In 2-4 double-spaced pages, describe your research project and conference participation, being sure to address the following proposal components:

1. Overview of conference and your conference participation (ex. description of your participation, motivation for attending, what you hope to gain from the experience, etc.)
2. How this conference presentation fits into your Harvard academic trajectory and/or research goals.
3. Objective, significance, and implications of research project
4. Detailed plan for research (the status of the research project in question; how does the conference presentation fit into the overall research plan?)
5. Use of funds (provide detailed estimates and quotes of all anticipated presentation expenses; it’s okay if total goes beyond $1000)
6. Faculty involvement

Dear URAF Conference Funding Committee,

My name is Iñaki Arango. I am a junior at Harvard college pursuing a joint concentration in Computer Science and Statistics and a concurrent masters in Computer Science at the Harvard John A. Paulson School of Engineering and Applied Sciences (SEAS).

Thank you for considering my request for URAF Conference Funding to support our project, "Multi-objective generative AI for designing novel brain-targeting small molecules." This work was conducted under the guidance of Prof. Nada Amin, Assistant Professor of Computer Science at Harvard University, as part of COMPSCI 252R: Advanced Topics in Programming Languages and COMPSCI 91R: Supervised Reading and Research during the Fall 2023 and Spring 2024 semesters.

In our study, we focus on identifying small molecule drug candidates for central nervous system (CNS) diseases that currently lack effective treatments. The blood-brain barrier (BBB) poses a significant challenge in delivering drugs to the CNS, hindering both diagnosis and treatment. Computational methods that generate BBB-permeable lead compounds *in silico* could be valuable tools in the CNS drug design process. However, in practical applications, BBB penetration alone is not sufficient; molecules must also perform a desired function, such as binding to a specific target or receptor in the brain, while being safe and non-toxic for human use.

To address these challenges, we employed multi-objective generative AI to synthesize drug-like, BBB-permeable small molecules with high predicted binding affinity to a disease-relevant CNS target. We specifically focused on designing molecules with predicted bioactivity against dopamine receptor D2, which is the primary target for most clinically effective antipsychotic drugs. After training several graph neural network-based property predictors, we adapted SyntheMol (Swanson et al., 2024), a recently developed Monte Carlo Tree Search-based algorithm originally designed for antibiotic discovery, to perform a multi-objective guided traversal over an easily synthesizable molecular space.

As a result, we designed a library of 26,581 novel and diverse small molecules containing hits with high predicted BBB permeability, favorable predicted safety and toxicity profiles, and the potential for straightforward synthesis and experimental validation in the wet lab. We also validated top-scoring molecules using molecular docking simulation against the D2 receptor, demonstrating predicted binding affinity comparable to risperidone, a clinically prescribed D2-targeting antipsychotic. We hope that our SyntheMol-based computational approach will enable the discovery of novel neurotherapeutics for currently intractable CNS disorders in the future.

I completed this project together with fellow undergraduate and co-first author Ayush Noori. Ayush and I divided the work for this project equally: for example, we designed the algorithmic paradigm together, Ayush curated the labeled data to train molecular property predictors (*e.g.*, see his contribution to the Therapeutics Data Commons at [mims-harvard/TDC PR #215](https://github.com/mims-harvard/TDC/pull/215)), I developed the training loop and adapted the Monte Carlo Tree Search algorithm, Ayush performed the molecular simulations, and we drafted the paper together. We also plan to jointly present the paper at any future conferences.

Our research paper was submitted to the 2024 Generative and Experimental Perspectives for Biomolecular Design (GEM) workshop, which is part of the 12th International Conference on Learning Representations (ICLR), scheduled to take place from May 7th to May 11th, 2024, in the city of Vienna, Austria. ICLR is considered one of the most influential and highly regarded machine learning conferences worldwide. We were both thrilled and grateful when our paper, which underwent a rigorous double-blind review process, was accepted for presentation at ICLR GEM.

