

A revolução silenciosa: como o AlphaFold redefiniu a biologia estrutural e abriu novas fronteiras na ciência

The silent revolution: how AlphaFold redefined structural biology and opened new frontiers in science

¹ Paulo Henrique Matayoshi Calixto  

¹ Instituto Federal Goiano

Resumo

Por meio século, a determinação da estrutura tridimensional de uma proteína a partir de sua sequência de aminoácidos representou um dos maiores desafios da biologia. A compreensão da função molecular, o desenvolvimento de fármacos e a engenharia de proteínas dependiam da superação desse obstáculo. Em 2020, o AlphaFold, um sistema de inteligência artificial desenvolvido pela DeepMind, emergiu como uma solução revolucionária, prevendo estruturas com precisão atômica até então inatingível por métodos computacionais. Esta revisão delineia a importância do problema do enovelamento de proteínas, os fundamentos conceituais do AlphaFold e seu impacto transformador. A disponibilização de centenas de milhões de estruturas através do AlphaFold DB facilitou o acesso à informação estrutural, acelerando a pesquisa em desenvolvimento de fármacos, biotecnologia e biologia fundamental. Contudo, apesar de revolucionário, o método possui limitações intrínsecas, como a predição de conformações estáticas e a incapacidade de modelar diretamente o efeito de ligantes, que exigem uma análise crítica e consciente dos seus resultados. O AlphaFold não é um ponto final, mas o início de uma nova era de «biologia digital», na qual a integração sinérgica entre a predição computacional e a validação experimental promete desvendar os mecanismos da vida em um nível de detalhe sem precedentes.

Palavras-chave

AlphaFold, Biologia Estrutural, Enovelamento Proteico, Inteligência Artificial.

Abstract

For half a century, determining the three-dimensional structure of a protein from its amino acid sequence has represented one of the greatest challenges in biology. The understanding of molecular function, drug development, and protein engineering depended on overcoming this obstacle. In 2020, AlphaFold, an artificial intelligence system developed by DeepMind, emerged as a revolutionary solution, predicting structures with atomic accuracy previously unattainable by computational methods. This review outlines the significance of the protein folding problem, the conceptual foundations of AlphaFold, and its transformative impact. The availability of hundreds of millions of structures through the AlphaFold DB has facilitated access to structural information, accelerating research in drug discovery, biotechnology, and fundamental biology. However, despite its revolutionary nature, the method possesses intrinsic limitations, such as the prediction of static conformations and the inability to directly model the effect of ligands, which demand a critical and conscientious analysis of its results. AlphaFold is not an endpoint, but the beginning of a new era of "digital biology," in which the synergistic integration of computational prediction and experimental validation promises to unveil the mechanisms of life at an unprecedented level of detail.

Keywords

AlphaFold, Structural Biology, Protein Folding, Artificial Intelligence.

1 INTRODUÇÃO

A biologia molecular moderna apoia-se no princípio fundamental de que existe uma relação íntima entre sequência, estrutura e função das proteínas. Em um trabalho clássico, Christian Anfinsen demonstrou que a informação necessária para que uma proteína adote sua conformação nativa funcional está inteiramente contida em sua sequência de aminoácidos, um conceito consagrado como o “Dogma de Anfinsen” (Anfinsen, 1973). Essa constatação abriu caminho para a ideia de que seria possível prever a estrutura tridimensional de uma proteína apenas a partir de sua sequência primária.

Entretanto, logo emergiu um desafio conceitual descrito por Cyrus Levinthal: se uma proteína testasse todas as conformações possíveis de maneira aleatória, o tempo necessário para alcançar o estado nativo ultrapassaria a própria idade do universo, o chamado “Paradoxo de Levinthal”. Essa aparente contradição reforçou que o enovelamento não ocorre ao acaso, mas segue trajetórias guiadas por um funil energético, no qual forças físico-químicas direcionam a proteína a estados intermediários estáveis até alcançar sua conformação funcional (Dill; Maccallum, 2012).

Apesar dos avanços nesse entendimento teórico, a predição da estrutura proteica permaneceu, por décadas, como um dos maiores problemas não resolvidos da biologia. Esse desafio mobilizou a comunidade científica em busca de métodos experimentais e computacionais capazes de decifrar o enigma do enovelamento, culminando em conquistas recentes que vêm transformando o campo da biologia estrutural.

