University of Toronto

Woodsworth College

Science Abroad

(CHM396y0)

13th Annual European Molecular Computational Summer School for Life Science Students

University of Miskolc,

Miskolc, Hungary, EU

May 12 – July 1

2018

RESEARCH PROJECTS

The CHM396Y0 course within the Summer Abroad Program offers the possibility to conduct research in different aspects of modern chemistry. Students who register will learn how to deal with quantum chemical problems, and will work on a project under one of the proposed topics.

In this course, students can expect to:

1) Learn and apply the basics of computational chemistry
2) Learn how to perform literature searches on their molecular system
3) Evaluate new scientific results
4) Prepare a poster for an international scientific conference
5) Write a report on the results they achieved in the hopes of publishing a scientific paper based on their results
6) Plan for further advanced study

Participating Lecturers and Assistants

Béla Viskolcz	University of Miskolc (Hungary)
Imre G. Csizmadia	University of Miskolc (Hungary)
	University of Toronto (Canada)
Anita Rágyanszki	University of Miskolc (Hungary)
	University of Toronto (Canada)
Milán Szőri	University of Miskolc (Hungary)
Béla Fiser	University of Miskolc (Hungary)
	Ferenc Rákóczi II. Transcarpathian Hungarian Institute (Ukraine)
John Justine Villar	University of the Philippines Diliman (Philippines)
Imre Jákli	MTA-ELTE Protein Modeling Research Group (Hungary)
Min-Yen Lu	University of Miskolc (Hungary)
Zsófia Borbála Rózsa	University of Miskolc (Hungary)
Andrea Guljas	University of Toronto (Canada)
Bálint Kiss	University of Miskolc (Hungary)
Attila Szabó	University of Miskolc (Hungary)

Green Chemistry

Imre G. Csizmadia
e-mail: icsizmad@hotmail.com

While the current chemical industry is producing many million tons of all types of carbon compounds from pharmaceuticals to plastics it is also producing a huge amount of waste. The chemical reagents used in the reactions end up in poisoning the waters and soil and the needed thermal energy, coming from combustion, is poisoning the atmosphere.

If the reactions of organic chemistry could be converted to redox reactions, then new electrochemistry could replace the existing obsolete thermal reactions used in industrial manufacturing. In this green chemistry instead of reagents catalysts, deposited on the electrodes, and instead of thermal energy electrical energy is used.

PROJECT EXAMPLES

$Ph\text{-}H + H\text{-}OH \rightarrow Ph\text{-}OH + 2H^{(+)} + 2e^{(-)}$

$Ph\text{-}H + H\text{-}NH_2 \rightarrow Ph\text{-}NH_2 + 2H^{(+)} + 2e^{(-)}$

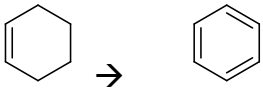
$Ph\text{-}H + H\text{-}CH{=}CH_2 \rightarrow Ph\text{-}CH{=}CH_2 + 2H^{(+)} + 2e^{(-)}$

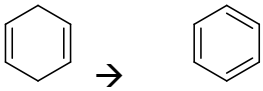
$2CH_4 \rightarrow H_3C\text{-}CH_3 + 2H^{(+)} + 2e^{(-)} \rightarrow H_2C{=}CH_2 + 4H^{(+)} + 4e^{(-)} \rightarrow HCCH + 6H^{(+)} + 6e^{(-)}$

$CH_4 + NH_3 \rightarrow H_3C\text{-}NH_2 + 2H^{(+)} + 2e^{(-)} \rightarrow H_2C{=}NH + 4H^{(+)} + 4e^{(-)} \rightarrow HCN + 6H^{(+)} + 6e^{(-)}$

$2\ H_2C{=}CH_2 \rightarrow H_2C{=}CH\text{-}HC{=}CH_2 + 2H^{(+)} + 2e^{(-)}$

$H_2C{=}CH\text{-}HC{=}CH_2 \rightarrow H\text{-}CC\text{-}CCH + 2H^{(+)} + 2e^{(-)}$

