar o entendimento em relação aos fenômenos químicos, nos seus níveis macroscópico, microscópico (partículas, átomos, íons e moléculas), assim como simbólico, este representado por equações, coeficientes, gráficos, imagens e números. Prausnitz (2002)²⁶ diz que uma aplicação particular da termodinâmica molecular diz respeito à separação de proteínas aquosas por precipitação seletiva. A termodinâmica molecular combina termodinâmica clássica e estatística juntamente com modelos de física e química para alcançar fins práticos.

Os programas de química computacional se valem da termodinâmica estatística para a computação de propriedades termoquímicas de sistemas moleculares. O ponto de partida da termodinâmica estatística é a construção da função de partição, que é uma grandeza que descreve as propriedades estatísticas de um sistema em equilíbrio termodinâmico.²²

A termodinâmica estatística tem uma aplicação direta no docking molecular, que para ser caracterizada, deve-se primei-

ro tratar da construção da função de partição, que é uma grandeza que descreve as propriedades estatísticas de um sistema em equilíbrio termodinâmico, sendo assim:

$$(22) \\ Z = \sum_j e^{-\beta E_j}$$

Onde E é a energia total do sistema quando está no microestado j e é a temperatura inversa definida como:

$$(22) \\ \beta = \frac{1}{K_B T}$$

Onde K_B , é a constante de Boltzmann

Segundo o autor, a função de partição, surge da soma dessas contribuições, podendo-se derivar algumas funções básicas da termodinâmica, tais quais a energia, capacidade calorífica, entropia, e energia de Helmholtz²²:

$$\langle E \rangle = - \frac{\partial \ln Z}{\partial \beta} \\ C_V = \frac{\partial \langle E \rangle}{\partial T} = \frac{1}{k_B T^2} \langle \partial E^2 \rangle \\ S = K_B (\ln Z + \beta \langle E \rangle) = \frac{\partial}{\partial T} (K_B \ln Z) \\ A = \langle E \rangle - TS = - K_B T \ln Z$$

Ao avaliarmos as definições de energia livre de Helmholtz, $A \equiv U - TS$, e a energia livre de Gibbs, $G \equiv U + pV - TS$, observamos que ambas são dependentes de duas grandezas comuns, a energia interna (ou energia potencial) do sistema e a entropia. Ou seja, as mudanças entrópicas e as interações interatômicas influenciam nas afinidades e estabilidades de ligação/interação e nas torções/dobramentos de proteínas.

3.2. DETERMINAÇÃO DA ENERGIA LIVRE DE LIGAÇÃO NO SISTEMA LIGANTE-PROTEÍNA

O processo de docking pode ser feito interagindo a proteína com um ligante por vez ou com múltiplos ligantes, o também chamado método *Ligand Simultaneous Docking* (MLSD). Normalmente para compreender a interação e descobrir as principais forças responsáveis pela ligação entre ligantes e proteínas, são aplicados os métodos baseados no docking de um único ligante molecular (fazendo-se a inserção de um ligante por cada vez). Segundo Sharma *et al* (2020)³³, esses métodos se mostram úteis no docking de um ligante em diferentes sistemas moleculares, porém são inúteis para o modo de docking e especialmente a previsão de afinidade quando mais de um ligante está envolvido. Ainda segundo os autores, o docking baseado em um único ligante molecular não representa a realidade física, ou seja, a dinâmica dos processos na vida real, pois geralmente as substâncias interagem com várias outras ao mesmo tempo, sendo raras as interações um a um na natureza. Por isso é mais adequado usar o docking com ligantes simultâneos (MLSD), por permitir uma melhor simulação dos processos reais de reconhecimento molecular. Este tipo de trabalho é de vital importância nas áreas de ciências de superfície, sistemas biológicos e aplicações industriais³³.

O docking molecular fornece diferentes conformações espaciais do ligante, dentre os quais existe um com maior probabilidade de se ligar ao sítio ativo da proteína-alvo. Os programas nos dão a possibilidade de visualizar as tão importantes energias livres de ligação (ligante-alvo) para cada conformação espacial gerada do ligante. Na interpretação destes dados de energia, é importante ter atenção que a menor energia é considerada a mais provável para justificar a conformação da interação³⁴.