Para a elaboração desta revisão narrativa, realizou-se um levantamento bibliográfico nas bases de dados PubMed, Web of Science e Google Scholar. A estratégia de busca empregou os descritores ‘AlphaFold’, ‘Protein Structure Prediction’, ‘Artificial Intelligence in Biology’ e ‘Deep Learning’, combinados através do operador booleano ‘AND’. O recorte temporal priorizou publicações compreendidas entre 2018 e 2025, período que abrange desde o surgimento da primeira versão do algoritmo até suas aplicações mais recentes. Adicionalmente, referências seminais anteriores a esse intervalo foram incluídas para a devida contextualização histórica dos desafios fundamentais da biologia estrutural.

2 EVOLUÇÃO DAS METODOLOGIAS DE PREDIÇÃO ESTRUTURAL E O IMPACTO DO ALPHAFOLD NA EFICIÊNCIA TEMPORAL

Historicamente, a elucidação *in silico* de estruturas proteicas dependia majoritariamente de métodos baseados em conhecimento prévio, como a modelagem por homologia e o *threading* (reconhecimento de dobramento). Tais abordagens exigiam a existência de uma estrutura experimentalmente resolvida com alto grau de similaridade sequencial depositada no *Protein Data Bank* (PDB) para servir como molde. Embora ferramentas clássicas como o MODELLER e o I-TASSER tenham estabelecido padrões fundamentais para a área, o processo era inerentemente limitado pela escassez de *templates* adequados para proteínas órfãs ou evolutivamente distantes (Sali; Blundell, 1993; Berman et al., 2000; Roy; Kucukural; Zhang, 2010). Além disso, a etapa de refinamento de modelos e a validação estereoquímica demandavam intervenção manual significativa e *expertise* do pesquisador, tornando o fluxo de trabalho suscetível a erros humanos e restrito a proteínas bem caracterizadas.

Na ausência de homólogos detectáveis, a obtenção de estruturas tridimensionais recorria a métodos *ab initio* ou *de novo*, que tentavam prever o dobramento baseando-se puramente em princípios físico-químicos e campos de força. Essas simulações eram computacionalmente dispendiosas, exigindo semanas ou até meses de processamento em *clusters* de alto desempenho para gerar modelos plausíveis para uma única cadeia polipeptídica pequena (Dill; Maccallum, 2012). O progresso nessa área foi monitorado bianualmente pela competição CASP (*Critical Assessment of Structure Prediction*), onde, por décadas, observou-se uma evolução lenta

na acurácia das predições para alvos difíceis, evidenciando o gargalo tecnológico existente entre a sequência primária e a estrutura terciária final (Simons et al., 1997; Alquraishi, 2019).

A introdução de arquiteturas de aprendizado profundo, consubstanciada no lançamento do *AlphaFold*, rompeu com as limitações temporais e de acurácia vigentes. Ao substituir a dependência estrita de moldes e simulações físicas longas por redes neurais treinadas em padrões evolutivos e coevolução de resíduos, o sistema permitiu a predição de estruturas com precisão atômica em questão de minutos ou horas, um ganho de eficiência de ordens de magnitude, quando comparado aos métodos tradicionais (Jumper et al., 2021). Essa aceleração viabilizou não apenas a análise de proteínas individuais, mas o mapeamento estrutural de proteomas inteiros, disponibilizando centenas de milhões de modelos no *AlphaFold DB* e transformando radicalmente o ritmo da pesquisa em biologia estrutural (Varadi et al., 2022; Thornton; Laskowski; Bork, 2021; Tunyasuvunakool et al., 2021).

