

# Research Portfolio

Kavin Ramadoss, Sunset High  
School, Portland, Oregon

---

**Intended Major:** Biochemistry

**Research Focus:** My research focuses on the intersection of biochemistry and computational modeling, focusing on drug discovery for infectious transmitted diseases (malaria, TB, HIV) and neurodegenerative diseases such as Alzheimer's. I use a multitude of computational including molecular docking and dynamic simulations, to study protein-ligand interactions.

# Table of Contents

---

## **Research I: Enhancing Ancient Greek Inscription Lemmatizer (2024)**

Research I. Section I: Research Overview

Research I. Section II: Preprint + Presentation Materials + Computational Models

## **Research II: Augmented Reality Surgical Navigation System (2023 - 2024)**

Research I. Section I: Research Overview

Research I. Section II: Research Paper + Presentation Materials + Computational Models

## **Research III: Biocomputational Framework for Identifying Novel Malaria Inhibitors (2024-2025)**

Research I. Section I: Research Overview

Research I. Section II: Research Paper + Presentation Materials + Computational Models

## **Research IV: Collagen Segmentation in Glaucoma-Affected Tissue (2024)**

Research I. Section I: Research Overview

Research I. Section II: Computational Models

## **Research V: Binding Site Discovery in Ab42 to Prevent Early-Onset Alzheimer's (2025 - Present)**

Research I. Section I: Research Overview

Research I. Section II: Research Abstract + Computational Materials

# Table of Contents Cont.

---

## **Research VI: Enhancing DiffSBDD for Optimal Pharmacokinetic Properties (2024 - 2025)**

Research I. Section I: Research Overview

Research I. Section II: Research Paper + Computational Models

## **Research VI: HTML Website for Nonprofit: HearUsOut (2024)**

Research I. Section I: Project Overview

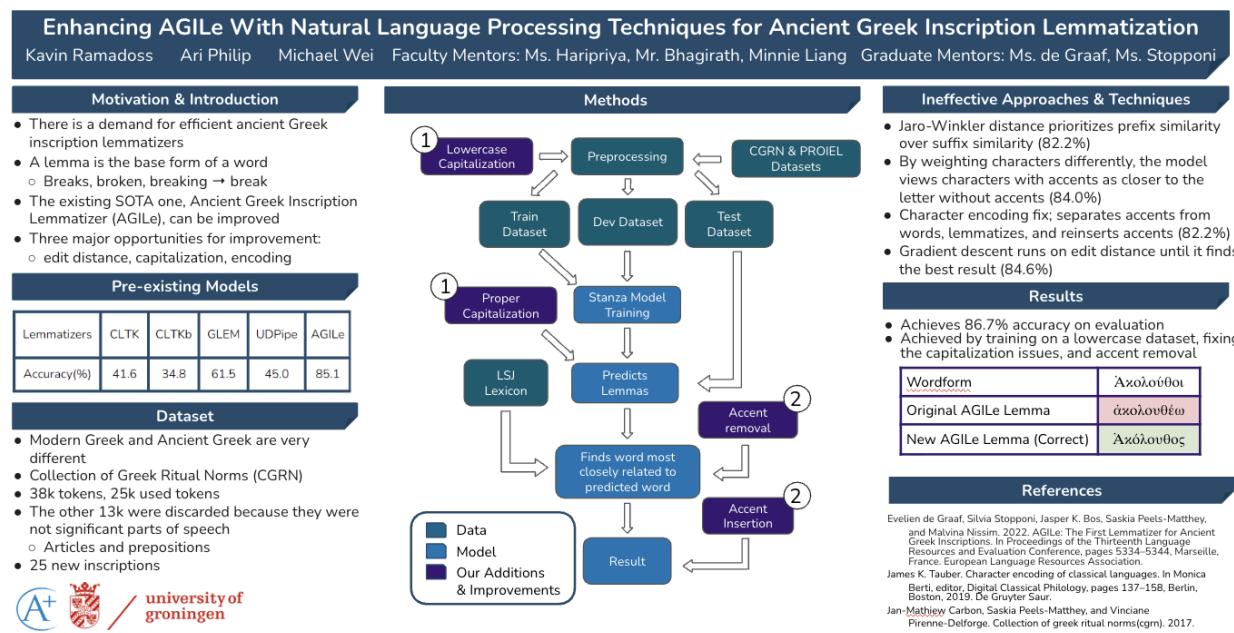
Research I. Section II: Presentation Materials

# I: Enhancing AGILe with NLP

I conducted research titled *Enhancing AGILe with Natural Language Processing Techniques for Ancient Greek Inscription Lemmatization* during high school in collaboration with researchers at the University of Groningen. As a co-author, I investigated systematic failures in the existing AGILe lemmatizer and directly modified its Python-based NLP pipeline, which is built on the Stanza framework and trained on the CGRN (Collection of Greek Ritual Norms) inscription dataset.