Na busca de possíveis novos ligantes para determinados alvos (ex.: busca de novos fármacos) por dinâmicas de docking, o investigador precisará fazer o preparo das estruturas químicas destes ligantes a serem testados, através de varredura conformacional e alinhamento tridimensional com o ligante nativo. O objetivo é atribuir a esses novos ligantes, disposições atômicas parecidas com as dos átomos do ligante nativo³⁴.

Como dito no tópico anterior, os mecanismos usados no docking molecular possuem uma relação com a termodinâmica, principalmente a termodinâmica estatística. Relatamos anteriormente que é possível determinar a espontaneidade e o equilíbrio de sistemas através do estudo de grandezas termodinâmicas, tais como Entalpia (ΔH), Entropia (ΔS), Energia Livre de Helmholtz (ΔA) e Energia Livre de Gibbs (ΔG). Todavia, dentre as citadas, a energia livre de Gibbs (ΔG) é a mais adequada para ser usada como critério de espontaneidade e equilíbrio de sistemas e reações⁴.

Sabemos que as interações entre ligantes e alvos proteicos se dão através de forças intermoleculares do tipo dipolo-dipolo induzido (forças de Van Der Waals ou London) e dipolo-dipolo permanente (ligações de hidrogênio). Neste sentido, um dos objetivos do docking é o cálculo da energia livre de ligação não covalente entre ligantes e a proteínas-alvo³⁴.

Assim, para interações do tipo “*Proteína + Ligante* \leftrightarrow *Proteína-Ligante*”, a energia livre de ligação pode ser expressa como:

$$\Delta G_{\text{Ligação}} = G_{\text{Proteína-Ligante}} - G_{\text{Proteína}} - G_{\text{Ligante}} \quad 34$$

Estudos anteriores, trazem uma abordagem próxima à que estabelecemos como objeto de estudo do presente trabalho, descrever o contributo da termodinâmica no processo de

docking molecular. Os autores buscaram os valores de energia livre envolvidas nas interações/ligações de diferentes proteínas e ligantes (disponíveis no *Protein Data Bank*) por duas vias: estimando esses valores por cálculos diretos e obtendo-os através de docking molecular.^{1,14,17}

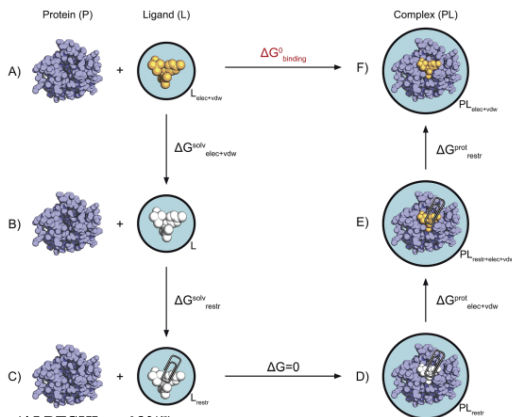
A espontaneidade de um processo é determinada por sua energia livre, porque, sem uma fonte externa de energia, os sistemas evoluem para seu estado de energia livre mais baixo. Por sua vez, as energias livres também são extremamente importantes na determinação de afinidades de ligação (ligação espontânea ou não) ou dobramento de proteínas (o estado dobrado é geralmente o mínimo de energia livre)¹⁴.

As simulações de dinâmica molecular (como é o caso do docking), permitem o acesso a diferenças de energia livre que governam a força motriz subjacente a todos os processos biológicos. Um aspecto importante da validação de uma configuração de simulação é ilustrado por meio do cálculo das diferenças de energia livre ao longo de um ciclo termodinâmico completo.¹⁴

A abordagem sobre cálculo de energia livre está associada à ciclos termodinâmicos. De uma forma simples, podemos definir ciclos termodinâmicos como sendo uma série de processos termodinâmicos (repetitivos) que um sistema realiza a fim de produzir trabalho, ou de se realizar trabalho sobre ele. Os ciclos termodinâmicos são a base de operação de máquinas térmicas/motores de combustão, havendo quatro tipos que se destacam: Ciclo de Otto, Ciclo de Diesel, Ciclo de Stirling e Ciclo de Carnot. Gapsys *et al* (2015) e Aldeghi *et al* (2015) em seus estudos, determinaram as energias livres com base em um ciclo termodinâmico (figura 1), para vários ligantes se unindo à uma única proteína.^{1,14}

