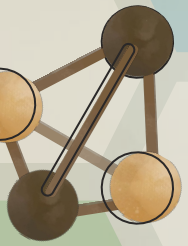
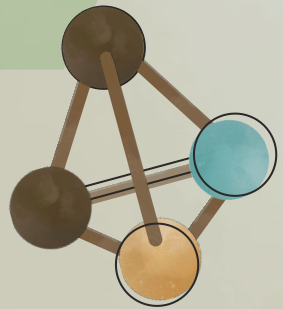




SIMCODES Midterm Presentation

Avery Martin

Acknowledgements

Funding from SIMCODES and the NSF

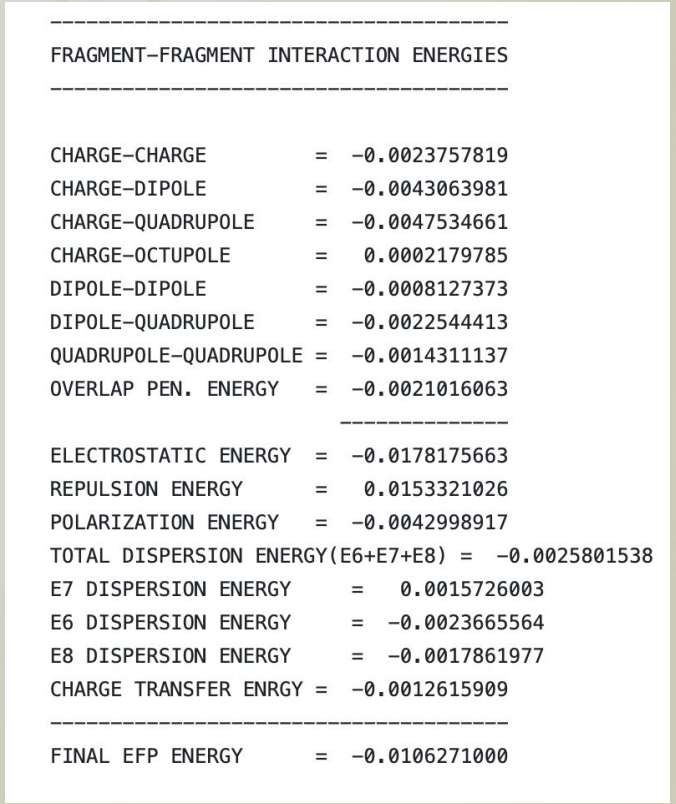
My Project

Title: Machine Learning-Enhanced Computational Modeling of Metal-Protein Interactions

Description: Focuses on enhancing computational models of metal-protein interactions by integrating quantum chemistry, fragmentation methods, and machine learning. We aim to validate a fragmentation-based strategy, develop machine learning algorithms to predict metal binding in biological systems, and create scalable workflows for future therapeutic screenings.

My Role in the Project

I am the chemistry student in my group. I mainly focus on generating data that the computer scientists use for their work. I do this by utilizing GAMESS and EFP to process systems of large molecules, which calculates the interaction energies between the fragments of molecules to find the total energy.



```
-------------------------------------
FRAGMENT-FRAGMENT INTERACTION ENERGIES
-------------------------------------

CHARGE-CHARGE          =  -0.0023757819
CHARGE-DIPOLE          =  -0.0043063981
CHARGE-QUADRUPOLE      =  -0.0047534661
CHARGE-OCTUPOLE        =   0.0002179785
DIPOLE-DIPOLE          =  -0.0008127373
DIPOLE-QUADRUPOLE      =  -0.0022544413
QUADRUPOLE-QUADRUPOLE  =  -0.0014311137
OVERLAP PEN. ENERGY    =  -0.0021016063
                          --------------
ELECTROSTATIC ENERGY   =  -0.0178175663
REPULSION ENERGY       =   0.0153321026
POLARIZATION ENERGY    =  -0.0042998917
TOTAL DISPERSION ENERGY(E6+E7+E8) =  -0.0025801538
E7 DISPERSION ENERGY      =    0.0015726003
E6 DISPERSION ENERGY      =  -0.0023665564
E8 DISPERSION ENERGY      =  -0.0017861977
CHARGE TRANSFER ENRGY  =  -0.0012615909
-------------------------------------
FINAL EFP ENERGY       =  -0.0106271000
```

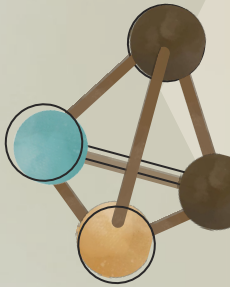
What I've Learned So Far:

Git/GitHub

I've been utilizing Git and GitHub throughout the entire process. I had never used it before.