The goal of the research was to improve lemmatization accuracy for Ancient Greek inscriptions, which pose unique challenges due to irregular spelling, accent variation, inconsistent capitalization, and nonstandard morphology not seen in literary Greek.

To address this, I implemented character encoding-aware normalization, capitalization-sensitive preprocessing, logic to handle missing predictions, and conducted controlled experiments comparing different edit-distance and string-similarity approaches. Through systematic evaluation, the final model achieved a 1.61% absolute increase in accuracy and reached a peak accuracy of 86.7%, resulting in a more reliable lemmatization tool for computational philology and epigraphic research.



# I: Enhancing AGILE with NLP

---

---

## ENHANCING AGILE WITH NATURAL LANGUAGE PROCESSING TECHNIQUES FOR ANCIENT GREEK INSCRIPTION LEMMATIZATION

---

**Kavin Ramadoss**  
Sunset High School

**Michael Wei**  
Ward Melville High School

**Ari Philip**  
Monte Vista High School

**ABSTRACT:** Pre-existing ancient Greek lemmatizers lack one crucial feature: the handling of inscriptions. To fill this gap, we present an improved AGILE, a lemmatizer for academics across the globe to translate Greek inscriptions accurately. We implement various techniques such as Jaro-Winkler distance, normalization processes, and the modification of the edit distance calculation to account for accents and capitalization. When these tactics are applied in testing, there is a notable 1.61 percent increase in accuracy. The improvement of AGILE not only strengthens its position as the superior model for the lemmatization of ancient Greek inscriptions but also makes a significant contribution to the field of computational philology and linguistic analysis. 24 trials definitively confirm the improved accuracy of AGILE over both the pre-existing model and all other competing inscription lemmatizers.

**KEYWORDS:** AGILE (Ancient Greek Inscription Greek Lemmatizer), Jaro-Winkler

### 1 Introduction

In the constantly evolving, ChatGPT-led present day, where Large Language Models (LLMs) dominate the current landscape of machine learning, lemmatization is irreplaceable. With it, artificial intelligence can understand the context of and accurately respond to the millions of queries people ask daily. Due to its importance, lemmatization has been swiftly implemented into Natural Language Processing (NLP) models in most languages. The same is the case for Ancient Greek. Historians still uncover artifacts from thousands of years ago, but deciphering the ancient Greek written on these materials is often tedious. There have been great strides in ancient Greek lemmatization, but the vast majority of these models have been trained with data from literary text written in ancient Greek manuscripts. Therefore, there remains a gaping hole in the scholarly world for the lemmatization of Ancient Greek inscriptions. AGILE, a new ancient Greek inscription lemmatizer with an accuracy of 85.09 percent, intends to fill this gap[1]. It is the most effective Ancient Greek inscription lemmatizer by a margin of 20 percent. A 'clean' dataset of literary texts is inadequate for inscriptions as the texts inscribed on materials such as ceramic and stone vary significantly from the dialects and morphology of the ancient Greek manuscripts. Hence, the dataset most appropriate for this project's purposes was *CGRN: A Collection of Greek Ritual Norms*[3]. Our goal was to improve the pre-existing model of AGILE to increase its efficiency and accuracy, determining that AGILE had four significant issues.

- The edit distance failed at multiple points, returning lexicon words unrelated to AGILE's predicted word. (*Edit distance*: how the model matches the predicted word with the most similar word from the lexicon through a numerical distance calculation.)
- Similar to the English language, proper nouns in Ancient Greek are capitalized. AGILE could not make an association between the lowercase and uppercase characters, often causing lexicon words to have incorrectly increased edit distances.
- In certain situations, AGILE fails to offer any prediction at all. In this case, nothing but a blank underscore is printed.
- Like capitalization, AGILE cannot recognize the similarity between two characters that only differ by accents. Consequently, likely words have a greater edit distance and are therefore less considerable.

## II: AR Surgical Navigation

---

I developed an augmented reality surgical navigation system for endoscopic endonasal skull base surgeries that integrates real-time anatomy visualization and distance feedback using the Microsoft HoloLens 2, with the goal of reducing the risk of carotid artery injury during minimally invasive neurosurgical procedures.

The project combined medical imaging, computer vision, and deep learning, including the design and training of a 3D U-Net–based convolutional neural network to segment carotid arteries from magnetic resonance angiography (MRA) scans from the OASIS-3 dataset, achieving 91% accuracy

I implemented a full AR pipeline in Unity and C# using the Mixed Reality Toolkit, enabling spatial registration of patient-specific vascular models, real-time hand and instrument tracking, and continuous distance calculations between surgical tools and the carotid arteries with visual safety warnings. The system overlays holographic reconstructions of critical vasculature directly into the surgeon's field of view, supports hands-free interaction via gestures and voice commands, and maintains real-time performance on untethered hardware.