No caso de Aldeghi *et al* (2015), estudaram um conjunto de diversos inibidores que se ligam à proteína apo 4 contendo bromodomínio (BRD4), e obtiveram um erro absoluto médio de 0,6 kcal mol⁻¹ e uma precisão de 1,0 kcal mol⁻¹. Por sua vez, Gapsys *et al* (2015) realizaram um estudo ilustrativo, de modos a estimar a afinidade de ligação de um conjunto de inibidores com a proteína trypsina, usando medições de calorimetria isotérmica (ITC).¹

Figura 1. Esquema do ciclo termodinâmico não físico usado para obter as energias livres de ligação. O ligante totalmente interativo (laranja) em solução no canto superior esquerdo (A) é transformado em um soluto não interativo (B, branco) durante uma série de simulações de equilíbrio onde suas interações eletrostáticas e de Van der Waals são escaladas para zero, fornecendo o termo $\Delta G^{\text{sol}v}_{\text{elec+vdw}}$. O ligante é então contido sem interagir com o meio ambiente (C). O ligante que não interage com o meio ambiente fica contido dentro da cavidade da proteína (D). O ligante contido e não interativo em complexo com a proteína tem suas interações eletrostáticas e de Van der Waals religadas (E), fornecendo o termo $\Delta G^{\text{prot}}_{\text{elec+vdw}}$. Então, as restrições de interação entre o ligante e a proteína são removidas ($\Delta G^{\text{sol}v}_{\text{restr}}$), fechando o ciclo termodinâmico, cujo estado final é a total e desenfreada interação entre o ligante em complexo com a proteína (F).



Fonte: (ALDEGHI *et al*,2015)

Geralmente a convergência entre os ligantes e as proteínas são lentas devido grandes perturbações que as energias livres envolvem. Por isso, se deve buscar a melhor configuração possível a partir da qual se tem maiores chances de estimar uma energia livre precisa. No caso de ser preciso determinar a mudança na afinidade de ligação após uma mutação de proteína ou modificação de ligante, se deve determinar a diferença entre a energia livre de ligação do ligante ou proteína no estado inicial e no estado final do ciclo fechado. A diferença de energia livre ao longo de um ciclo fechado de reações sempre somará zero¹⁴.

Souza (2015) nos apresenta três classes de tipos de funções para avaliar a energia livre de ligação de sistemas receptor-ligante utilizadas pelos programas computacionais de docking molecular:³⁴

- 1) Classe de funções baseadas em campos de força: o campo de força cria uma representação física do sistema estudado através de simulações computacionais, e suas equações procuram disfarçar o potencial de interação entre os átomos.

- 2) Classe de funções empírica: as equações utilizam coeficientes pré-otimizados vindos de dados experimentais de interações proteína-ligante.

- 3) Classe de funções baseadas em conhecimento: utilizam informações vindas de estruturas determinadas experimentalmente, para se basear e criar possíveis geometrias assumidas pelos ligantes no sítio ativo.

3.3. INTERPRETAÇÃO DE DADOS DE UM DOCKING MOLECULAR

Depois de estudos de otimização desta área da química, nos dias de hoje se conhecem diversas técnicas e ferramentas computacionais para a realização do docking. Dependendo das moléculas que se está usando ou dos objetivos que se quer alcançar, o investigador pode escolher a metodologia de docking através de quatro maneiras: 1) ligante rígido e proteína rígida, 2) ligante rígido e proteína flexível, 3) ligante flexível e proteína rígida e 4) ligante flexível e proteína flexível³⁴. E como vimos anteriormente, dependendo da quantidade de ligantes, o processo de docking pode ser feito interagindo a proteína com um ligante por vez ou com múltiplos ligantes (Ligand Simultaneous Docking, MLSD)³³.

Foram consultadas várias referências^{33; 19; 14 11; 22; 17; 34} para compreender o processo de interpretação dos dados gerados em um docking molecular. Cada investigador estabelece um objetivo particular para sua pesquisa, e por isso, tanto os métodos quanto a interpretação de dados de um docking varia. Como ilustração para observação da análise dos dados gerados no docking, principalmente a análise dos valores de energia livre de ligação, apresentamos o estudo de Sharma *et al* (2020)³³.