GAMESS

I use GAMESS throughout the process of calculating energies.

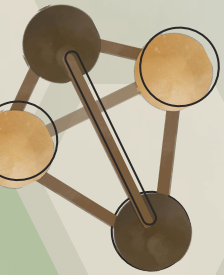


EFP Method

I use EFP to calculate the energies of each fragment in a molecule.

Terminal

I use a Linux terminal everyday. I've learned basic commands and usage.



Progress

GAMESS and EFP

I've ran many databases through GAMESS using EFP. This data has been used by the computer science students to make a neural network.

```
---------------------------------------
FRAGMENT-FRAGMENT INTERACTION ENERGIES
---------------------------------------

CHARGE-CHARGE          =  -0.0023757819
CHARGE-DIPOLE          =  -0.0043063981
CHARGE-QUADRUPOLE      =  -0.0047534661
CHARGE-OCTUPOLE        =   0.0002179785
DIPOLE-DIPOLE          =  -0.0008127373
DIPOLE-QUADRUPOLE      =  -0.0022544413
QUADRUPOLE-QUADRUPOLE  =  -0.0014311137
OVERLAP PEN. ENERGY    =  -0.0021016063
```

Amino Acid Optimization

I've created a database of coordinates for every amino acid and all of their conformers. I also optimized these conformers. These will be used for metal-protein calculations.

```
-----------------------------------------------
-------   RESULTS OF PCM CALCULATION   -------
-----------------------------------------------

FREE ENERGY IN SOLVENT = <PSI| H(0)+V/2 |PSI>       =     -322.9374758713 A.U.
INTERNAL ENERGY IN SOLVENT = <PSI| H(0) |PSI>       =     -322.9219772680 A.U.
DELTA INTERNAL ENERGY =  <D-PSI| H(0) |D-PSI>       =        0.0000000000 A.U.
ELECTROSTATIC INTERACTION                           =       -0.0154986033 A.U.
PIEROTTI CAVITATION ENERGY                          =        0.0000000000 A.U.
DISPERSION FREE ENERGY                              =        0.0000000000 A.U.
REPULSION FREE ENERGY                               =        0.0000000000 A.U.
TOTAL INTERACTION (DELTA + ES + CAV + DISP + REP)   =       -0.0154986033 A.U.
TOTAL FREE ENERGY IN SOLVENT                        =     -322.9374758713 A.U.

FREE ENERGY IN SOLVENT       =     -202646.35 KCAL/MOL
INTERNAL ENERGY IN SOLVENT   =     -202636.62 KCAL/MOL
DELTA INTERNAL ENERGY        =          0.00 KCAL/MOL
ELECTROSTATIC INTERACTION    =         -9.73 KCAL/MOL
PIEROTTI CAVITATION ENERGY   =          0.00 KCAL/MOL
DISPERSION FREE ENERGY       =          0.00 KCAL/MOL
REPULSION FREE ENERGY        =          0.00 KCAL/MOL
TOTAL INTERACTION            =         -9.73 KCAL/MOL
TOTAL FREE ENERGY IN SOLVENT =     -202646.35 KCAL/MOL
```

Challenges

- What I've found most challenging so far is the more detailed computer related tasks. I've come across many different issues and error messages that I've needed help debugging.
- Finding a general understanding of the computer science and math focused aspects of my project has also been very difficult. I have no background in comp sci and my only math experience is Calculus I.

What I Aim To Accomplish

Data Generation

I want to develop a lot more data for the computer science side so that they can have more training data to refine their ML model.

Research Skills

I hope to learn more research-related skills; specifically, how to prepare for a symposium, how to make a poster, how to write an abstract, and how to present at a symposium.