

## Opening for an undergraduate research scholar in the Ghose Laboratory ([www.ghoselab.org](http://www.ghoselab.org))

### DESCRIPTION

The Ghose laboratory is looking for a skilled programmer with knowledge of common coding paradigms to perform analysis of the structural properties of biomolecules using statistical and/or machine learning approaches. The ideal candidate will possess a strong mathematics background and have knowledge of at least one of the commonly used programming languages (i.e. Python, R, C++, etc.). No prior knowledge of chemistry or biology is necessary since the necessary skills will be acquired during the course of the project. Junior or Seniors are preferred since it is expected that the proposed would become their senior project. Ambitious undergraduates who wish to pursue future careers in Computational Biology/Chemistry/Physics or Applied Mathematics are preferred. The student will be expected to work under close supervision of an advanced PhD student.

### REQUIREMENTS

- The student should be proficient in at least one coding language (Python, R or C++ preferred).
- The student should have a good working knowledge of basic statistics, vector analysis and linear algebra.
- The student must be comfortable with working in a Linux environment (e.g. Ubuntu or OSX).
- Any knowledge of machine learning approaches to data mining is a bonus but not necessary. It is expected that the student will become proficient in these approaches during the course of the project.
- While there are no specific GPA requirements for this position, it is expected that the student will have meet the GPA requirements for most nationally competitive fellowships.

### RESPONSIBILITIES

- The student will be expected to be present in the laboratory on a regular basis. The student will be provided with the necessary computational resources including hardware and software.
- The student will be expected to write code that reads large sets of protein structures and extracts specific structural/geometric properties.
- The student will generate and annotate tables of the above properties.
- The student will perform appropriate encoding of these properties for input into unsupervised machine learning classifier.
- The student will write code that uses any of the popular machine learning libraries for either R, C++ or Python to perform classification of this data.
- As part of their training the student will be expected to write periodic progress reports as the project advances.
- As part of their training the student will be expected to present their results in local and national meetings.

Interested students should contact Ronnie Ghose (Professor of Chemistry & Biochemistry) by email at [rghose@ccny.cuny.edu](mailto:rghose@ccny.cuny.edu) or [rghose@ghoselab.org](mailto:rghose@ghoselab.org) attaching a copy of their transcript and curriculum vitae. This position is available immediately.