$H_2C{=}CH\text{-}HC{=}CH_2 + H_2C{=}CH_2 \rightarrow$  $+ 4H^{(+)} + 4e^{(-}$

$H_2C{=}CH\text{-}HC{=}CH_2 + H\text{-}CC\text{-}H \rightarrow$  $+ 2H^{(+)} + 2e^{(-)}$

Each of the above anodic oxidations may be coupled by the following cathodic reduction:

$CO_2 + 6H^{(+)} + 6e^{(-)} \rightarrow HCOOH + 4H^{(+)} + 4e^{(-)} \rightarrow H_2C{=}O + H_2O + 2H^{(+)} + 2e^{(-)} \rightarrow H_3C\text{-}OH + H_2O$

Computer Aided Biofuel Design: Systematic Theoretical Investigation for High Energy Molecules

Milán Szőri
e-mail: milan.szori@uni-miskolc.hu

The need for more environmentally friendly energy sources and the limited fossil resources encourage current research to convert biomass into renewable products. Next-generation production of biofuels from lignocellulosic biomass is a topic of intense research and a range of oxygenates is being proposed as future fuels or components for fuel blends [1,2]. Ethers and esters are such alternative biofuels or fuel additives [3].

Figure 1 *Optimization layers of fuel components [5].*

Oxygenates derived through the selective catalytic refunctionalization of carbohydrates of lignocellulosic biomass can be tailored to exhibit desired physico-chemical fuel properties that unlock the full potential of advanced internal combustion engines [4,5]. Considering the molecular structure of a fuel component as adjustable parameter, it is possible to find the most promising molecular entities, if computational property prediction is employed to virtually screen the generated structures with regard to key physico-chemical fuel properties.

In the current proposal, following stoichiometry are selected for such investigation providing individual project for students. All the possible constitutional isomers of the selected molecular formula will be generated. Then, the relevant physical-chemical properties of these isomers, such as heat of formation and heat of combustion, will be computed. The students involved in this project can contribute with their findings to the deeper understanding of biofuel and combustion.

References

1. Janssen, A. J.; Kremer, F.W.; Baron, J. H.; Muether, M.; Pischinger, S.; Klankermayer, J. *Energy Fuels* **2011**, 25 4734–4744.
2. Kohse-Höinghaus, K.; Oßwald, P.; Cool, T. A.; Kasper, T.; Hansen, N.; Qi, F.; Westbrook, C. K.; Westmoreland, P. R. Biofuel Combustion Chemistry: From Ethanol to Biodiesel. *Angew. Chem., Int. Ed.* **2010**, 49, 3572-3597.
3. Bansch, C.; Kiecherer, J.; Szőri, M.; Olzmann, M. Reaction of Dimethyl Ether with Hydroxyl Radicals: Kinetic Isotope Effect and Prereactive Complex Formation *J. Phys. Chem. A* **2013**, 117, 8343-8351.
4. Dahmen, M.; Marquardt, W. Model-Based Design of Tailor-Made Biofuels, *Energy Fuels*, **2016**, 30, 1109–1134.
5. Di Li, Xinyu Li, and Jinlong Gong Catalytic Reforming of Oxygenates: State of the Art and Future Prospects, *Chem. Rev.*, **2016**, 116, 11529–11653

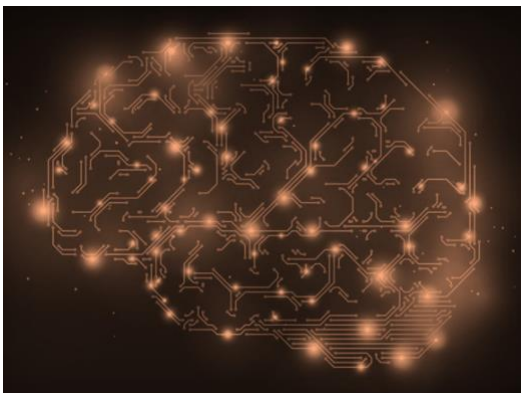
Artificial Neural Networks as a tool to describe conformational changes

Anita Rágyanszki
e-mail: a.ragyanszki@utoronto.ca

"Far out in the uncharted backwaters of the unfashionable end of the Western Spiral arm of the Galaxy lies a small unregarded yellow sun. Orbiting this at a distance of roughly ninety-eight million miles is an utterly insignificant little blue-green planet whose ape-descended life forms are so amazingly primitive that they still think digital watches are a pretty neat idea."

