Summer Internship in Computational Chemistry Programming at Ipsen Bioscience, Cambridge MA


Technical & Soft-Skill Requirements:

Student studying computer science/engineering at a Boston-area university. The intern should have experience and a strong interest in scientific programming and machine learning. S/he will apply the programming skillset to real problems in computational drug design.

Skill Requirements:

- Experience with Python required (R, Scipy, Numpy also preferred)
- Experience training machine learning models (SVM, Neural nets, clustering, regression, etc.) in courses or projects
- Experience validating models (cross-validation, ROC analysis, etc.)
- Experience with building web interfaces and database management strongly preferred

Project Outline:

The modeling team at Ipsen is developing computational tools to facilitate in silico drug design. An important part of that process is the curation of large databases of information relevant to peptide medicinal chemistry and using machine learning to convert this data into predictive models for peptide design. Determining the best-performing kernel functions, choosing the right features/descriptors for the model, and developing the appropriate training/test sets for cross validation are critical areas of focus.

WHAT THE INTERN WILL DO

1. Learn the composition and property attributes of our internal peptide and amino acid datasets
2. Use machine learning to train models for peptide design and validate the predictive accuracy of these models on experimental test sets
3. Build a user interface that allows access to the model
4. Write software/scripts to search public databases (PDB, NCBI, etc.) to cross-reference PPI structures with more extensive human sequence paralogs.
5. Write software/scripts to process *in silico* PPI structures to search for specific peptide scaffolds that mediate the interface interaction.
6. Benchmark the accuracy of the method to predict existing PPI-mediating peptide scaffolds.
7. Build a graphical user interface (GUI) to access and search the PEPTIDE-PPI database.

Ipsen will offer a young, talented software developer the unique opportunity to learn how her/his skills can be used to advance drug design efforts within the context of a pharmaceutical company. Knowing how to code versus knowing what *needs* to be coded is a critical part of transitioning from an academic environment to industry. The job will involve answering real questions and building tools that benefit early stage drug design. It will involve collaboration with the molecular modeling team, biologists, bioinformaticians, and other stakeholders to help achieve project goals.