Sharma *et al* (2020) realizaram um estudo físico-químico e conformacional do efeito de dois co-solventes, Dimetil Sulfoxido (DMSO) e Glicerol, nas interações moleculares de um sistema proteína-surfactante com carga semelhante, constituído pela proteína Albumina de Soro Bovino (BSA) e o surfactante aniônico Dodecil Benzeno Sulfonato de Sódio (SDBS). A presença dos co-solventes atribuiu ao sistema BSA-SDBS, propriedades de um novo sistema de distribuição de drogas.

Os estudos das interações foram realizados mediante várias abordagens teóricas e experimentais. Se fez o estudo da condutividade, fluorescência, se realizou um docking com ligante único para o sistema BSA-SDBS e outro docking aplicado a abordagem Ligand Simultaneous Docking (MLSD), acoplado a proteína BSA com os ligantes SDBS + Glicerol e SDBS + DMSO simultaneamente.

Os resultados do docking mostraram que a proteína BSA possui 8 cavidades, cujas coordenadas foram agrupadas hierarquicamente por um algoritmo simples, de acordo com sua similaridade espacial (com base na distância). Também foram identificados três domínios homólogos (I, II e III), que foram avaliados, e cada local foi encaixado com substrato único e múltiplo simultâneo na estrutura 3D do BSA usando Auto-Dock³⁵. O programa gerou os valores de energia livre (ΔG), dos quais foram apresentadas as energias mais baixas (Tabela 1):

Tabela 1. ΔG de ligação ($\text{kJ}\cdot\text{mol}^{-1}$) para o docking BSA-Substratos*.

SUBSTRATOS	CAVIDADES							
	Cv 1	Cv 2	Cv 3	Cv 4	Cv 5	Cv 6	Cv 7	Cv 8
SDBS ^b	-29,71	-19,66	-26,78	-20,92	-10,88	-28,45	-25,94	-28,03
SDBS + Glicerol ^c	-12,55	-15,89	-13,38	-15,06	-13,38	-12,13	-9,20	-12,97
SDBS + DMSO ^c	-9,20	-10,87	-10,04	-10,46	-7,53	-7,94	-7,11	-10,46

*Valores calculados com uma incerteza padrão de (ΔG) = 2,70 $\text{kJ}\cdot\text{mol}^{-1}$, sob as seguintes condições: temperatura (T) = 0,05 K, pressão (p) = 10 kPa e medições termodinâmicas com nível de confiança de 0,68.

^aA energia foi calculada em cada lugar para verificar qual apresentava o maior grau de união com o ligante.

^bDocking de ligante único.

^cDocking simultâneo de múltiplos substratos.

Fonte: (SHARMA *et al*, 2020)³⁵

Analisando os dados de energia livre (ΔG) apresentados na tabela, os autores puderam interpretar que a ordem da energia encontrada para cada um dos locais de ligação foi a seguinte: Cv2 < Cv4 < Cv3 < Cv5 < Cv8 < Cv1 < Cv6 < Cv7. A partir desta constatação, concluíram que os Cv2 representam os melhores sítios/cavidade de ligação para surfactantes e co-solventes. Após análise, percebeu-se que os sítios de ligação Cv2 da proteína BSA consistem em 2 resíduos hidrofóbicos, 4 polares, 2 ácidos e 4 básicos, demonstrando dessa maneira que nelas acontecem interações hidrofóbicas e que desfavorecem a formação de micelas devido à presença de co-solventes³³.

A partir dos dados apresentados na tabela, também se observa uma diminuição na afinidade de ligação depois da adição de Glicerol ao SDBS, que cresce ainda mais quando DMSO foi adicionado ao SDBS. Isto deve-se ao fato de que em interação com surfactantes iônicos, como é o caso do SDBS, o Glicerol apresenta propriedades cosmotrópicas e o DMSO apresenta propriedades caotrópicas. Ressalta-se também o fato de que o local de ligação é modificado pela presença de dois ligantes simultâneos e depende da natureza do co-solvente e do surfactante³³.