Douglas Adams, Hitchhiker's Guide to the Galaxy

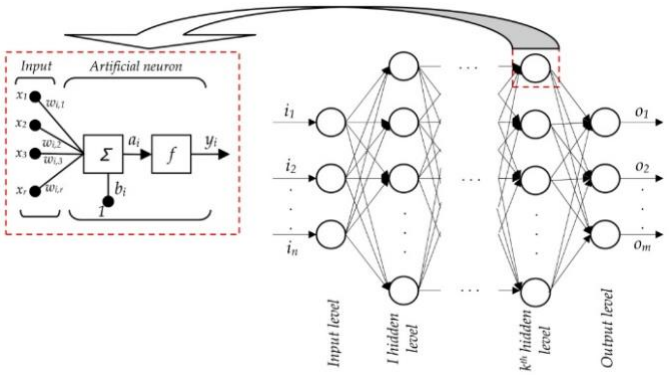
The progress of scientific discovery has accelerated in the last decades of the 20th century, especially with the development of third generation computers. This technological advance presents the opportunity to develop artificial intelligence or machine learning, which includes the creation of artificial neural networks. This development is largely based on numerical simulations that have been created to model and study natural phenomena.



Atoms in nature can be described in part by their electronic structure, which refers to the arrangement and energy of electrons that exist in specialized orbitals surrounding a positively-charged nucleus. By applying electronic structure calculations, we can describe atomic and molecular properties and reaction mechanisms. These computations form the basis of the calculation of Potential Energy Surfaces (PES), where the potential energy of a molecule is modelled as a function of the positions of atoms. PES contain information about the molecular structure, thermodynamic properties, and the chemical reactivities of the molecules they describe. Although all chemical changes may be described by a reaction PES, conformational PES represent a special case, since they do not require calculations involving the breaking and forming of chemical bonds. Instead, they suggest a practical starting point for introduction of mathematical minimization techniques in predicting chemical properties [1]. Based on these approaches, it is possible to develop new mathematical models that may be used both to represent conformation changes in molecules and to build conformational networks. Using different mathematical and computational methods, we can describe the folding pathways of, for instance, a peptide or a protein, or find the folded structure of molecules in a vacuum and in different type of solvents. The current focus of our interests is the development of Artificial Neural Networks (ANNs) to simplify the prediction protein folding pathways.

ANNs are computing systems that consist of many simple processors that transmit information to each other through their many interconnections. An ANN model attempts to use similar "organizational" principles that are believed to be used in the human brain; their interconnections parallel the complex network of neuronal

pathways that are formed by brain cells, and depend on the specific connections between processors similarly to how that the brain depends on synaptic connections for memory and learning. An important application of ANNs is in machine learning to perform complex tasks.



In our research, we attempt to use ANNs to discover and understand relationships between molecular geometry and the forces acting on atoms. Based on the work of Behler *et al.* as well as other teams [3], we plan to develop new geometry descriptor variables for input into an ANN, and eventually implement a more effective method for optimizing ANN. We will then use the ANN to compute accurate forces at a small computational cost. This will allow us to simulate conformational changes and state transitions in polymers and clusters [4]. Overall, the goal of our project during the summer school is to use newly developed mathematical models to describe the conformation changes of peptide molecules, the simplest units of fully-formed proteins, and to understand conformational networks using mathematical and computational methods.

References

[1] J. J. S. Villar, A. R. L. Valdez, D. H. Setiadi, I. G. Csizmadia, B. Fiser, B. Viskolcz, A. Rágyanszki, Can. J. Chem. 95 (2017). 830-836.

[2] A. Rágyanszki, K. Z. Gerlei, A. Suranyi, A. Kelemen, S. J. K. Jensen, I. G. Csizmadia, B. Viskolcz, Chem. Phys. Lett. 625 (2015) 91-97.