3 POR QUE A ESTRUTURA IMPORTA?

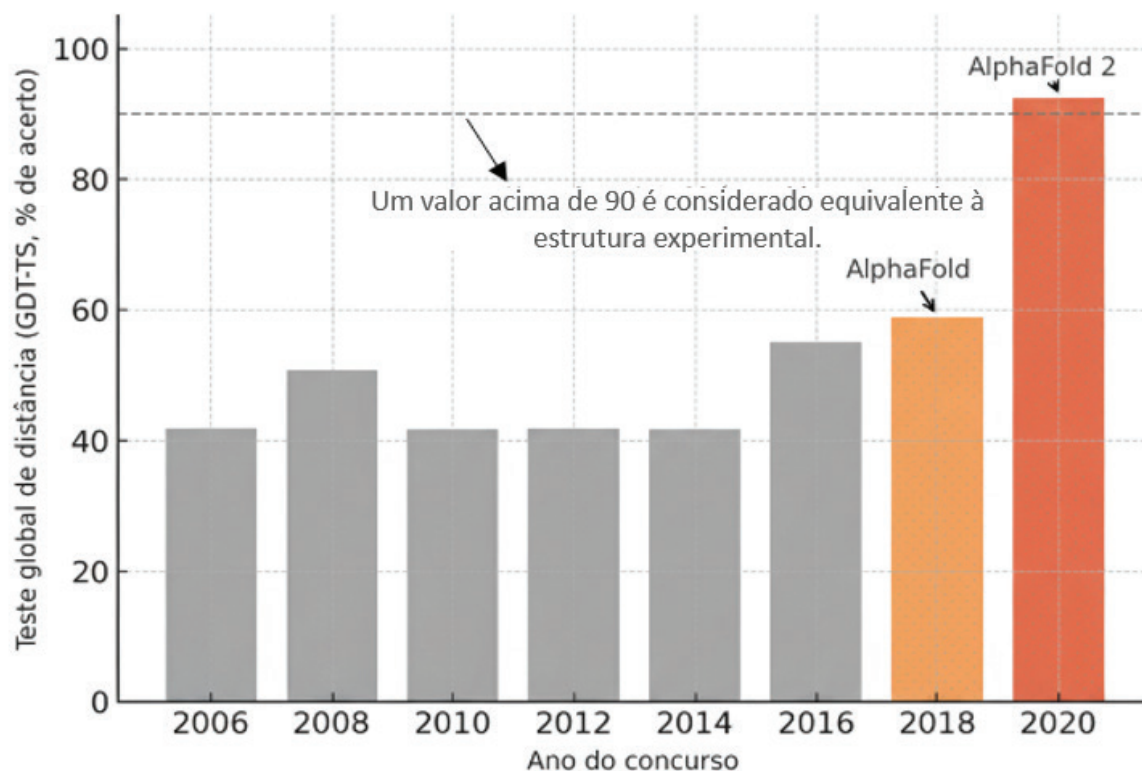
A máxima “estrutura dita a função” é central na bioquímica. A disposição precisa dos átomos em uma proteína cria sítios ativos em enzimas, superfícies de interação para a formação de complexos moleculares e receptores capazes de se ligar a hormônios ou fármacos. Sem o conhecimento estrutural, a compreensão dos processos biológicos é incompleta e o desenho racional de medicamentos ou novas enzimas é praticamente impossível (Khoury; Baliban; Floudas, 2014).

Por décadas, a biologia estrutural dependeu de técnicas experimentais sofisticadas, como a cristalografia de raios-X, a ressonância magnética nuclear (RMN) e, mais recentemente, a crio-microscopia eletrônica (crio-ME). Embora eficazes na determinação da estrutura tridimensional, essas técnicas enfrentam gargalos significativos: são trabalhosas, custosas e nem todas as proteínas, especialmente as de membrana, são passíveis de análise, deixando vastas áreas do proteoma humano e de outros organismos como ponto cego estrutural (Wlodawer et al., 2008).

Para preencher essas lacunas, métodos computacionais, como a modelagem por homologia, que utiliza uma estrutura conhecida de uma proteína similar como molde, e métodos *ab initio*, que tentam prever a estrutura a partir de princípios físicos, foram desenvolvidos. Embora úteis, sua precisão era limitada, especialmente para proteínas sem homólogos estruturais conhecidos (Guzenko et al., 2019).

O ponto de virada aconteceu no evento CASP (*Critical Assessment of Protein Structure Prediction*), uma competição que ocorre a cada dois anos, cujo objetivo é determinar estruturas tridimensionais de proteínas, visando mensurar o progresso científico na área. Em 2020, no CASP14, a *performance* do *AlphaFold 2* foi descrita como transformadora e sem precedentes pela comunidade científica (Figura 1). O sistema alcançou uma precisão comparável a de métodos experimentais para uma grande fração das proteínas-alvo, resolvendo efetivamente o problema do enovelamento de proteínas em uma escala antes inimaginável (Jumper et al., 2021; Cramer, 2021).