Quantitative evaluation demonstrated improvements in estimated surgical precision, safety, efficiency, and cost compared to traditional optical and electromagnetic navigation systems, highlighting the potential of combining AI-driven anatomical modeling with augmented reality to enhance surgical decision-making and reduce operative risk.

**Awards Received:** National Junior Science and Humanities Symposium (NJSHS) Finalist; Regeneron Biomedical Science Award (\$500);

# II: AR Surgical Navigation

## ARNES: A novel machine learning based navigation system for endoscopic endonasal surgeries

By Kavin Ramadoss

### INTRODUCTION: BACKGROUND

- Endoscopic endonasal (EE) skull base surgeries aim to remove brain / pituitary tumors and other lesions from the skull base.
- Tumor is very close to the carotid arteries.
- Surgeons reach the brain through the nasal cavity and sinuses using an endoscope.
- An average of 1000 EE endonasal surgeries are performed each year.
- Electromagnetic (EM) system emits a magnetic field.
- The registration process creates a map of various points.
- Surgical tools to obtain location of instrument relative to the patient pre-op images.
- Distance to carotid artery unavailable.

### METHODOLOGY

#### PHASE 1: CAROTID ARTERY MODELING

#### PHASE 2: IMPLEMENTATION

**RB-CNN custom architecture**

**Hypiperparameter Tuning**

**Loss Function**

**Dataset**

**Feature Extraction from 2D Images**

**Distance Calculation**

### DATASET

The Open Access Series of Imaging Studies (OASIS) project was used for this project.

- OASIS provides access to large MRI datasets.
- The OASIS-3 dataset was used for this project and spans over 1,379 patients aged 42 to 95.
- A total of 2,842 MRI sessions.
- The OASIS-3 dataset contains various imaging modalities such as T1 and T2.
- 4 different MRI machines were used.
- The scans provided high-resolution structural brain imaging data with 1mm isotropic voxels and 256x256 acquisition matrices.

### PROPOSED SOLUTION

**Phase one of ARNES:** These work in addition with the RB-CNN to make ARNES.

**Phase two of ARNES:** This phase uses the RB-CNN output vs original MRI images to calculate the distance between the surgeon's hand and the carotid artery vertices.

**Measures:** Measures the distance between the surgeon's hand and the carotid artery vertices.

**Immersive and aids decision making key anatomical distances in 3D.**

### CONCLUSION

- The convolutional neural network (CNN) demonstrated very good results achieving a 95% accuracy.
- The introduction of the surgical navigation system resulted in significant improvements across all evaluation criteria, surpassing existing benchmarks and setting new standards for accuracy and reliability.
- The cost has reduced from 600K to 25K which is a very significant reduction.
- This means that potentially the cost that the patient also has to pay could be reduced significantly as the surgical equipment doesn't cost as much as before.

### FUTURE WORK

- Even though this navigation system has shown great promise some things have to be reworked for the surgical setting.
- Factors like dynamic distortion and line-of-sight obstruction may complicate augmented overlays in surgical settings.
- In the future, once minor tweaks are made to the navigation system, it could surpass current surgical navigation systems.
- So this AR surgical navigation system can be used worldwide.

### RESULTS

**98.3%** Safety with ARNES

**93%** Precision with ARNES

**2 hrs** Time with ARNES

**10K** Cost with ARNES

CITERIA	USING CURRENT TECHNOLOGY	USING AR NAVIGATION SYSTEM
TIME	Around 3 hours	Around 2 hours
SAFETY	95.2%	95.7% (3.1% relative increase)
PRECISION	73%	20%
COST	300-600K	98% decrease

Table comparing criterias currently vs with ARNES

### APPLICATIONS

- AR surgical navigation systems hold significant potential not only in neuroscience but also across various surgical disciplines.
- The method of distance calculation integrated with the surgical view can be used for a range of surgical procedures beyond endoscopic endonasal surgeries.
- For instance, in surgeries near vital organs like the heart, the navigation system can be used to measure distances to the heart.
- Once adjustments are made to tailor the navigation system for different surgical contexts, it could be implemented universally across diverse surgical specialties.
- This represents a shift in surgical navigation, potentially leading to substantial reductions in mortality rates and improved patient outcomes on a global scale.
- By providing surgeons with real-time spatial awareness and precise guidance, AR surgical navigation systems have the potential to minimize surgical complications and enhance patient safety.
- The widespread adoption of such technology has the capacity to positively impact patient care worldwide, saving lives and improving overall surgical outcomes.