3.4. CONSIDERAÇÕES FINAIS

O presente capítulo procurou esclarecer o objetivo traçado que era o de fazer uma análise sobre a aplicação da termodinâmica no docking molecular. Na literatura, existem obras que abordam sobre o assunto, mas em um número limitado. Este estudo é de extrema importância, pois contribui para a busca de novas descobertas científicas que tem surgido significativamente nos últimos anos, sendo uma nova área de pesquisa emergente.

REFERÊNCIAS

1. ALDEGUI, Matteo *et al.* Accurate calculation of the absolute free energy of binding for drug molecules. **The Royal Society of Chemistry**, v. 7, p. 207-218, 25 sept. 2015. DOI: 10.1039/C5SC02678D.
2. Silva, A. G., & Andrade e Silva, J. (1982). *Origens históricas da física moderna: introdução abreviada* (1st ed.). Fundação Calouste Gulbenkian.
3. ATKINS, Peter; JONES, Loretta. **Princípios de Química: questionando a vida moderna e o meio ambiente – 3ª ed.** Porto Alegre: Bookman, 2006.
4. ATKINS, Peter; PAULA, Júlio de. **Físico-química 1 – 9ª ed.** Rio de Janeiro: LTC, 2012.
5. BARREIRO, E. J.; RODRIGUES, C. R. Modelagem Molecular: Uma ferramenta para o planejamento racional de fármacos em química medicinal I. **Química Nova**, v. 20, p. 1-17, 1996.
6. CAVALCANTI, Higo L. B.; FERREIRA, Edvan A.; ABRANTES, Paloma G.; CAVALCANTI, Gláucia N. **As muitas interpretações da entropia e a criação de um material didático para o ensino da interpretação probabilística da entropia.** São Paulo: Química Nova na Escola, 2018.
7. Colin A. Ronan, *História Ilustrada da Ciência.* Univ. de Cambridge (Zahar, 1987).
8. D. S. L. Cardwell, *La Recherche* 5 (48), 726 - 733 (Sept. 1974).
9. DE OLIVEIRA, André Maurício. **Introdução à modelagem molecular para química, engenharia e biomédicas.** 1. ed. Curitiba, PR: Appris Editora, 2018. 235 p. ISBN 978-85-473-1294-7.
10. FERNANDES, Fabiano A. N.; PIZZO, Sandro M.; JR., Deovaldo Moraes. **Termodinâmica Química – 1ª ed.** UFC, 2006.
11. FONSECA, Aluísio M. *et al.* Synthesis, Biological Activity, and *In silico* Study of Bioesters Derived from Bixin by the CALB Enzyme. **Biointerface research**, [S.l.], v. 12, n. 5, p. 5901-5917, 24 out. 2021. DOI: 10.33263/BRIAC125.59015917.

12. Freeman J. Dyson, "What is Heat?"; *Scientific American* 191 (3) 58 – 63 (1954).
13. G.N. Hatsopoulos, J. H. Keenan, *Principles of General Thermodynamics* (Wiley 1965).
14. Gapsys, Vytautas *et al.* Calculation of Binding Free Energies. *In: Kukol A. (eds). Molecular Modeling of Proteins: Methods in Molecular Biology (Methods and Protocols.* vol 1215. New York, NY: Humana Press, 2015. p. 173-209. DOI: 10.1007/978-1-4939-1465-4_9
15. GIL, Antônio Carlos, *Como elaborar projetos de pesquisa.* 4ª ed. São Paulo: Atlas, 2002.
16. Gulbenkian, Lisboa, 1982.
17. HAUSER, Annet *et al.* Combining free energy calculations with tailored enzyme activity assays to elucidate substrate binding of a phospho-lysine phosphatase. **The Royal Society of Chemistry**, v. 11, n. 47, p. 12655–12661, 8 sept. 2020. DOI: 10.1039/D0SC03930F.
18. Jose Maria F. Bassalo, *Crônicas da Física*, Tomo 3 (Editora UFPA, Belém, 1992).
19. KHATUN, Samima *et al.* Thermodynamics, molecular modelling and denaturation studies on exploring the binding mechanism of tetramethylpyrazine with human serum albumin. **The Journal of Chemical Thermodynamics**, v. 140, n. 105915, jan. 2020. DOI: 10.1016/j.jct.2019.105915.
20. MARCONI, Marina de Andrade; LAKATOS, Eva Maria. *Fundamentos de metodologia científica.* 5 ed. São Paulo: Atlas S.A. 2003.
21. Nelson Studart: *História da Física.* Apresentação em Power point (UFSCar).
22. OLIVEIRA, Daniel A. B. **Estudo teórico de inibidores da proteína quinase de adesão focal.** 2013. 116 f. Tese (Doutorado em Química) - Instituto de Química, Universidade de Brasília, Brasília-DF, 2013. Disponível em: <https://repositorio.unb.br/handle/10482/15190>. Acesso em 18 dez. 2021.