Figura 1. O avanço do AlphaFold na previsão da estrutura de proteínas. O gráfico abaixo ilustra a pontuação média no Teste de Distância Global (GDT_TS) ao longo dos anos na competição CASP. Em 2020, o algoritmo *AlphaFold 2*, da *DeepMind*, alcançou uma pontuação próxima de 90, um patamar considerado equivalente à precisão de métodos experimentais. Esse resultado representa um salto significativo em relação à sua versão anterior, o *AlphaFold* (2018), e a outros algoritmos competidores.



Fonte: Callaway (2020) (Adaptado)

4 O SALTO QUÂNTICO: O QUE É O ALPHAFOLD E COMO FUNCIONA?

A inovação do *AlphaFold* não reside em uma nova compreensão da física do enovelamento, mas na aplicação da inteligência artificial (IA), através de redes neurais profundas (*deep learning*) para reconhecer padrões complexos em dados biológicos. Ele aprendeu as regras do enovelamento, analisando o vasto banco de dados de estruturas e sequências de proteínas já conhecidas (Bouatta *et al.*, 2021).

O sistema *AlphaFold 2* opera através de dois módulos principais. O primeiro, denominado *Evoformer*, processa informações genéticas e evolutivas. Ele utiliza alinhamentos de múltiplas sequências (MSAs) para identificar sequências homólogas e inferir a coevolução entre resíduos de aminoácidos. A lógica subjacente é que a coevolução entre resíduos espacialmente distantes na sequência primária é um indicativo de proximidade na estrutura tridimensional. O *Evoformer*, por meio de um mecanismo de atenção, ou seja, redes neurais que permitem ao modelo ponderar a importância de diferentes partes dos dados de entrada, integra esses dados para capturar relações espaciais inter-residuais. O segundo módulo, a rede de estrutura, converte essas relações abstratas em um conjunto de coordenadas atômicas tridimensionais, gerando um modelo estrutural da proteína (Jumper *et al.*, 2021).

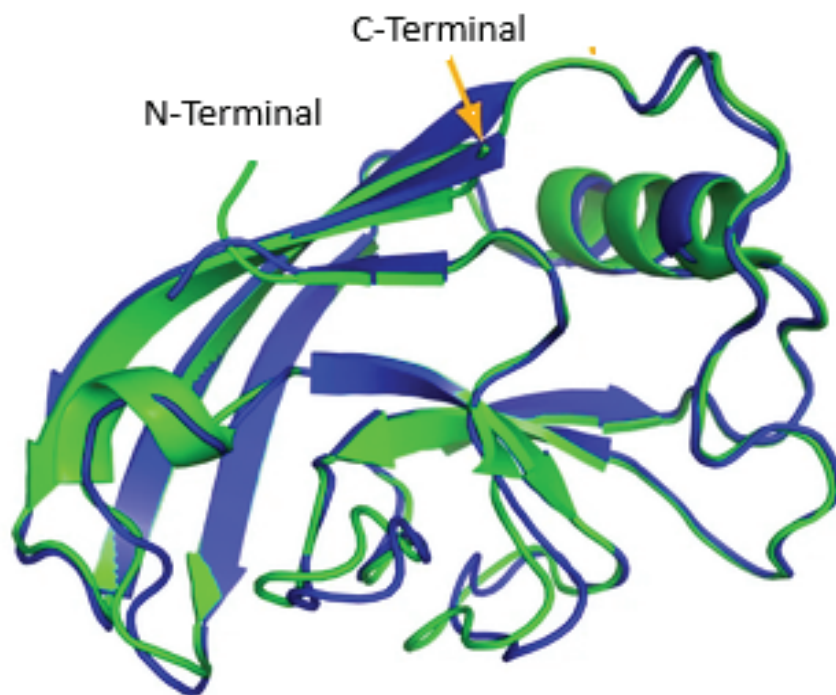
Um dos avanços mais importantes do *AlphaFold* é sua capacidade de autoavaliar a precisão de suas previsões, fornecendo métricas essenciais para o uso consciente dos modelos, a partir das análises do pLDDT (*predicted local distance difference test*) e do PAE (*predicted aligned error*). O pLDDT é um escore de confiança local que varia de 0 a 100, indicando a precisão da previsão para cada resíduo da proteína. Valores mais altos de pLDDT, geralmente acima de 90, sugerem uma alta confiança na previsão da posição atômica, enquanto valores abaixo de 50 indicam uma baixa precisão, com a conformação local sendo essencialmente aleatória. Já o PAE, é uma matriz que quantifica a distância esperada entre os resíduos, fornecendo um erro de alinhamento predito. Essa métrica é crucial para entender a incerteza posicional relativa entre os domínios da proteína, permitindo a identificação de regiões flexíveis ou de domínios mal preditos. Em conjunto, o pLDDT e o PAE são ferramentas indispensáveis para a análise e interpretação rigorosa das previsões de estruturas proteicas geradas pelo *AlphaFold* (Jumper *et al.*, 2021; Varadi *et al.*, 2022).