### KEY REFERENCES

Beadle, E., Schroeder, J., Morris, B., & S. Sorenson. (2008). An overview of Remote Entry and some potential applications. <https://doi.org/10.1109/leisec.2008.5074715>

Children's Cerebellopharyngeal Tethering (PDB92). (2023, October 5). National Cancer Institute. Retrieved November 1, 2023, from <https://www.cancer.gov/cancertopics/pdq/cancer/PDB92>

Choudry, F. A., Grantham, J. T., Rai, A. T., & Hodge, J. P. (2015). Vascular geometry of the intracranial carotid arteries: A study of length, diameter, and tortuosity. *Journal of Neurosurgical Surgery*, 85, 52–58. <https://doi.org/10.1136/neurospur-2015-011671>

Find smth to add here

# II: AR Surgical Navigation

---

---

## AUGMENTED REALITY SURGICAL NAVIGATION SYSTEM WITH REAL-TIME CAROTID ARTERY DISTANCE CALCULATION AND TISSUE VISUALIZATION IN ENDOSCOPIC ENDONASAL SURGERIES USING THE MICROSOFT HOLOLENS

---

Kavin Ramadoss  
Sunset High School

**ABSTRACT:** The intricate nature of endoscopic endonasal skull base surgeries, where surgeons reach the brain through the nasal cavity and sinuses, necessitates precise navigation to avoid inadvertent encounters with the delicate carotid arteries, which can precipitate severe neurological consequences. This is why an innovative system that combines imaging modalities with a distance calculation algorithm is being built to solve this critical problem. The program enables real-time visualization of the surgical field, allowing surgeons to have sight of the patient's critical tissues and organs through a camera embedded in their instruments, coupled with a continuous assessment of the range and separation between the surgeon's instruments and the carotid arteries. The interface is designed to issue immediate warnings when the calculated distance approaches or breaches a predetermined safety threshold, thus equipping surgeons with invaluable, instantaneous feedback. The surgeons wear the Microsoft Hololens, an augmented reality headset that allows them to see holograms displaying the distance and their surroundings. Furthermore, a custom convolutional neural network (CNN) was developed to extract intraluminal carotid artery models reliably from standard preoperative scans. By mimicking human visual processing, the algorithm achieves segmentation accuracy at around 91%. Patient-specific models then become projected holographically onto the operative scene through the Microsoft Hololens. This comprehensive program enables tremendous change in the medical field, causing surgeries to be accomplished quicker and significantly improving the mortality rate.

**KEYWORDS:** Endoscopic endonasal skull base surgeries; Imaging modalities, Convolutional Neural Network(CNN); Carotid arteries; Microsoft Hololens

### 1 Introduction

Surgical navigation systems aim to provide real-time tracking and visualization to assist surgeons in safely navigating complex anatomy during minimally invasive procedures. While external optical and electromagnetic tracking systems have been utilized, they require line-of-sight and suffer from occlusion and accuracy degradation [19]. Augmented reality (AR) presents a promising alternative by integrating visualization directly within the surgeon's field of view. Recent advances in AR headsets enable inside-out tracking without external hardware, as well as seamless blending of virtual imagery with the real environment. This research explores building an AR-based surgical navigation system using the Microsoft HoloLens 2, a mixed-reality headset, and Unity, a software platform with 3D capabilities. The system relies on registration through fiducial markers placed around the skull, which tells the surgeon the exact size of the carotid arteries. The surgeons can then scale down the carotid arteries hologram on the Microsoft Hololens to align with the patient's carotid arteries. Crucial structures at risk, such as the carotid arteries segmented from magnetic resonance images, are visualized directly within the surgeon's view, along with metrics like proximity between the surgeon's instrument and the carotid arteries. The ability to display patient-specific anatomical reconstructions and metrics directly in real-time could improve precision, speed, and spatial awareness during surgery.

A head-mounted AR navigation system also leaves the surgeon's hands-free and does not require constant attention shifts between a static monitor and the operating field. Furthermore, a self-contained headset solution increases

# III: Novel Malaria Inhibitors

---

I independently designed and led this research project, with mentorship from a faculty advisor, to address whether an integrated, low-cost computational pipeline could rapidly identify novel inhibitors of *Plasmodium falciparum*.

The goal was to accelerate antimalarial drug discovery by systematically filtering massive chemical libraries for compounds with strong binding, drug-like properties, and low toxicity.

I developed AutoFilter, a multi-stage framework combining rule-based chemical filtration (Lipinski, Veber, PAINS), high-throughput molecular docking (AutoDock Vina), ADME analysis (SwissADME), machine-learning–based toxicity and synthetic accessibility prediction (eToxPred), and molecular dynamics simulations (GROMACS).

Applying AutoFilter to the 2.4 million–compound ChEMBL database against *P. falciparum* apicoplast DNA polymerase, I reduced the search space to five high-confidence inhibitor candidates.

These compounds exhibited strong binding affinities, favorable pharmacokinetics, low predicted toxicity, and stable protein–ligand interactions, with preliminary *in vitro* assays showing one compound achieving 50% inhibition.

This work demonstrates AutoFilter as a scalable, generalizable platform that can significantly reduce the time and cost of early-stage drug discovery.

