Introduction

In recent years, environmental consciousness has been of utmost importance in various disciplines, including chemistry, where the wasteful practice of using hazardous solvents persists. Although a class of environmentally friendly solvents, ionic liquids, have been discovered, properties of these substances are not widely understood. By running ab initio Molecular Dynamics (AIMD) simulations, the energetic properties of these liquids can be understood; however, these simulations are computationally expensive in terms of time and processing power. Using the concept of machine learning, Artificial Neural Networks (ANNs) can be developed to minimize these costs while maintaining a sufficient level of accuracy.

What is Machine Learning?

In broad terms, machine learning is a division of artificial intelligence that relies on the principle that "systems can learn from data, identify patterns, and make decisions with minimal human intervention." ANNs serve as a major component of machine learning and consist of a series of layers each containing nodes that take inputs from previous layers and output results to subsequent layers, as shown in Figure 1. These structures improve upon themselves by changing the weights of certain nodes in the overall computations during each iteration until they show a sufficient level of accuracy as compared to a training set of data.

What is an Ionic Liquid?

An ionic liquid is an ionic compound, which occurs as a result of the attraction of opposite charges between a cation and an anion, that exists as a liquid at a wide range of temperatures, including room temperature. A typical characteristic of these substances is a relatively large, organic cation and a small anion, which causes the molecule to be asymmetric and thus unable to crystallize at normal temperatures. Ionic liquids are considered environmentally friendly as they are non-volatile compounds and can be “designed and tuned to optimize yield (and) selectivity."1

Methodology

The following procedure was used to run AIMD simulations and train ANNs for the two substances studied, water and BMIM-Cl:

1. The chemical drawing software MarvinSketch® was used to develop a representation of the substance to be studied.
2. The molecule was then used to generate an input file for a solvent box (consisting, separately, of 32 water molecules and 1 molecule of BMIM-Cl) using the Packmol® software, which creates initial coordinates for AIMD simulations.
3. Separate simulations were run using CP2K®, an AIMD simulation program, to find the cutoff radius for each atom regarding electron interactions with other atoms to be used in AIMD simulations.
4. Geometry optimization simulations were then run in CP2K® to find the lowest energy state of the solvent box and the final coordinates of the simulations, then used as the starting coordinates of the equilibration step, which finds a stable position for the system given the temperature and pressure constraints: 298 K and 1 atm.
5. After the equilibration, AIMD simulations were run using CP2K® to generate coordinates and energy for each iteration. A portion of the input files were then used to train the ANN using AENET®, a software package that allows for the construction of ANNs related to atomic and molecular structures. This information is displayed in Figure 4. At the time of writing, the ANN for the experiment of the 32 water molecules has not completed an iteration of its training process, meaning that data regarding its training progress percentage and the number of test cases passed could not be collected at this time.

Discussion and Future Directions

Although, due to time constraints, the ANNs for each experiment were not able to be trained completely, the data collected in this work thus far shows the potential that these constructs have in understanding the energetic properties of ionic liquids. The AIMD simulations, though consistently accurate as they use quantum mechanical calculations to determine the energy of a given orientation of molecules, were run for approximately a week, indicating an excessive amount of computational resources, whereas the data collected from the training of the ANNs was performed after a single day. Although the accuracy of the ANNs may seem astounding, the amount of time that would be saved by using these instead of running simulations is significant and should be further studied. Future research on this subject should involve running separate AIMD simulations, extracting the coordinates, and comparing the obtained energies to those predicted by a fully trained ANN using the same set of coordinates. By doing so, the advantages of using ANNs to understand the properties of ionic liquids instead of only running simulations can be better understood.

Conclusion

• Ionic liquids serve as an environmentally friendly alternative to commonly used solvents for organic reactions.
• Several chemical properties of ionic liquids, including energetic properties, are currently not understood.
• The widely used method of running AIMD simulations to understand the energetics of these substances, though highly accurate, is computationally expensive.
• Using the concept of machine learning, ANNs can be trained to predict the energies associated with molecular orientations in a fraction of the time used with AIMD simulations.
• Although the accuracy of the ANNs was shown to be only 60.7% in this work, they show great potential to serve as viable alternatives to running costly simulations by limiting computational expenses, which will allow research regarding ionic liquids to be conducted at a faster rate.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1659144.

Citations