5 O IMPACTO TRANSFORMADOR DO ALPHAFOLD NA PESQUISA BIOLÓGICA

Historicamente, como mencionado, a determinação da estrutura tridimensional de proteínas é um processo caro, demorado e laborioso, exigindo técnicas experimentais sofisticadas. Consequentemente, a pesquisa em biologia estrutural era restrita a laboratórios de ponta com vastos recursos financeiros e tecnológicos. O *AlphaFold* quebrou essa barreira, possibilitando que estudos de biologia estrutural sejam realizados nos mais diversos contextos. A disponibilidade do banco de dados de estruturas e, mais recentemente, de ferramentas de previsão acessíveis a qualquer pesquisador com acesso à internet, independentemente de sua localização geográfica ou capacidade de financiamento, nivelou o campo da pesquisa. Laboratórios em economias emergentes e instituições com recursos limitados agora têm acesso a um vasto acervo de dados estruturais, permitindo que a pesquisa de ponta floresça em escala global, nos mais diferentes níveis da educação. Nesse sentido, o *AlphaFold* está acelerando o ritmo das descobertas, incentivando uma colaboração científica mais ampla e diversificada (Cramer, 2021; Cui *et al.*, 2025).

O impacto do *AlphaFold* vai muito além de seu uso direto na biologia estrutural. A capacidade de prever a forma de uma proteína a partir de sua sequência de aminoácidos com alta precisão (Figura 2) acelera significativamente a pesquisa básica em biologia evolutiva e biologia celular (Abriata, 2024). Na biologia evolutiva, o *AlphaFold* possibilita a reconstrução da história evolutiva de proteínas e famílias de proteínas de forma mais precisa. A comparação das estruturas de proteínas homólogas em diferentes espécies pode revelar como as funções biológicas evoluíram ao longo do tempo. Além disso, a previsão de estruturas para proteínas de organismos extintos, com base em sequências de DNA antigas, oferece um vislumbre sem precedentes da biologia do passado (Lin *et al.*, 2025).

Figura 2. Alinhamento estrutural entre o domínio de ligação ao receptor dependente de zinco de *Proteus mirabilis*. A estrutura em verde (PDB: 6Y4F) representa a estrutura obtida por métodos experimentais. Em azul, trata-se da estrutura resolvida pelo AlphaFold.



Fonte: Jumper et al. (2021) (Adaptado)

Na biologia celular, o *AlphaFold* oferece um novo e poderoso panorama para o funcionamento da célula. O conhecimento da estrutura de proteínas-chave, como enzimas, receptores e componentes do citoesqueleto, é fundamental para entender suas funções e interações. A predição rápida de estruturas acelera a formulação de hipóteses sobre como as proteínas atuam em vias metabólicas e processos celulares complexos. Isso é particularmente relevante para o estudo de doenças, em que a mutação de uma única proteína pode ter consequências graves. Com o *AlphaFold*, pesquisadores podem rapidamente modelar a estrutura de proteínas mutadas, gerando conhecimento sobre o mecanismo molecular de patologias como o câncer, doenças neurodegenerativas e infecciosas, e acelerando o desenvolvimento de terapias e medicamentos mais eficazes (Mubeen et al., 2024; Zhao et al., 2024).

