

## 15º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2024

### DESENVOLVIMENTO DE FERRAMENTAS COMPUTACIONAIS PARA A INVESTIGAÇÃO DA CONDUÇÃO QUÂNTICA EM PROTEÍNAS

DANIEL CASTILHO DE OLIVEIRA NETO<sup>1</sup>, FILIPE CAMARGO DALMATTI ALVES LIMA<sup>2</sup>.

<sup>1</sup> Graduando em Licenciatura em Química, Bolsista PIBIFSP, IFSP, Campus Matão, d.castilho@aluno.ifsp.edu.br.

<sup>2</sup> Professor Orientador, IFSP, Campus Matão, fdlima@ifsp.edu.br.

Área de conhecimento (Tabela CNPq): 1.06.03.07-7 Ciências Exatas e da Terra / Química / Físico-Química / Química Teórica

**RESUMO:** A compreensão dos princípios fundamentais que regem o movimento dos elétrons nas proteínas ainda apresenta muitas questões em aberto, que estão sendo gradualmente exploradas no contexto científico. Uma técnica utilizada para o estudo da condução quântica de proteínas é a microscopia de tunelamento (STM), que visa identificar as regiões mais sensíveis à passagem de elétrons. No entanto, o estudo desses mecanismos pode ser aprimorado com o uso de ferramentas computacionais que auxiliem na interpretação dos experimentos. Estamos desenvolvendo um software para apoiar os resultados experimentais de STM, modelando potenciais regiões tridimensionais de proteínas a partir de dados do Protein Data Bank (PDB). O código está sendo escrito em Python 3 e incorpora funções de busca de resíduos de aminoácidos, utilizando uma abordagem interdisciplinar que combina geometria analítica, bioquímica e banco de dados. Além disso, os resultados são integrados para visualização imediata no software de código aberto Pymol. Os resultados indicam a formação de grids de seleção das regiões das proteínas, que foram validados por meio de planilhas e verificação por *debug*. As próximas etapas deste trabalho em andamento incluem a correlação com medidas experimentais de tunelamento eletrônico. Espera-se que essas informações inéditas contribuam significativamente para o avanço no desenvolvimento de biossensores e dispositivos nanotecnológicos, além de oferecer novas compreensões sobre os mecanismos que regulam a condução quântica em proteínas.

**PALAVRAS-CHAVE:** condução eletrônica; nanotecnologia; simulações moleculares; bioeletrônica; dinâmica quântica; biomoléculas.

### DEVELOPMENT OF COMPUTATIONAL TOOLS FOR THE INVESTIGATION OF QUANTUM CONDUCTION IN PROTEINS

**ABSTRACT:** The understanding of the fundamental principles governing electron movement in proteins still presents many open questions, which are being gradually explored within the scientific context. A technique used for studying quantum conduction in proteins is scanning tunneling microscopy (STM), which aims to identify regions most sensitive to electron passage. However, the study of these mechanisms can be enhanced with the use of computational tools that assist in experiment interpretation. We are developing software to support STM experimental results, modeling potential three-dimensional regions of proteins based on data from the Protein Data Bank (PDB). The code is being written in Python 3 and incorporates functions for amino acid residue search, using an interdisciplinary approach that combines analytical geometry, biochemistry, and database systems. Additionally, the results are integrated for immediate visualization using the open-source software

PyMOL. The results indicate the formation of selection grids for protein regions, which were validated through spreadsheets and debug verification. The next steps in this ongoing work include correlation with experimental measurements of electronic tunneling. These novel insights are expected to significantly contribute to the development of biosensors and nanotechnological devices, as well as provide a deeper understanding of the mechanisms that regulate quantum conduction in proteins.

**KEYWORDS:** electronic conduction; nanotechnology; molecular simulations; bioelectronics; quantum dynamics; biomolecules.

