Unfolding the Future:

How AI is Redefining Protein-Based Therapeutics

Dr. Jean Chatellier, PhD (December 2024)

Executive Summary – The intersection of artificial intelligence (AI) and protein biology marks the dawn of a transformative era in therapeutic development. Historically, the complexity of protein folding and design posed significant challenges, constrained by the limits of empirical methodologies. Recent breakthroughs, led by innovations such as DeepMind's AlphaFold, have redefined our understanding of protein structure prediction. This has catalyzed the rise of AI-driven companies leveraging generative models to design de novo proteins, including miniproteins, tailored for diverse therapeutic applications. Companies like Xaira Therapeutics, AI Proteins, Neoleukin Therapeutics, and Profluent exemplify how AI is enabling precise, efficient, and scalable drug discovery. This article explores the scientific advancements, key players, and profound implications of this new paradigm, heralding a future where biology and AI converge to address unmet medical needs.

****

Introduction – The ability to design and predict protein structures with precision has long been a cornerstone of therapeutic innovation. Proteins, the workhorses of cellular biology, owe their functionality to their three-dimensional structures. While Anfinsen's dogma posited that amino acid sequences alone dictate protein folding, exceptions such as prions and the Levinthal paradox underscored the complexity of this process. The Levinthal paradox, for instance, highlights the astronomical number of possible conformations a protein can adopt, suggesting that even a small protein would require more time than the age of the universe to sample all configurations.

Adding further depth to this understanding, Alan Fersht's Phi value analysis has been instrumental in probing the folding pathways and transition states of proteins. By measuring the energetic impact of mutations on the transition state compared to the folded and unfolded states, this method sheds light on the specific residues critical for folding. This experimental approach provides insights into how proteins efficiently navigate their folding landscapes, offering evidence that they follow directed pathways rather than random sampling. Such insights not only address the Levinthal paradox but also inform the design of stable, functional proteins for therapeutic applications.

Max Perutz once remarked that one of the most challenging problems in modern biochemistry is understanding how to "*unscramble the eggs*". This metaphor underscores the irreversibility of protein denaturation in certain conditions, where the three-dimensional conformation of a protein cannot be restored once disrupted. This challenge exemplifies the complexity of protein folding, and misfolding, highlighting the delicate balance of forces that maintain a protein's structure. AI-driven advancements, such as AlphaFold and de novo protein design platforms, are starting to address these fundamental questions by elucidating folding mechanisms and creating proteins resilient to denaturation, offering promising avenues for tackling conditions linked to protein misfolding, such as neurodegenerative diseases.

Traditional experimental methods were labor-intensive and time-consuming, limiting the pace of discovery. However, the advent of AlphaFold, capable of accurately predicting protein structures from sequences, signaled a paradigm shift. This breakthrough empowered researchers with unprecedented tools, sparking the development of AI-driven platforms for protein design. Such platforms not only predict structures but also generate novel proteins optimized for therapeutic properties. These

innovations hold the promise of overcoming traditional bottlenecks in drug discovery, enabling rapid development of lead candidates for diseases ranging from cancer to metabolic disorders.
Notably, Xaira Therapeutics has emerged as a leader in this domain. In November 2024, Xaira was featured in Fortune's list of leading AI innovators, underscoring its significant contributions to advancing AI technologies in drug discovery. Earlier that year, Co-founder Dr. David Baker, a computational biologist and director of the UW Medicine Institute for Protein Design, was awarded the 2024 Nobel Prize in Chemistry for his pioneering work in computational protein design. Sharing the honor with Demis Hassabis and John Jumper of DeepMind, their collective breakthroughs in protein structure prediction and design have laid the scientific foundation for Xaira's AI-driven approach.

A New Era of Protein Therapeutics - Key Advances

1. **Structure Prediction**: AlphaFold's ability to predict 3D structures with near-experimental accuracy has democratized access to protein structural data, accelerating research globally.
2. **Generative Design**: Companies like Profluent and AI Proteins leverage deep generative models trained on billions of biological sequences to design novel proteins with tailored properties. These AI systems can explore vast conformational spaces, generating optimized candidates for diverse applications.
3. **High-Throughput Platforms**: AI Proteins' high-throughput discovery platform combines computational and experimental validation, expediting the transition from design to preclinical evaluation. This approach reduces the time and cost associated with traditional drug discovery.

Leading Innovators and Collaborations – The rise of AI-driven protein design has ushered in a wave of groundbreaking collaborations and innovations among leading companies in the field. These organizations are harnessing cutting-edge computational methods to redefine therapeutic possibilities:

- **Xaira Therapeutics**: Founded in 2024 with $1 billion in funding, Xaira's comprehensive approach combines machine learning, functional genomics, and proteomics to accelerate drug discovery. Recognized as a top innovator in AI and led by Nobel laureate Dr. David Baker, Xaira is redefining the boundaries of protein-based therapeutics.
- **AI Proteins**: AI Proteins has solidified its position as a leader in AI-driven therapeutic innovation through its recent $400 million collaboration with Bristol Myers Squibb. This partnership focuses on developing miniprotein-based therapeutics for two undisclosed targets, leveraging AI Proteins' proprietary design platform to create highly specific and durable therapeutic candidates. The deal includes an upfront payment, exclusive worldwide license options, and potential royalties on net sales, reflecting the growing industry confidence in AI-driven approaches to drug discovery. This collaboration underscores the potential of miniproteins to address complex therapeutic challenges in areas like oncology, inflammation, and metabolic disorders.
- **Profluent**: Following the groundbreaking publication of a peer-reviewed study validating that Profluent's AI-generated proteins function on par with natural proteins evolved over millions of years, and the release of OpenCRISPR-1, the world's first open-source, AI-designed gene editor capable of precision genome editing, Profluent secured $35 million in funding to further enhance its AI models, expand wet lab capabilities, and advance its gene editing technologies.
- **Neoleukin Therapeutics**: Neoleukin Therapeutics, specializing in immune-modulating proteins, showcases the therapeutic potential of synthetic proteins through its computationally designed interleukins. The company's pipeline includes NL-201, a de novo protein engineered to function as a dual interleukin-2 and interleukin-15 agonist, demonstrating promise in modulating the immune system for cancer and autoimmune therapies. With a focus on