6 APLICAÇÕES DO ALPHAFOLD

O *AlphaFold* demonstrou uma capacidade sem precedentes de prever estruturas de proteínas com alta precisão, um feito que culminou na disponibilização pública de um banco de dados com milhões de estruturas preditas, abrangendo praticamente todo o proteoma humano e de organismos modelo (Varadi et al., 2022). A alta disponibilidade de estruturas resolvidas com alto percentual de qualidade gerou grande impacto na produção científica, tanto em aplicações consolidadas quanto emergentes, como desenho racional de fármacos, biotecnologia e medicina de precisão (Rennie; Oliver, 2025).

6.1 Aceleração da identificação e validação de alvos terapêuticos

A primeira aplicação para as estruturas tridimensionais geradas pelo *AlphaFold* é a identificação de novos fármacos ou compostos. O desenho racional de fármacos baseia-se no conhecimento da estrutura da proteína-alvo para desenvolver moléculas que possam modular sua atividade de forma específica. Dessa forma, o *AlphaFold* está potencializando avanços notáveis nessa área, através da aceleração da identificação de novos fármacos, desenho de inibidores e moduladores (Ren *et al.*, 2023; Chen *et al.*, 2024). Um exemplo notável foi a sua aplicação durante a pandemia de COVID-19, na qual a predição rápida da estrutura de proteínas do SARS-CoV-2, como a proteína S (*spike*), auxiliou no desenvolvimento de vacinas e antivirais (Rennie; Oliver, 2025).

A partir de um modelo estrutural preciso em mãos, é possível empregar técnicas de docagem molecular (*docking*) para triar virtualmente bibliotecas de compostos em busca de moléculas que se liguem ao alvo. O *AlphaFold* tem sido utilizado para prever a estrutura de alvos sem estrutura experimental resolvida, permitindo o desenho de inibidores. Ren e colaboradores demonstraram a descoberta de um inibidor da quinase dependente de ciclina 20 (CDK20), um alvo para o carcinoma hepatocelular, utilizando uma estrutura predita pelo *AlphaFold* (Ren *et al.*, 2023; Foudah *et al.*, 2025).

Muitas doenças são causadas por interações anômalas entre proteínas. O *AlphaFold*, especialmente em suas versões mais recentes como o *AlphaFold-Multimer*, pode prever a estrutura de complexos proteicos, oferecendo um vislumbre detalhado das interfaces de interação. Esse conhecimento é crucial para o desenvolvimento de moléculas que possam estabilizar ou desestabilizar seletivamente essas interações (Guan; Keating, 2025).

6.2 Aplicações em Biotecnologia

A engenharia de proteínas para fins industriais, ambientais e terapêuticos é outra área que se beneficia enormemente da predição estrutural acurada. A otimização de enzimas para maior estabilidade, atividade ou especificidade é um objetivo central da biotecnologia. O conhecimento estrutural fornecido pelo *AlphaFold* permite a identificação de resíduos de aminoácidos-chave que podem ser mutados para aprimorar as características desejadas. Aplicações notáveis incluem o desenho de enzimas para a degradação de plásticos, como a PETase, na qual a compreensão da estrutura do sítio ativo é fundamental para melhorar sua eficiência (Lu *et al.*, 2022; Kim *et al.*, 2025).

O “*design de novo*” de proteínas com funções totalmente novas é uma fronteira da biotecnologia. Ao prever como sequências de aminoácidos se enovelarão, o *AlphaFold* serve como uma ferramenta de validação no ciclo de design, permitindo que os pesquisadores refinem suas sequências para alcançar o enovelamento e a função desejadas. Essa abordagem, combinada com modelos generativos de IA, abre portas para a criação de novos biossensores, catalisadores e materiais biomiméticos (Zhang *et al.*, 2025).

Além disso, o *design* de vacinas baseadas em epítopos estruturais tem sido impulsionado pela capacidade do *AlphaFold* de modelar antígenos de patógenos. Por exemplo, a estrutura predita de proteínas da superfície do parasito da malária (*Plasmodium falciparum*) está sendo usada para identificar regiões vulneráveis e projetar imunógenos capazes de gerar uma resposta imune protetora (Laurenson *et al.*, 2023).

6.3 Aplicações em medicina de precisão

A medicina de precisão visa adequar o tratamento médico às características individuais de cada paciente, incluindo seu perfil genético. O *AlphaFold* oferece uma ferramenta poderosa para interpretar o impacto funcional de variantes genéticas (Grech *et al.*, 2025).