## INTRODUÇÃO

A condução quântica em proteínas tem sido amplamente investigada nos últimos anos, com foco em entender como os elétrons se movem através de estruturas proteicas, particularmente em proteínas redox, como a azurina (SEPUNARU et al., 2011, p. 8). Essa condução é relevante devido ao seu potencial de aplicação em dispositivos bioeletrônicos, como biosensores, explorando as propriedades eletrônicas das biomoléculas (ZOTTI et al., 2019, p. 10). Estudos iniciais sugerem que proteínas redox facilitam o transporte eletrônico a longas distâncias, o que impulsiona o desenvolvimento de sistemas eletrônicos baseados em proteínas (BOSTICK et al., 2018, p. 5). Apesar disso, a medição desse transporte eletrônico em proteínas ainda enfrenta dificuldades técnicas, especialmente na manutenção de condições experimentais controladas e na formação de junções moleculares estáveis (ZOTTI et al., 2019, p. 11). Neste contexto, este trabalho propõe uma ferramenta computacional para simular o transporte eletrônico em proteínas, utilizando dados do Protein Data Bank (PDB) para explorar possíveis canais de condução quântica.

## MATERIAL E MÉTODOS

Neste estudo, foi desenvolvido e utilizado um código computacional em Python para investigar o transporte eletrônico em proteínas. A metodologia foi dividida em duas etapas principais: a busca conformacional e a separação de grids tridimensionais. A busca conformacional foi conduzida utilizando algoritmos de otimização para explorar diferentes configurações estruturais das proteínas. Esse processo permitiu identificar as conformações mais estáveis e compreender como pequenas variações na estrutura influenciam o transporte de elétrons, fornecendo informações cruciais para a análise das rotas de condução. Na sequência, a separação de grids tridimensionais foi realizada para subdividir a estrutura proteica em pequenas células cúbicas, facilitando a análise detalhada das regiões com maior relevância eletrônica.

Os grids foram calculados de forma a garantir que cada célula capturasse informações essenciais sobre a distribuição espacial dos aminoácidos e o potencial de condução. Essa abordagem permitiu identificar com precisão áreas de maior concentração de elétrons e potencial de transporte eletrônico. Para otimizar o processo e reduzir o esforço manual, o código foi ajustado para automatizar a geração dos grids e a análise gráfica dos resultados, permitindo uma variação rápida dos parâmetros de estudo e aumentando a eficiência. As proteínas analisadas foram selecionadas devido às suas características relacionadas ao transporte de elétrons, e a hipótese é que variações conformacionais, detectadas através dos grids, revelem áreas de alta condutividade.

O método adotado visa isolar e examinar regiões críticas de condução, proporcionando uma compreensão mais aprofundada das rotas eletrônicas. Por fim, os cálculos envolveram a correlação entre as posições estruturais dos aminoácidos e as possíveis rotas de condução quântica, levando em consideração a proximidade entre os átomos e suas propriedades espaciais. O cálculo das distâncias entre os átomos foi empregado para prever as rotas preferenciais de transporte de elétrons.

Modelo de equação:

$$Tamanho_{grid} = \frac{Dimensão_{máxima} - Dimensão_{mínima}}{Número\ de\ grids}$$

(1)

em que,

*Dimensão*<sub>máxima</sub> – valor máximo das coordenadas *x*, *y*, e *z*;

*Dimensão*<sub>mínima</sub> – valor mínimo das coordenadas *x*, *y*, e *z*;