[3] J Behler, M Parrinello, Phys. Rev. Lett. 98 (2007) 1–4.

[4] V Botu, R Batra, J Chapman, R Ramprasad, J. Phys. Chem. C 121 (2017) 511–522.

MOLECULAR DESIGN AND SYNTHESIS OF GREEN POLYURETHANES

Béla Fiser, Min-Yen Lu
fiser@uni-miskolc.hu , emilie.lu@mail.utoronto.ca

Turning traditional design to green design is important to the sustainability of industry. Many processes and material selection start to consider for the long-term and short-term environmental impact [1]. Polyurethanes (PUs) are polymers made by reacting diisocyanates with polyols to form various versatile materials such as heat insulators in construction, seat cushion in automotive and elastomeric materials in medical industries [2].

Conventional PUs are made from petrochemical based starting materials which raised severe health, environmental concerns. Thus, intensive research and development has been carried out to prepare polyurethanes based on natural polyols that can be sustainable for production. Natural polyols are abundant and renewable. In particular, the substitution of petro-based polyol with carbohydrate polyols have shown to improve biodegradability and mechanical properties [3]. Carbohydrates such as sugars, oligo- and polysaccharides can successfully replace polyether polyols in the bio-PU foam systems [4].

This work aims to perform a mechanistic study on polyurethane systems by using computational chemical tools. It explores the reactivity of natural polyols towards conventional isocyanates. In the end, the most promising molecules will be tested experimentally. The students involved in this project can contribute with their findings to the molecular design of environmentally friendly polyurethanes. During the course they will gain hands-on experience in the computational study and synthesis of green polyurethanes.

References

1. U. M. Diwekar and Y. N. Shastri, "Green process design, green energy, and sustainability: A systems analysis perspective," *Comput. Chem. Eng.*, vol. 34, pp. 1348–1355, 2010.

2. B. Nohra, L. Candy, J. F. Blanco, C. Guerin, Y. Raoul, and Z. Mouloungui, "From petrochemical polyurethanes to biobased polyhydroxyurethanes," *Macromolecules*, vol. 46, pp. 3771–3792, 2013.

3. J. David, L. Vojtová, K. Bednařík, J. Kučerík, M. Vávrová, and J. Jančář, "Development of novel environmental friendly polyurethane foams," *Environ. Chem. Lett.*, vol. 8, pp. 381–385, 2010.

4. M-Y. Lu, "Molecular Design of Sugar-Based Polyurethanes", Master Thesis, University of Miskolc, 2017.

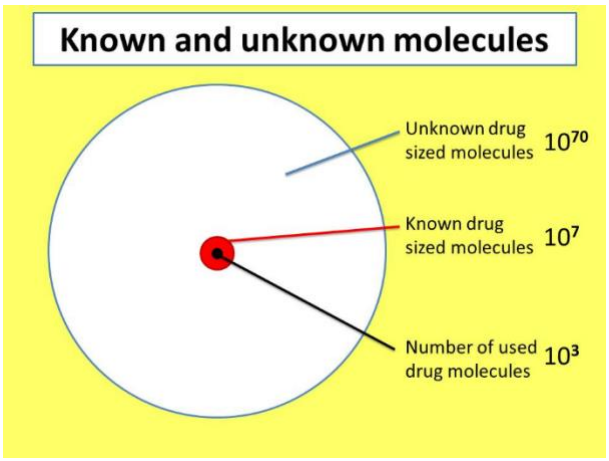
A Prelude to "Modular" Protein Folding:
Building Models for Understanding Peptide Conformational Transformations

John Justine Villar
e-mail: jjvillar@uni-miskolc.hu

"Problems are best solved not on the level where they appear to occur but on the next level above them…. Problems are best solved by transcending them and looking at them from a higher viewpoint. At the higher level, the problems automatically resolve themselves because of that shift in point of view, or one might see there was no problem at all."