O sequenciamento genômico frequentemente identifica mutações cujo impacto clínico é desconhecido. Ao modelar a estrutura da proteína com e sem a mutação, o *AlphaFold* pode prever como a variante afeta o enovelamento, a estabilidade ou as interações da proteína. Essa análise estrutural ajuda a classificar as variantes de significado incerto (VSI) como patogênicas ou benignas, auxiliando no diagnóstico de doenças genéticas e na estratificação de risco (Molotkov et al., 2024; Caswell et al., 2022).

A compreensão de como mutações específicas levam a fenótipos de doenças, como em diversos tipos de câncer e doenças neurodegenerativas, pode ser aprimorada com modelos estruturais precisos. Além disso, o *AlphaFold* pode ajudar a prever como variantes genéticas em enzimas metabolizadoras de fármacos ou nos próprios alvos terapêuticos podem influenciar a resposta de um indivíduo a um determinado medicamento, um pilar da farmacogenômica (Chang et al., 2024; Sarvepalli; Vadarevu, 2025).

7 CONSIDERAÇÕES FINAIS

O *AlphaFold* transformou a biologia estrutural de uma área de baixa produtividade para uma era de exploração em larga escala. A natureza silenciosa dessa revolução reside em sua origem puramente computacional: uma mudança de paradigma que, embora discreta em sua execução algorítmica, provocou um impacto transformador ao facilitar o acesso à informação estrutural e catalisar a descoberta científica em uma escala sem precedentes. O *AlphaFold* não substitui ou torna os pesquisadores obsoletos, pelo contrário, os empoderou com uma ferramenta extraordinária para gerar hipóteses, interpretar dados complexos e focar seus esforços experimentais em questões biológicas mais profundas. Estamos testemunhando o nascimento de uma nova biologia, na qual a sinergia entre a inteligência artificial e a experimentação humana nos permitirá mapear os mecanismos moleculares da vida com uma clareza nunca antes imaginada.

Declaração de uso de IA generativa

O autor declara o uso de IA generativa na revisão textual e formatação.

REFERÊNCIAS

- ABRIATA, L. The Nobel Prize in Chemistry: past, present, and future of AI in biology. **Communications Biology**, v. 7, n. 1, p. 1409-1402, October 2024.
- ANFINSEN, C. B. Principles that govern the folding of protein chains. **Science**, v. 181, n. 4096, p. 223-230, July 1973.
- ALQURAISHI, M. AlphaFold at CASP13. **Bioinformatics**, v. 35, n. 22, p. 4862–4865, November 2019.
- BERMAN, H. M. et al. The Protein Data Bank. **Nucleic Acids Research**, v. 28, n. 1, p. 235–242, January 2000.
- BOUATTA, N. et al. Protein structure prediction by AlphaFold2: are attention and symmetries all you need?. **Biological Crystallography**, v. 77, n. 8, p. 982-991, August 2021.
- CALLAWAY, Ewen. 'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. **Nature**, v. 588, p. 203-204, November 2020.
- CASWELL, Richard C. et al. Assessing the clinical utility of protein structural analysis in genomic variant classification: experiences from a diagnostic laboratory. **Genome Medicine**, v. 14, n. 1, p. 77-93, July 2022.
- CHANG, L. et al. Revolutionizing peptide-based drug discovery: advances in the post-AlphaFold era. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, v. 14, n. 1, p. e1693-1715, January 2024.
- CHEN, L. et al. AI-driven deep learning techniques in protein structure prediction. **International Journal of Molecular Sciences**, v. 25, n. 15, p. 8426-8447, August 2024.
- CRAMER, P. AlphaFold2 and the future of structural biology. **Nature Structural & Molecular Biology**, v. 28, n. 9, p. 704-705, August 2021.
- CUI, X. et al. Beyond static structures: protein dynamic conformations modeling in the post-AlphaFold era. **Briefings in Bioinformatics**, v. 26, n. 4, p. bbaf340-352, July 2025.
- DILL, K. A.; MACCALLUM, J. L. The Protein-Folding Problem, 50 Years On. **Science**, v. 338, n. 6110, p. 1042-1046, November 2012.
- FOUDAH, A. et al. Quantum chemical optimization and residue-specific stabilization of CDK20 inhibitors in hepatocellular carcinoma. **Molecular Diversity**, p. 1-22, September 2025.
- GRECH, S. et al. The Role of Artificial Intelligence in Identifying NF1 Gene Variants and Improving Diagnosis. **Genes**, v. 16, n. 5, p. 560-588, May 2025.
- GUAN, L.; KEATING, A. How AlphaFold and related models predict protein-peptide complex structures. **bioRxiv**, p. 1-34, June 2025.
- GUZENKO, D. et al. Assessment of protein assembly prediction in CASP13. **Proteins: Structure, Function, and Bioinformatics**, v. 87, n. 12, p. 1190–1199, December 2019.
- JUMPER, J. et al. Highly accurate protein structure prediction with AlphaFold. **Nature**, v. 596, n. 7873, p. 583-589, July 2021.