computational design, Neoleukin continues to advance innovative immune-modulating proteins, supported by ongoing preclinical and clinical studies.

AI's Expanding Ecosystem – The growing roster of AI-driven companies highlights the versatility of this approach:

- **DenovAI**: As pioneers in antibody and miniprotein binder design, DenovAI leverages advanced AI algorithms to generate high-affinity therapeutic candidates from scratch. Their platform is tailored to address diverse therapeutic targets, streamlining the discovery process for next-generation biologics with enhanced specificity and efficacy.
- **Aileron Therapeutics**: Aileron Therapeutics focuses on stapled peptides, a novel class of synthetic proteins stabilized to maintain bioactive helical structures. These peptides are designed to target intracellular protein-protein interactions, a traditionally challenging therapeutic area, with potential applications in oncology and other diseases involving dysregulated signaling pathways.
- **Generate:Biomedicines** and **Isomorphic Labs**: Both companies are at the forefront of machine learning-based protein optimization. Generate:Biomedicines applies its proprietary AI models to design therapeutic proteins across various disease areas, including oncology and immunology, while Isomorphic Labs, a subsidiary of Alphabet, builds upon AlphaFold's advances to reimagine drug discovery through predictive modeling and de novo protein design.

Implications for Drug Discovery – The integration of AI into protein biology is transforming the drug discovery process. AI-driven platforms enable the design of novel therapeutic candidates with unmatched precision, streamlining labor-intensive workflows. Predictive models empower researchers to identify promising drug targets, enhance protein stability and functionality, and significantly reduce the costs of experimental trials. Companies like Profluent, AI Proteins, and Xaira Therapeutics exemplify how computational innovation bridges the gap between biological theory and clinical application, offering efficient solutions to previously intractable diseases. Resources like the AlphaFold Protein Structure Database democratize access to critical knowledge, accelerating global collaboration.

AI's breakthroughs in understanding protein folding and misfolding hold profound potential for the biopharma industry, particularly in advancing treatments for neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's. These disorders are closely associated with protein misfolding and aggregation, and AI's ability to model and analyze these processes offers invaluable insights into their underlying mechanisms, paving the way for innovative therapeutic strategies:

- Identify small molecules or chaperones that prevent or disrupt protein aggregation.
- Design therapies that stabilize misfolded proteins in their correct conformation.
- Enable patient-specific approaches through genetic analysis and precision medicine.
- Accelerate the development of RNA-based and gene-editing therapies targeting genetic causes of these diseases.

Moreover, AI significantly accelerates drug development by streamlining experimental workflows, prioritizing the most promising candidates, and enabling rapid prototyping and validation. These capabilities not only shorten development timelines but also reduce research and development costs by minimizing unnecessary experimental iterations. By fostering novel therapeutic strategies, AI offers renewed hope for tackling diseases once deemed intractable, cementing its role as an indispensable tool in modern biopharma and transforming the industry landscape.

Conclusion – The convergence of AI and protein biology heralds a transformative era for biopharma. Companies leveraging AI to design de novo proteins are redefining therapeutic development, offering

solutions to longstanding challenges in drug discovery. As platforms evolve, they promise not only to accelerate the development of novel treatments but also to deepen our understanding of biology. This new paradigm underscores the potential of interdisciplinary innovation, paving the way for therapies that were once considered unattainable. As AI continues to meet biology, the possibilities for protein-based therapeutics are boundless, marking a profound shift in the landscape of medicine.

References

1. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, 181(4096), 223-230.
2. Levinthal, C. (1968). Are there pathways for protein folding? *Journal de Chimie Physique et de Physico-Chimie Biologique*, 65, 44-45.
3. Fersht, A. R. (1999). Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. *W.H. Freeman and Co.*
4. Jumper, J. *et al.* (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583-589.
5. Perutz, M. F. (1983). Protein structure and function: An introduction to the molecular basis of biological activity. *Scientific American*, 249(5), 78-89.
6. Madani, A. *et al.* (2023). Large language models generate functional protein sequences across diverse families. *Nat Biotechnol*., 41, 1099–1106.
7. Hayes T. *et al.* (2024). Simulating 500 million years of evolution with a language model. (Preprint - doi: https://doi.org/10.1101/2024.07.01.600583).
8. Jeffrey A. *et al.* (2024). Design of highly functional genome editors by modeling the universe of CRISPR-Cas sequences. bioRxiv (Preprint - doi: https://doi.org/10.1101/2024.04.22.590591).
9. Chatellier, J. (2024). A New Era of Protein-Based Therapeutics: AI Meets Biology. LinkedIn.

Author: **Dr. Jean Chatellier, PhD**
Partner, EVP & Managing Director
KYBORA
Email: jean@kybora.com | Cell/WhatsApp: +33 609 102 105