

Density Functional Theory and Information Theory Based Indices as Tools to Investigate the Reactivity of Chemical Systems and Their Applications

Meressa Abrha Welearegay

Supervisor: Prof. dr. hab. Andrzej Holas
Subsidiary supervisor: Dr. Robert Balawender

Abstract of PhD Thesis

A general theme of this dissertation is showing usefulness of the density functional theory (DFT) and the information theory (IT) based indices as a source and carrier of the information about molecular structure and reactivity.

One of the main targets undertaken in this work is opening the gate to the exploration of the chemical space through alchemical derivatives. This is done by developing the methodological framework of alchemical derivatives and their chemical reactivity based indices. The crucial conditions for the qualitative and quantitative accuracies of the alchemical predictions are recognized. As illustrative transmutations showing the potential of the method in exploring chemical space, some examples of increasing complexity are presented, starting with the deprotonation, continuing with the transmutation of the nitrogen molecule, and ending with the substitution of isoelectronic (B,N) units for (C,C) units and N units for C-H units in carbocyclic systems. The overall trends observed for the alchemical deprotonation energy prove the usefulness of the alchemical indices as the probe in the chemical reactivity investigations. The results of calculations for the BN derivatives of benzene and pyrene show that this method has great potential for efficient and accurate scanning of chemical space. The tentative results for the carcinogenic activity of the polycyclic aromatic hydrocarbons (PAHs) reinforce this opinion.

The information theory based methodological framework necessary for extracting useful information for atomic and molecular systems is deeply examined. The concepts of the transferability and additivity of atoms or functional groups is used as a test in extensive and detailed analyses of the electronic density and the shape function as a functional argument for the IT based measures. It is also shown that in the shape representation, the observed trends for the IT measures and complexities are in contradiction with chemical intuition.

It is important to point out that linear correlations are obtained between the kinetic energy and the Fisher information and Onicescu information measures as well as between the atomization energy and the atomization entropy. The analysis of the IT based measures of information planes shows that the Shannon-Fisher information plane provides "richer" information about the pattern, organization, similarity of molecules than the Shannon-Onicescu and Fisher-Onicescu planes. The final conclusion for this part is that, the IT measures can be used in the chemical reactivity investigation as the source of the information about the pattern, organization, similarity of molecules, while not as direct indicators of their reactivity.

Finally, a support vector machine based models of classification and regression are developed for the carcinogenic effect of PAHs. The accuracy of 93% of correct classification is achieved using selected structural and molecular descriptors. The correlation coefficient for the predicted versus experimental index of carcinogenicity is 0.9475.

The used set of molecule is large enough (in total around 1000 molecules) and diverse to improve the previous understating of subject undertaken in this thesis and to generalize obtained conclusions.