**Awards Received:** Gloria Barron Young Hero; National Junior Science and Humanities Symposium (NJSHS) Finalist; 3rd place regional JSHS finalist (\$1500 scholarship); Biomedical Engineering Society (BMES) Conference Attendee; Johns Hopkins University (JHU) Global Health Leadership Conference Presenter (available on YouTube); Awarded \$2000 scholarship to Oregon State University; 2nd place in Biochemistry category at regional science fair (BHSE); 2nd place in Microbiology category at state science fair (NWSE);

# III: Novel Malaria Inhibitors

**Autofilter: A Low-Cost Biocomputational Framework For High-Throughput Screening Of Chemical Databases**

**INTRODUCTION**

**Malaria**

Malaria is a disease caused by parasites of the genus *Plasmodium* and is transported through the bites of mosquitoes. Malaria is the third deadliest disease with a death rate of over 17% and it had over 249M cases in 2018. 94% of cases are seen in the African region. CDC confirms that *P. falciparum* has also developed resistance to ALL of the drugs currently targeting the disease.

Chemical databases contain chemical information about bioactive drugs. However effective methods of screening these large chemical databases (n = 2.4 million) haven't been discovered.

Lack of efficient drugs + High drug-resistance → No Malaria drugs → 600K annual deaths from malaria

**Current Drug Discovery**

- Scientists collect biological samples from natural environments.
- Solvents separate compounds based on solubility.
- Chromatography techniques purify natural compounds for testing.
- Analytical instruments determine molecular structures of active compounds.
- Disease models are prepared using cells or organisms.
- Compounds are tested directly on biological systems.
- Microscopy captures cellular responses to test compounds.
- Validate promising compounds through repeated experiments.
- Target identification reveals how active compounds affect cells.

**Limitations**

- Takes around 10 years to synthesize one drug.
- Over 1 billion to test one drug through various clinical trials.
- 90% of drugs fail in clinical trials.
- Many drugs lack usefulness over time due to drug resistance.
- Plasmodium* has a complex life cycle.

**RESEARCH OBJECTIVES**

**PROBLEM STATEMENT** ~ Current drug discovery methods take many years and are very costly due to inefficient methods to synthesize various drugs. Drug resistance also causes drugs to lose their usefulness over time. Inefficiency methods coupled with drug resistance leads to more deaths worldwide.

**CLAIM** ~ This research project will develop a cost-effective framework to screen large chemical databases for better drug discovery and develop malaria drugs through the use of the framework.

**Engineering Goals**

**Engineering Goal #1:** To build the first ever biocomputational framework (AutoFilter) for the high-throughput screening of potential malaria inhibitors.

**Engineering Goal #2:** Make AutoFilter cost-effective, high-throughput, and timely to address the biggest issues of current drug discovery: high cost and time.

**Engineering Goal #3:** Through the use of AutoFilter, develop 5 potential malaria inhibitors to help millions of individuals suffering from malaria on a global scale.

**PROPOSED SOLUTION**

The flowchart illustrates the integrated workflow for drug discovery:

- User enters input via UI.
- Input goes to Database Preparation (CHEMBL).
- Database Preparation feeds into Protein Preparation and Molecular Docking.
- Protein Preparation feeds into ADMET Filtration, Toxicity and SA Filtration, and Molecular Dynamics.
- Molecular Docking feeds into ADMET Filtration, Toxicity and SA Filtration, and Molecular Dynamics.
- ADMET Filtration, Toxicity and SA Filtration, and Molecular Dynamics all feed into DNA Polymerization.
- DNA Polymerization feeds into Gel Electrophoresis.
- Gel Electrophoresis feeds into Measure Cytotoxicity & IC50.
- Measure Cytotoxicity & IC50 feeds back into Molecular Docking.
- Measure Cytotoxicity & IC50 also feeds into a final step labeled "Results".

**METHODOLOGY**

**Database Preparation**

Step 1: Lipinski's Rule of 5, Veber's Rule, PMNS Rule, Curated Dataset. Step 2: Ro5 used to remove unstable compounds, Veber's rule used to remove non-bioavailable compounds. Step 3: Removes compounds likely to produce false positives (n = 2.2 M). Step 4: Curated database ready for screening phases (n = 2.2 M).

**Protein Preparation**

- Hydrogen Atoms: Hydrogen atoms are added to ensure protein structure.
- Kollman Charges: Add Kollman charges for electrostatic interactions.
- Prepared Protein: Prepared protein ready for Molecular Docking.

**Molecular Docking**

Technique that predicts how well small molecules will bind to target receptor proteins. Consists of a scoring function and a search algorithm.

**Scoring Algorithm**

- Docking uses the Gradient Optimization Conformational Search Algorithm.
- Scoring function:  $A = \text{value} \sum_{i=1}^n \left( \frac{\partial}{\partial x_i} E_i \right)^2$
- Score each pose and then optimizes to find the best pose of the compound.

