CHEM GA 2671 - DIVE IN TO COMPUTATIONAL PHYSICAL CHEMISTRY
FALL 2024

Class Time/Room: Thursday, 2PM-4:30 PM. Computer lab, 541 LaGuardia Place

Instructor:	Glen Hocky	hockyg@nyu.edu
	Nicole Smina	ns3539@nyu.edu
Office Hours:	Glen Hocky	TBD, Silver 1018
	Nicole Smina	TBD, Silver 1018

Office hours start the second week of classes.

Course Overview: The goal of this class is to prepare you to use computational tools for chemistry research. It will also lay a foundation for data science and scientific computing more generally. A short lecture portion will introduce concepts behind what we are doing, but the primary emphasis of the class will be on *doing*. The goal by the end of the class is that you can jump into using new tools with ease and confidence.

This is a 4-credit course intended for first or second-year graduate students or advanced undergraduates. There are no pre-requisite courses, but knowledge of undergraduate physical chemistry is expected.

Specific skills developed in the class: Some skills you will definitely come away with in this class are: (a) use of command line tools and general understanding of Linux-based architectures, plus running and submitting jobs on high performance computing resources (b) software version control with git and GitHub, (c) AI guided programming with Copilot + ChatGPT/Bing Search etc, (d) python and jupyter notebooks for data analysis, including use of numpy, matplotlib, pandas, scipy, scikitlearn, (e) molecular dynamics simulations of proteins (with gromacs or amber), (f) molecular dynamics simulations for materials (with LAMMPS or HOOMD-Blue), (g) analyzing MD simulation results using python libraries (mdtraj or MDAnalysis), (h) enhanced sampling simulations with PLUMED, (i) quantum calculations in python, using wavefunction or DFT (with PySCF or Psi4), (j) visualizing molecular structures (primarily with VMD).

Course communication: Like in a research lab, we will use **slack** for communication. You will be invited to the slack specific to the class. *Please install the slack application on your computer*.

Slack has different "channels" for different topics. Please post your questions and discussions in public channels so that everyone can work together.

Course websites: Course material and assignments will be available on from the group *github* website, and will be announced in slack. The course github is: https://github.com/hockyg/comp-lab-class-2024 and the course website is: https://hockygroup.hosting.nyu.edu/teaching/comp.html

Textbook: Useful reference textbooks for this course are Introduction to Computational Physical Chemistry by Joshua Schrier and Statistical Mechanics, Theory and Molecular Simulation, 2nd edition, by Mark Tuckerman. All required material will be provided by the instructor and available online.

Lectures: The first 20-30 minutes of each section will be used to review key concepts, as well as to go over past hands-on exercises in more detail.

Office hours: There will be regularly scheduled office hours if you wish to discuss topics with me in person. You will also have access to me and all of your classmates' knowledge through slack.

Purpose of office hours. What is the point of office hours? Many people never attend office hours, or they do so only right before or right after an exam. But office hours can be so much more than for emergencies!

This is a great chance for us to get to know each-other better. It is also a good time to clear up confusions you (or we) have about the material. We never want people to feel like they are behind. It's also a great time to discuss how the course material relates to your other interests, or to discuss more advance topics. So please consider attending as many office hours as you can from the beginning!

Course assignments: The only assignments will be completion of each set of hands-on exercises. For each one, you will push a git repository of a certain format that includes a jupyter notebook or other documents explaining your results and perhaps answering various questions. *Git commits should be done with the command line tool to prevent issues later, and you should practice this early and often at the beginning.*

Class Attendance: This is a hand-on class, even though it is on the computer. Attendance is mandatory unless you get specific permission. Each week you will be paired with another class member (or two) to work with, which is another reason that attendance is crucial. *You should stay the entire time and work on your labs.*

Respect and inclusion: Another goal of this course is to create a learning environment that is inclusive and fosters contributions from all students. No one in the course should be made to feel uncomfortable because of the identity or background. If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me.

Grading: Each module will be equally weighted. It is expected that you will complete the entire module with high quality figures and explanations. Each module *must be turned in before class before 1 week later*. You will then have *1 week to revise* after receiving feedback. The exception is the final assignment which must be completed by the end of finals week.

An unexcused absence will lose half-credit on that module and so two unexcused absences will likely not result in a passing grade.

Note on keeping up: This is a cumulative course. The reason we are asking each lab to be turned in the subsequent week is so that you continually practice all the required skills, and by the end of the semester basic skills will be much easier!

Academic Integrity and Plagiarism: As you know, we take academic honesty very seriously at NYU. The instructors for this course have no tolerance for plagiarism or cheating. The NYU policy on plagiarism will be enforced. Students who fail to conform to NYU's standards on academic integrity will be subject to stringent disciplinary action. Inform yourself in advance of proper academic conduct. In brief (and quoting from the College of Arts & Science policy), "Academic honesty means that the work you submit – in whatever form – is original." When in doubt, ask. Please consult: http://cas.nyu.edu/page/academicintegrity

Since there are no exams, and you will be working together and using resources from the internet, I do not expect this to be an issue. Still, you should turn in your own work.

Disability Disclosure Statement: Academic accommodations are available for students with disabilities. Please contact the Moses Center for Students with Disabilities (212-998-4980 or mosescsd@nyu.edu) for further information. Students who are requesting academic accommodations are advised to reach out to the Moses Center as early as possible in the semester for assistance.

Wellness exchange: Additional resources related to mental health and wellness are available from the Wellness Exchange. Their 24 hour hotline is (212) 443-9999. Information can be found at https://www.nyu.edu/students/health-and-wellness/wellness-exchange.html

Modules overview

Week 1 (Sept 5) – What is computational chemistry. How do we use computers to do chemistry? Use of the bash, command line text editing, ssh/ssh keys, GitHub, VMD. ChatGPT etc

Week 2 (Sept 12) – Data analysis and visualization, data organization. Visualizing real molecular dynamics simulation data.

Week 3 (Sept 19) – What is a molecular dynamics simulation? Setting up and running your own MD simulation using GROMACS.

Week 4 (Sept 26) – More complex MD setups. Using CHARMM-GUI to build complex inputs; use these data to study diffusion of ions. [Probable: lecture for next lab]

**Week 5 (Oct 3) – Enhanced sampling molecular dynamics simulations, and parallel tempering. Perform and analyze parallel tempering in GROMACS. [Holiday]

Week 6 (Oct 10) – Enhanced sampling by Metadynamics, and the PLUMED library for enhanced sampling. Perform Metadynamics simulations in GROMACS+PLUMED.

Week 7 (Oct 17) – Protein structure prediction and Alphafold. Use Alphafold to predict the structures of proteins only from sequence, and analyze the results.

Week 8 (Oct 24) – Molecular dynamics simulations for materials. Studying crystallization using LAMMPS.

Week 9 (Oct 31) – More advanced usage of LAMMPS. Using LAMMPS to study supercooled liquids.

Week 10 (Nov 7) – Introduction to electronic structure. Usage of the Psi4 quantum chemistry package to compute the interaction energy of an ideal gas dimer, and fit to Lennard-Jones interaction.

Week 11 (Nov 14) / Week 12 (Nov 21) – Introduction to RDKit and its usage for medicinal chemistry. Computing energy of small molecules by combining RDKit, SMILES, and Psi4.

Nov 28 – No class, Thanksgiving!

Week 12 (Dec 5) – Ab initio molecular dynamics simulations. Simulating water and an excess proton in CP2K.

Week 13 (Dec 12) – Mixed quantum classical simulations. Combining QM and MM into QM/MM to study an enzyme in CP2K.

Finals week – Wrap up and submit all assignments/projects