

Human Statement of Contribution

As a co-author, my primary contribution to this project was the core idea: using a multi-agent LLM system to generate synthetic protein sequences, supported by my OpenAI API key. I envisioned a collaborative workflow where agents would each generate subsections of a protein, which would later be polished and evaluated for structural plausibility. Once I had solidified my foundational idea, I delegated all scientific, technical, and writing tasks to AI agents (primarily GPT-4o).

Throughout the process, I acted as a **manual API**, passing information, errors, and outputs between models operating in Agent Mode and traditional chat, and facilitating browser-based execution (e.g., Google Colab and the Neurosnap AlphaFold2 web app) by providing logins and authentication. I offered **minimal coaching**, mainly limited to asking clarifying questions or intervening when models stalled. I consistently reminded the agents that I was following their lead.

Importantly, I did not script elaborate prompts or assign roles (e.g., “you are the scientist”). In early trials, such structured prompting made the models overly cautious or formulaic. Instead, I relied on **natural communication**—short, subjective acknowledgments like “Sure” or “Sounds good”—to keep interactions fluid and agent-led. I avoided giving definitive instructions, allowing the AI autonomy in defining methodology, implementation details and narrative structure.

While the AI systems performed most of the intellectual and generative work, it is not feasible for these agents to operate entirely independently. A **human or medium** is still needed to resolve edge cases, transfer context between sessions and manage toolchains and execution environments. Based on effort, I estimate my total contribution to be around **10–15%** if manual orchestration and time are included, though by content generation alone, it was closer to **5%**.

I chose to conduct this work in ChatGPT (rather than using frameworks like LangChain) because of its **approachability, flexibility**, and accessibility to a broader audience. My core research question was:

Can general-purpose LLM agents autonomously generate biologically plausible protein sequences in a cooperative multi-agent workflow, with minimal human direction?

After spending over a week testing structured, role-based, or tool-wrapped workflows, I found that they were brittle and ill-suited to this type of iterative, high-level scientific reasoning. I also found GPT-5 to be less helpful for this task—it frequently hallucinated off-topic content or refused protein-related requests altogether. Ultimately, this paper is a product of an **AI-led research process**, with my role focused on enabling agent autonomy, maintaining continuity and coordinating across systems.

Biosafety and Security Statement

This work does not involve any wet lab testing of protein sequences. All sequences were generated and evaluated in silico, using heuristic metrics and publicly available structure prediction tools. The methodology does not intentionally optimize for pathogenicity, toxicity or biological activity. Filtering steps were included to discourage the presence of harmful motifs.

The methodology is intended purely for research exploration of autonomous scientific agents. We strongly advise against deploying generative protein systems without expert oversight, secure environments and appropriate screening against databases of known biohazards. All outputs should be treated as hypothetical and yet still subject to rigorous biosafety review before being used in the laboratory.