One of the anonymous reviewers provided glowing feedback, stating, “This work introduces a groundbreaking AI-driven approach for designing brain-targeting small molecules, overcoming BBB challenges and enhancing CNS drug delivery, notably through its novel methodology and promising implications for treating CNS disorders... The paper presents a novel and potentially impactful approach to CNS drug discovery, with strong methodological rigor and clear presentation. Given its contributions to the field and the promising results demonstrated, I recommend acceptance.” The reviewer further indicated that our paper was ranked in the top 15% of accepted papers and gave it a strong recommendation for acceptance. We are truly humbled by this recognition and eagerly anticipate the opportunity to share our work with the scientific community at ICLR GEM.

As an international student from Argentina, attending the GEM workshop is an incredible opportunity, but it also presents significant financial challenges. Argentina's current economic instability, characterized by high inflation rates and currency devaluation, has greatly limited my family's ability to financially support my education in the United States.

I have been working diligently to secure internships and research opportunities to help fund my studies, but these efforts fall short of covering the substantial costs associated with attending an international conference like GEM. The expenses, including international travel, accommodation, and registration fees, are considerably higher than what I can afford on my own, especially given the workshop's location in Vienna, Austria.

The GEM workshop offers a unique platform for me to engage with the global scientific community, present my research, and gain valuable insights that are crucial for my academic and professional growth. Attending the GEM workshop would be a transformative experience for me as a young researcher. My work on "Multi-Objective Generative AI for Designing Novel Brain-Targeting Small Molecules" represents my first substantial academic research experience, and it has ignited a strong interest in potentially pursuing a PhD in the future. This project has opened my eyes to the incredible potential of combining computational methods with biology to make a meaningful impact on people's lives.

By presenting my research at such a prestigious international conference, I hope to receive invaluable feedback and insights from experts in the field. Their input could significantly refine my approach and elevate the quality of my future research endeavors. Moreover, the workshop provides a unique opportunity to connect with other passionate individuals who share my enthusiasm for using computational tools to advance biological research and ultimately improve patient outcomes.

Many of the invited speakers and panelists have done, or are doing, things that I am extremely interested in, and that I want to do in the future. I am interested in meeting with Rohit Singh, and discussing how he found his computational biology skills useful for finance, and why he decided to start a company for predicting default rates. I want to talk to Neolia Ferruz, and learn how her lab is using generative models for designing proteins, which could potentially be very applicable to the use of diffusion models for de novo drug design. I would also like to connect with Gevorg Grigoryan, and hear the story of how he spun out his research into a successful startup. Finally, I want to meet with Alex Rives and understand how he started two companies that went public before he even got his PhD. I believe everyone at the conference has a different story to tell, with a different learning to be made, and I would like to tell mine.

The discussions, sessions, and networking opportunities at the conference will be incredibly valuable for a developing researcher like me. It is an ideal place to explore my growing interest in research and better understand whether a PhD aligns with my long-term goals. This experience could be a defining moment in my academic journey, providing me with the clarity and direction I need at this early but critical stage.

Above all, the GEM workshop will allow me to immerse myself in a community of like-minded individuals who share my goals and aspirations. The connections and collaborations I establish there have the potential to shape my entire future in this field, and I am excited about the prospect of being part of such a vibrant and impactful community.

I would be deeply grateful for the financial assistance provided by URAF, which would enable me to attend this remarkable event. If granted, the URAF conference funding would cover the following expenses:

1. Round-trip airfare from Boston, MA to Vienna, Austria, which is estimated to cost approximately $2,160.
2. Accommodation in Vienna for the duration of the conference, from May 6th to May 11th, 2024, with an estimated cost of $2,490, including taxes.
3. The discounted academic registration fee for full-time students attending ICLR 2024, which amounts to $475.

The total amount of funding requested for this opportunity is $5,125. Once again, I extend my sincere gratitude for considering my application. I eagerly hope to have the chance to participate in this exceptional and unparalleled opportunity for learning and professional development.

Sincerely,

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Iñaki Arango

A.B./S.M. Candidate in Computer Science and Statistics

Harvard College