KHOURY, G. A.; BALIBAN, R. C.; FLOUDAS, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. **Scientific Reports**, v. 4, n. 1, p. 1-8, September 2014.

KIM, D. et al. Parametrically guided design of beta barrels and transmembrane nanopores using deep learning. **BioRxiv**, p. 1-23, January 2025.

LAURENSEN, A. et al. *In Silico* Prediction of *Plasmodium falciparum* PfRipr Epitopes as Vaccine Candidates. **Journal of the Pediatric Infectious Diseases Society**, v. 12, n. Supplement 1, p. S5-S6, November 2023.

LU, H. et al. Machine learning-aided engineering of hydrolases for PET degradation. **Nature**, v. 604, n. 7907, p. 662-667, April 2022.

LIN, Pei-Yu et al. Analysing protein complexes in plant science: insights and limitation with AlphaFold 3. **Botanical Studies**, v. 66, n. 1, p. 14-26, May 2025.

MOLOTKOV, I. et al. Making sense of missense: challenges and opportunities in variant pathogenicity prediction. **Disease Models & Mechanisms**, v. 17, n. 12, p. 1-4, December 2024.

MUBEEN, H. et al. Insights into AlphaFold's breakthrough in neurodegenerative diseases. **Irish Journal of Medical Science**, v. 193, n. 5, p. 2577-2588, June 2024.

REN, F. et al. AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor. **Chemical Science**, v. 14, n. 6, p. 1443-1452, January 2023.

RENNIE, L.; OLIVER, R. Emerging frontiers in protein structure prediction following the AlphaFold revolution. **Journal of the Royal Society Interface**, v. 22, n. 225, p. 20240886-20240903, April 2025.

ROY, A.; KUCUKURAL, A.; ZHANG, Y. I-TASSER: a unified platform for automated protein structure and function prediction. **Nature Protocols**, v. 5, n. 4, p. 725–738, April 2010.

SALI, A.; BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **Journal of Molecular Biology**, v. 234, n. 3, p. 779–815, December 1993.

SARVEPALLI, S.; VADAREVU, S. Role of artificial intelligence in cancer drug discovery and development. **Cancer Letters**, p. 217821-217832, September 2025.

SIMONS, K. T. et al. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. **Journal of Molecular Biology**, v. 268, n. 1, p. 209–225, April 1997.

THORNTON, J. M.; LASKOWSKI, R. A.; BORK, P. AlphaFold heralds a data-driven revolution in biology and medicine. **Nature Medicine**, v. 27, n. 10, p. 1666–1669, October 2021.

TUNYASUVUNAKOOL, K. et al. Highly accurate protein structure prediction for the human proteome. **Nature**, v. 596, n. 7873, p. 590–596, July 2021.

VARADI, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. **Nucleic acids research**, v. 50, n. D1, p. D439-444, January 2022.

WLODAWER, A. *et al.* Protein crystallography for non-crystallographers, or how to get the best (but not more) from published structures. **The FEBS Journal**, v. 274, n. 19, p. 4971-4992, January 2008.

ZHANG, G. *et al.* The Role of AI-Driven De Novo Protein Design in the Exploration of the Protein Functional Universe. **Biology**, v. 14, n. 9, p. 1268-1298, September 2025.

ZHAO, X. *et al.* AlphaFold's Predictive Revolution in Precision Oncology. **AI in Precision Oncology**, v. 1, n. 3, p. 160-167, June 2024.