**Molecular Docking Process**

- Compute conformer to Molecular String.
- Protein and prepare compound.
- Receptor active site, grid box are set to 10 Å.
- Dock compound in novel binding sites.
- Save 1 ligand conformer in PDBQT format.
- Active Site: X: -93.1, Y: -34.2, Z: 112.4.
- 2.2 million potential drugs.
- 3200 potential drugs.
- Negative scores indicate high binding affinity.

**RESULTS**

**Toxicity and SA Filtration**

Selected compounds exhibit low toxicity and high SA scores.

- Toxicity (0-1 scale) with 1 as most toxic and 0 as least toxic.
- SA (0-1) with 0 as most easily disrupted.
- The lowest SA score was roughly 0.18 (L4).

**Molecular Dynamics**

A RMSD below 1.5 is indicative of ligand stability and interaction strength.

A RMSD below 1.5 indicates that the protein is very stable over various residues (atoms).

Each compound (L1-L5) good binding and interaction strength.

The protein is stable allowing for better binding with selected compounds.

**in vitro Validation**

Testing is complete for 1 ligand, with 4 more currently underway. Preliminary testing confirms that the 1st ligand is a working inhibitor and inhibits malaria at low concentrations.

Further metrics are being tested like IC50 and cytotoxicity.

**CONCLUSION**

In this research, I was able to create the first-ever biocomputational framework for screening large chemical databases.

**Revisiting Engineering Goals**

- The framework uses a novel integrated screening solution to achieve high throughput. No equivalent commercial solution exists currently.
- The framework is highly cost-effective because of its *in silico* nature. It is predicted to reduce drug discovery cost by around 75% or more.
- Revisiting Engineering Goal #3, AutoFilter is able to screen an entire chemical database (n = 2.4 million compounds) and narrow it down to five inhibitors within two days. It is predicted to reduce the total time for drug discovery by around 75%.

**Impacts**

Discovered a new malaria inhibitor that inhibits the parasite by at least 50%

- Discovered a new novel allosteric binding site to prevent malaria from developing resistance to drugs.
- AutoFilter is a universal framework that can be applied to all diseases given a protein target and its active site such as HIV, cancer, TB, and more.

**APPLICATIONS**

- Use the framework with other diseases like cancer, HIV, Tuberculosis (TB) for more cures on a global scale.
- Since framework is usable with other chemical databases such as ZINC250, I want to expand my drug discovery with a larger chemical space of more than 2.4 million through the custom UI.
- Adapt framework to be used for pesticide discovery to potentially discover a new pesticide.

**FUTURE WORK**

I am very happy with the progress I made in this project and I want to increase its strength in the future to help even more people!

- Test 4 other drugs for their efficacy against malaria.
- Pursue next stages of testing for working inhibitor including *in vivo* testing and clinical trials.
- If the results are viable, I will take steps to publish a drug for malaria that can be used by people suffering from malaria worldwide.
- Publish the research in a high-end journal and take steps to make the framework available for commercial and public use.

**REFERENCES**

Gershon, Samuel, and Ulf Ryde. "The MM/PBSA and MM/CSA Methods to Estimate Ligand-Binding Affinities." Expert Opinion on Drug Discovery, vol. 10, no. 5, 2015, pp. 449–461. <https://doi.org/10.1517/17460441.2015.103936>.

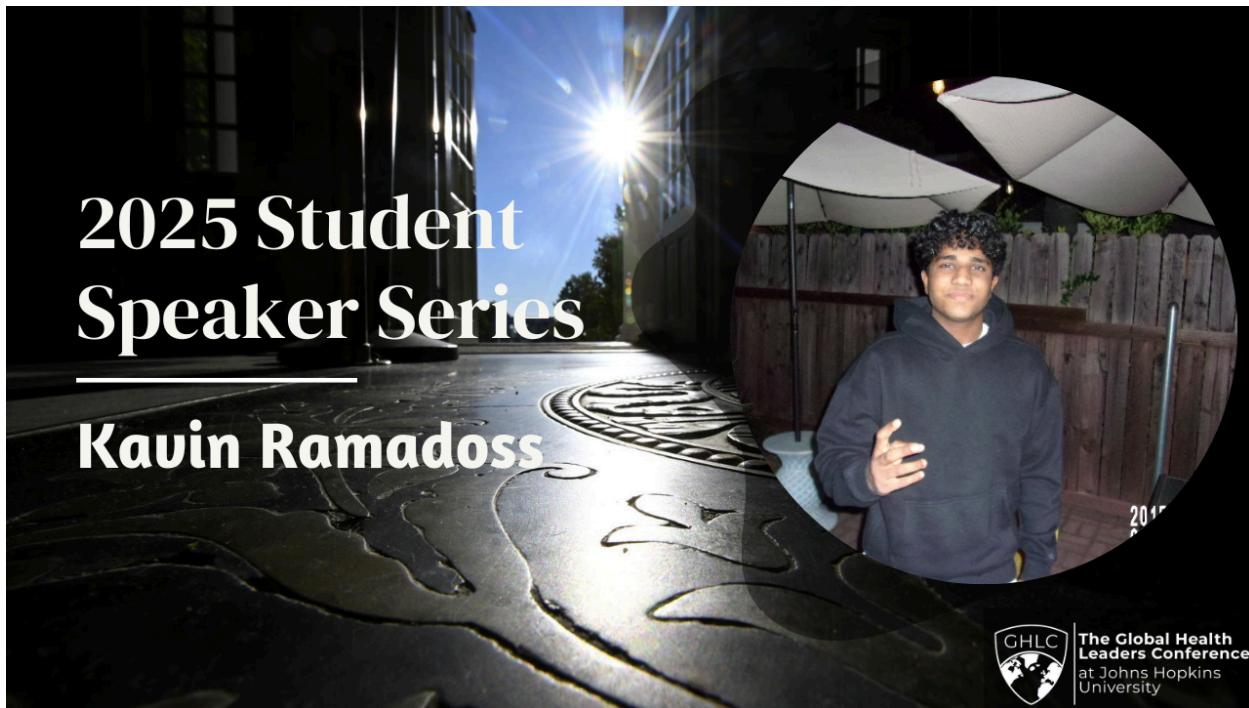
Zhou, Jun, and Ming Tang. "Zinc250: A Large-scale Library of Small Organic Molecules." *Journal of Computational Chemistry*, vol. 32, no. 11, 2011, pp. 2359–2368. <https://doi.org/10.1002/jcc.21816>.