23. OLIVEIRA, Bruno P. de; FECHINE, Pierre B. A.; ROMERO, F. Belmino; SOARES, A. Cury. **Concepções alternativas de conceitos termodinâmicos: uma investigação sobre a espontaneidade dos fenômenos naturais com os alunos dos cursos licenciatura em química.** Revista Interface, Edição nº 18, p. 35 – 46, dezembro de 2019.
24. Oriol Planas - Engenheiro Técnico Industrial, especialidade em mecânica, Data de publicação: 7 de março de 2018. Última revisão: 10 de outubro de 2019. <https://pt.solar-energia.net/termodinamica/historia-da-termodinamica>. Acesso: 03/01/2022.
25. PILLA, Luiz. **Físico-química I: termodinâmica química e equilíbrio químico – 2ª ed.** Porto Alegre: UFRGS, 2010.
26. PRAUSNITZ, John M. Molecular thermodynamics for some applications in biotechnology. **The Journal of Chemical Thermodynamics**, California, v. 1, ed. 1, p. 21-39, 11 feb. 2002. DOI: 10.1016/S0021-9614(02)00305-1.
27. PRODANOV, C. C.; FREITAS, E. C. Metodologia do trabalho científico [**recurso eletrônico**]: métodos e técnicas da pesquisa e do trabalho acadêmico. 2. ed. Novo Hamburgo: Feevale, 2013.
28. Prof. Dr. Danoso J. Pedro, Datas e Personagens da História da Termodinâmica. Univ. de São Paulo, Instituto de Física São Carlos- IFSC.
29. Raymond A. Serway, Physics (4th edition, Saunders College, 1996).
30. René Taton (direção), Historia Geral das Ciências (Difusão Europeia do Livro, 1960).
31. Robert Locqueneux, Histoire de la Physique. Collection que sais je? (PUF 1987).
32. SÁ, Ézio. Termodinâmica: Uma Proposta de Ensino a partir da Química Computacional. **Revista Virtual de Química**, Piauí, v. 1, ed. 1, p. 1-14, 26 maio 2020. DOI <https://doi.org/10.21577/1984-6835.20200062>. Disponível em: <http://rvq.s bq.org.br/>. Acesso em: 18 dez. 2021.
33. SHARMA, Vivek *et al.* A physicochemical and conformational study of co-solvent effect on the molecular interactions between

- similarly charged protein surfactant (BSA-SDBS) system. **The Journal of Chemical Thermodynamics**, v. 142, n. 106022, jan. 2020. DOI: 10.1016/j.jct.2019.106022
34. SOUZA, Raphael L. **Aplicação da técnica de ancoragem molecular na otimização do fármaco hipoglicemiante metformina**. 2015. 41 f. Trabalho de Conclusão de Curso (Graduação em Ciências Farmacêuticas - Farmácia) - Centro Universitário Luterano de Palmas, Universidade Luterana do Brasil, Palmas, TO, 2015. Disponível em: <https://ulbra-to.br/bibliotecadigital/uploads/document56e6fec51e112.pdf>. Acesso em: 20 dez. 2021.
 35. TROTT, Oleg; OLSON, Arthur 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. 455-461, 2010. DOI: 10.1002/jcc.21334
 36. Vinicius Kapudjian Carabett - RA 931101, Dados Biográficos: Carnot, Nicolas Leonard (1796-1832). **Bibliografia**; Encycloaepdia Britânica, Vol. 4, pp 940, 941. Vol. 11, pp 237 a 239 Dictionary of Scientific Biography Vol.3 e 4, pp 79 a 83. <http://www.fem.unicamp.br/~em313/paginas/person/carnot.htm>. Acesso: 03/01/2022.