*Número de grids* – número de subdivisões aplicadas em cada eixo

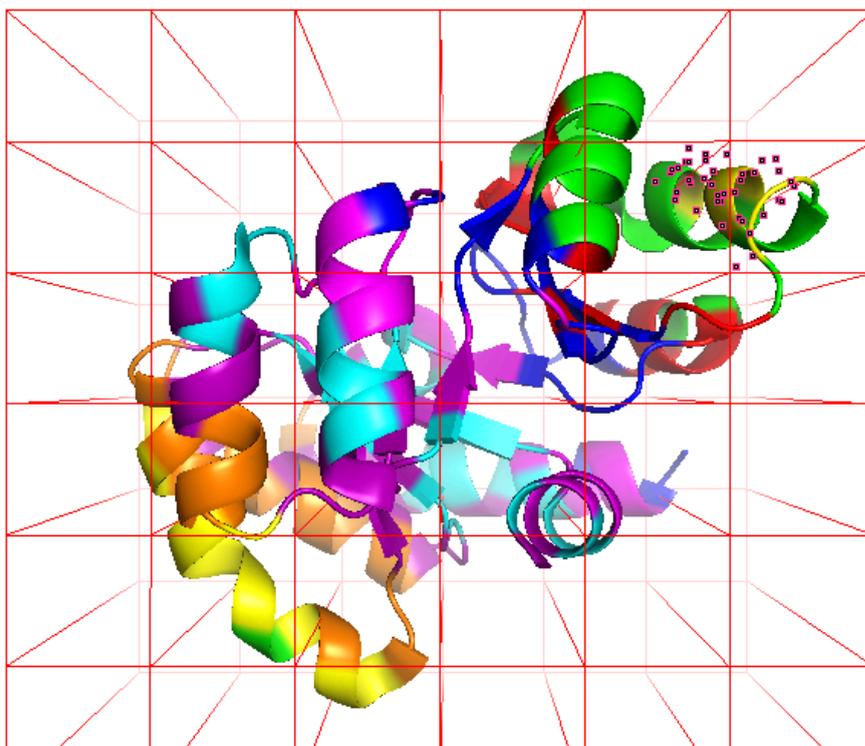
## RESULTADOS E DISCUSSÃO

Foi desenvolvido um algoritmo para criar o programa "Analisador de Estruturas PDB". Este programa foi concebido para analisar arquivos no formato PDB, com foco na extração e manipulação de dados relativos a átomos em estruturas moleculares. As coordenadas cartesianas de cada átomo são medidas a partir de um ponto de origem arbitrário, comumente representado pelas coordenadas (0,0,0). Cada coordenada indica a distância do átomo ao longo dos eixos *x*, *y* e *z*, respectivamente. Matematicamente, as coordenadas cartesianas são essenciais para a representação única da posição de cada átomo na estrutura tridimensional da molécula. A partir destas coordenadas, é possível realizar cálculos rigorosos que permitem determinar distâncias intermoleculares, ângulos de ligação e torção, bem como identificar padrões estruturais recorrentes. A determinação da distância entre dois átomos, por exemplo, pode ser realizada utilizando a fórmula da distância euclidiana em três dimensões:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

onde  $(x_1, y_1, z_1)$  e  $(x_2, y_2, z_2)$  são as coordenadas cartesianas dos dois átomos. Esta fórmula permite calcular a distância direta entre qualquer par de átomos na estrutura molecular.

O software atualmente possui as buscas de átomos por critérios específicos, o cálculo das dimensões da estrutura, a pesquisa de átomos por intervalo de coordenadas, a identificação de átomos por distância e a busca por átomos dentro de intervalos coordenados simultaneamente, denominados grids. A Figura 1 apresenta um exemplo de um grid 6x6x1 do código PDB 7B3Y (estrutura de uma nanopartícula para uma candidata a vacina contra a COVID-19 [Disponível em: <https://www.rcsb.org/structure/7B3Y>]). As esferas cinzas demarcam os limites dos grids.



**FIGURA 1.** Grid 6x6x1 selecionado na estrutura 7B3Y, Protein Data Bank (PDB).

A segmentação tridimensional aplicada na estrutura proteica facilita a análise das regiões de interesse, especialmente aquelas com maior potencial de condução eletrônica. A subdivisão em grids, além de fornecer uma representação clara da distribuição espacial dos aminoácidos, permite uma abordagem mais sistemática ao isolar áreas específicas da proteína que podem influenciar diretamente o transporte eletrônico. Esse tipo de análise é fundamental, pois a condutividade eletrônica em proteínas pode ser fortemente dependente da disposição exata dos aminoácidos e de suas interações eletrônicas com o ambiente.

Ao isolar regiões específicas da estrutura, é possível não apenas mapear as áreas mais favoráveis ao transporte de elétrons, mas também entender melhor como diferentes conformações da proteína podem impactar o comportamento eletrônico geral. Esse método possibilita que áreas de alta densidade eletrônica e proximidade entre átomos sejam analisadas em detalhe, permitindo prever rotas preferenciais de condução quântica.