Deale, Andrew, et al. "Autofilter: A Biocomputational Framework Web Tool to Evaluate Pharmacokinetics, Drug-Likeness and Medicinal Chemistry Friendliness of Small Molecules." *Scientific Reports*, vol. 7, no. 1, 2017, p. 42717. <https://doi.org/10.1038/srep42717>. Full list is in my Research Page.

Kavin Ramadoss, Sunset High School

# III: Novel Malaria Inhibitors

---



Kavin Ramadoss, Sunset High School

# III: Novel Malaria Inhibitors

---

---

## AUTOFILTER: A LOW-COST BIOCOMPUTATIONAL FRAMEWORK FOR HIGH-THROUGHPUT SCREENING OF CHEMICAL DATABASES AND IDENTIFICATION OF NOVEL MALARIA INHIBITORS TARGETING PLASMODIUM FALCIPARUM

---

Kavin Ramadoss  
Sunset High School

**ABSTRACT:** Malaria is the third deadliest disease, with approximately 249 million cases annually, particularly in tropical regions. Caused by Plasmodium parasites transmitted through the bite of Anopheles mosquitoes, malaria remains a significant global health burden and is increasingly difficult to treat due to rising drug resistance. Drug discovery for malaria is both costly and time-consuming, typically requiring over a decade and around \$3 billion before a compound gains approval. To address this challenge, AutoFilter was developed: a low-cost and novel biocomputational framework that integrates machine learning (ML) and screening tools to streamline the filtering of large chemical databases for more efficient drug discovery. AutoFilter sequentially screens compounds that violate basic chemical filters such as Lipinski's Rule of 5 (Lipinski et al., 1997), Veber's (Veber et al., 2002), and PAINS; performs molecular docking and analyzes post-docking interactions; conducts ADME filtration to identify compounds with favorable drug-like properties; employs an ML model to predict toxicity and synthetic accessibility; and finally applies molecular dynamics (MD) simulations to refine compound stability. AutoFilter was applied to screen the ChEMBL database, which contains 2.4 million bioactive compounds, to identify malaria inhibitors targeting Plasmodium falciparum apPOL. The five selected compounds demonstrated high inhibition performance and favorable drug-like properties and are currently undergoing *in vitro* trials. As the first integrated biocomputational framework for chemical database screening, AutoFilter is a transformative tool for drug discovery across diverse diseases, efficiently identifying inhibitors while reducing costs and time by 50%, with the profound potential to save lives worldwide.

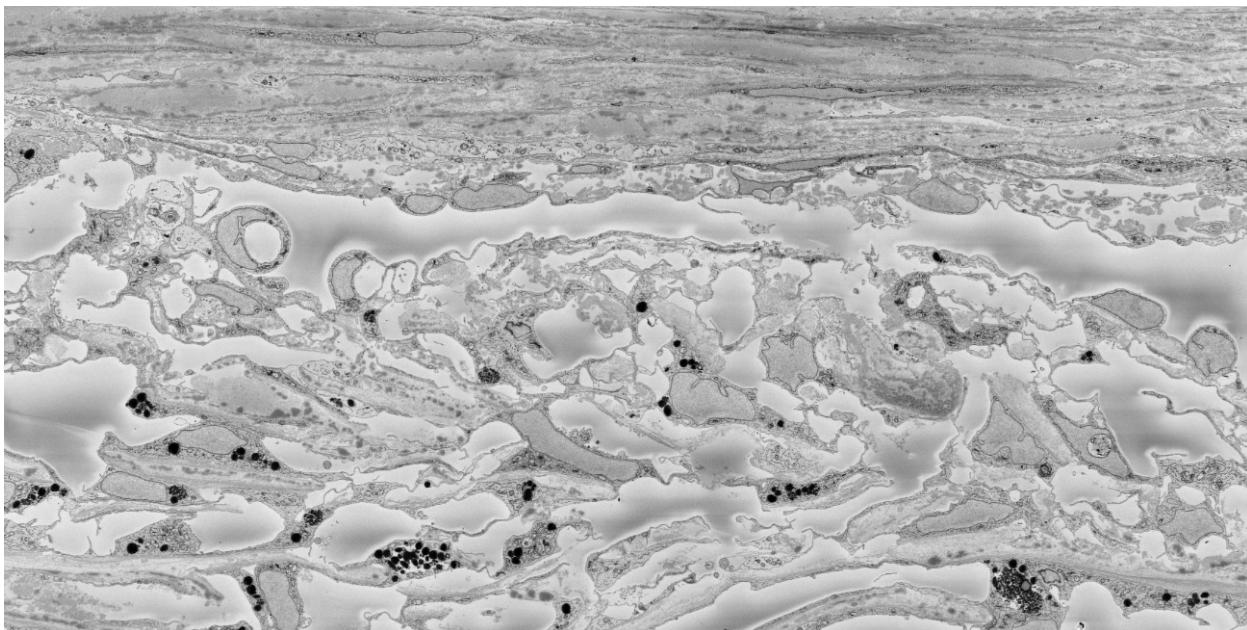
# IV: Collagen Segmentation

---

I conducted a research project in conjunction with the Oregon Health and Science University (OHSU) focused on collagen segmentation in glaucoma-affected patients, with the goal of determining whether automated image analysis could reliably identify and quantify structural collagen changes associated with disease progression.

Using labeled ophthalmic imaging data, I designed, trained, and evaluated convolutional neural network-based segmentation models to distinguish collagen fibers from surrounding tissue with high spatial accuracy.

My workflow included image preprocessing, data augmentation, model training and validation, and performance evaluation using standard segmentation metrics. The resulting model successfully captured collagen organization patterns that are difficult to assess manually, demonstrating the feasibility of deep learning as a tool for objective, high-throughput analysis of glaucomatous tissue.



# V: Alzheimer's Binding Site

---

This ongoing research at Caltech investigates how small-molecule compounds may interact with and disrupt amyloid-beta 42 (A $\beta$ 42) aggregation, a central pathological feature of Alzheimer's disease.

Building on prior evidence that certain drugs can chemically promote amyloid disaggregation, the project focuses on identifying and characterizing potential binding regions on A $\beta$ 42.

A computational pipeline combining blind molecular docking and molecular dynamics simulations is used to map favorable ligand–peptide interaction sites across the A $\beta$ 42 surface and to evaluate the stability and behavior of these complexes over time.

Structural fluctuations, interaction energies, and hydrogen-bonding patterns are analyzed to identify key residues involved in ligand stabilization and potential disruption of  $\beta$ -sheet assembly. Together, this research investigates the molecular mechanisms underlying chemical-driven amyloid clearance.

## **Computational Identification of Small-Molecule Binding Sites on Amyloid- $\beta$ 42 for Alzheimer's Disease Therapeutics**

Kavin Ramadoss, Shuwen Lei, William A. Goddard III

### **Abstract:**

Alzheimer's disease (AD) is characterized by the aggregation of amyloid-beta (A $\beta$ ) peptides, particularly A $\beta$ 42, into toxic oligomers and fibrils that disrupt neuronal function. Recent studies have highlighted the therapeutic potential of small molecules capable of disaggregating A $\beta$  aggregates, including compounds such as EPPS, Quinacrine, Rhizolutin, Borrelidin, Necrostatin-1, and others that promote chemical-driven amyloid clearance. Building on this foundation, our study aims to identify potential binding sites within A $\beta$ 42 for nine candidate small-molecule drugs previously reported to induce amyloid disaggregation. Blind molecular docking was employed to comprehensively scan the A $\beta$ 42 surface and predict favorable binding regions for each ligand, with docking energies used to prioritize potential interaction hotspots. Subsequent molecular dynamics simulations, performed using GROMACS, were used to refine binding poses and evaluate complex stability through energy minimization, RMSD analysis, and hydrogen bonding dynamics. This combined computational approach provides insights into atomical features and how these small molecules interact with A $\beta$ 42, revealing critical residues involved in ligand stabilization and potential disruption of  $\beta$ -sheet aggregation. These results advance our understanding of chemical-driven amyloid clearance mechanisms and support the design of novel therapeutics targeting A $\beta$  aggregation in Alzheimer's disease.

# VI: Alzheimer's Binding Site

---