Além disso, o uso de grids tridimensionais simplifica a interpretação dos dados e proporciona uma visão mais clara das interações entre as diferentes partes da proteína, facilitando a correlação entre dados experimentais e modelos computacionais. Isso contribui diretamente para o avanço no entendimento dos mecanismos de condução quântica nas proteínas e pode auxiliar no desenvolvimento de dispositivos bioeletrônicos baseados nas propriedades de condução dessas biomoléculas.

## CONCLUSÕES

Diante do que foi observado até o momento, o desenvolvimento do código computacional está avançando de forma consistente com os objetivos inicialmente estabelecidos. Embora o projeto ainda esteja em fase de ajustes finais, os dados obtidos até agora indicam que o método adotado possui potencial para cumprir o papel de análise das rotas de condução quântica. Além disso, a automação do processo de geração de grids e o processamento dos dados têm evoluído de maneira satisfatória, sugerindo que o modelo pode, de fato, otimizar a investigação do transporte eletrônico em proteínas.

## CONTRIBUIÇÕES DOS AUTORES

D.C. desenvolveu o software, realizou análise dos dados, redigiu o trabalho, tratou as figuras. F.C. realizou a conceitualização do projeto, testou o código e fez a revisão do trabalho. Todos os autores contribuíram com a revisão do trabalho e aprovaram a versão submetida.

## AGRADECIMENTOS

Agradecemos ao PIBIFSP pelo financiamento do projeto de iniciação científica, ao INCT/INEO e à FAPESP, pelo financiamento de materiais consumíveis e permanentes; ao CENAPAD/SP e ao LNCC Santos Dumont pelo tempo computacional disponibilizado.

## REFERÊNCIAS

AMDURSKY, N.; MARCHAK, D.; SEPUNARU, L.; PECHT, I.; SHEVES, M.; CAHEN, D. *Electronic transport via proteins. Advanced Materials*, v. 26, n. 42, 2014, p. 1-20. DOI: <https://doi.org/10.1002/adma.201402304>.

BERTAGLIA, T.; KERR, E. F.; SEDENHO, G. C.; WONG, A. A.; COLOMBO, R. N. P.; MACEDO, L. J. A.; IOST, R. M.; FARIA, L. C. I.; LIMA, F. C. D. A.; TEOBALDO, G. B. M.; OLIVEIRA, C. L. P.; AZIZ, M. J.; GORDON, R. G.; CRESPILO, F. N. *Self-Gelling Quinone-Based Wearable Microbattery. Advanced Materials Technologies*, v. 9, n. 10, 2024, p. 35-48. DOI: <https://doi.org/10.1002/admt.202400623>.

BOSTICK, C. D.; MUKHOPADHYAY, S.; PECHT, I.; SHEVES, M.; CAHEN, D.; LEDERMAN, D. *Protein bioelectronics: a review of what we do and do not know. Reports on Progress in Physics*, v. 81, n. 2, 2018, p. 1-18. DOI: <https://doi.org/10.1088/1361-6633/aa85f2>.

SEPUNARU, L.; PECHT, I.; SHEVES, M.; CAHEN, D. *Solid-state electron transport across azurin: from a temperature-independent to a temperature-activated mechanism. Journal of the American Chemical Society*, v. 133, n. 6, 2011, p. 1-3. DOI: <https://doi.org/10.1021/ja109989f>.

ZHANG, B.; SONG, W.; PANG, P.; LAI, H.; CHEN, Q.; ZHANG, P.; LINDSAY, S. *The role of contacts in long-range protein conductance. Proceedings of the National Academy of Sciences*, v. 116, n. 13, 2019, p. 14-19. DOI: <https://doi.org/10.1073/pnas.1819674116>.

ZOTTI, L. A.; CABELLOS, J. L.; MIGUEL, D.; EVANGELIO, E.; ARAGÓ, J.; ORTÍ, E.; CUEVAS, J. C.; CERCÓS, A.; FERRER, J. *Long-range charge transport in single G-quadruplex DNA molecules. Biomolecules*, v. 9, n. 10, 2019, p. 5-9. DOI: <https://doi.org/10.3390/biom9100580>.

SMITH, J.; LEE, A.; BROWN, R. *Can One Define the Conductance of Amino Acids? Journal of Molecular Electronics*, v. 10, n. 4, 2023, p. 1-13. DOI: <https://www.mdpi.com/2218-273X/9/10/580>.