-David R. Hawkins

Proteins are the machines and building blocks of living cells. These are predominantly assembled from 20 amino acids, and are used for structural support, storage, transport of other substances, signaling from one part of the organism to another, movement, and defense against foreign substances, among others. There are huge numbers of different proteins, with each one performing its specific task. After the successful deciphering of the genetic code that defines how the amino acid sequences of proteins are coded in the DNA, one of the major missing steps in understanding the chemical basis of life is the *protein folding problem* – the task of understanding and predicting how the information coded in the amino acid sequence of proteins at the time of their formation translates into the three-dimensional structure of the biologically active protein. Knowledge of how such a protein would fold would allow one to predict its chemical and biological properties. This would help the researchers understand the mechanism of hereditary and infectious diseases, aid in designing drugs with specific therapeutic properties, and of growing biological polymers with specific material properties. Misfolded peptides are the causes of multiple neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, amyloidosis, and ALS [1]. Furthermore, there is up to 70 orders of magnitude, i.e., 10^{70}, more drug-sized molecules that are still unexplored, that may have a potential to provide cure to the said disorders.



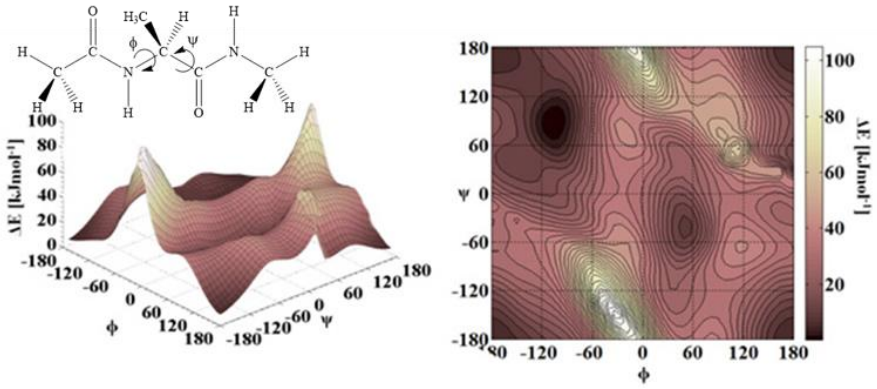
Different structures of the same molecule have distinct physicochemical properties. One of the most explicit a posteriori properties of the structure of a molecule is its energy. The energy of a molecule is dependent on parameters such as its environment and conformation.

One aspect of dissecting this problem is then to observe the energy profile of a protein of interest, as the potential energy of a foldamer allows us to determine the relative stability of each possible conformation.

In principle, finding the stable foldamers of a protein requires an efficient sampling of the entire conformational space of the protein, to which there is an associated *potential energy surface* (PES). A local minimum of the associated PES may correspond to the energy of native fold.

Building the conformational potential energy surface (PES) of a molecule are important because they aid us in visualizing and understanding the relationship between potential energy and molecular geometry, and in understanding how prediction methods locate and characterize structures of interest. However, the time and space complexity of electronic structure calculations, commonly used to generate PES, increases exponentially with an increasing number of atoms [2,3,4,5].

The goal of this study is to understand the topology and dynamics of protein folding through constructing models of conformational transformations in small peptides that precisely mimic the associated potential energy surfaces [2,3,4].



This will provide valuable insights on the mechanisms of protein folding, such as minima and transition states, while saving computational time and resources that is necessary with current popular methods.

References:

[1] Reynaud, E. Protein Misfolding and Degenerative Diseases. *Nature Education* 3 (2010) 28.

[2] J. J. S. Villar, A. R. L. Valdez, D. H. Setiadi, I. G. Csizmadia, B. Fiser, B. Viskolcz, A. Rágyanszki. Dimension Reduction in Protein Conformational Analysis: A Two-Rotor Mathematical Model of Amino Acid Diamide Conformational Potential Energy Surfaces. *Can. J. Chem.* 95 (2017). 830-836.

[3] J. J. S. Villar, A. R. L. Valdez, D. H. Setiadi, I. G. Csizmadia, B. Viskolcz, A. Rágyanszki. An Improved Two-Rotor Function for Conformational Potential Energy Surfaces of 20 Amino Acid Diamides. *Can. J. Chem* (2017). In press.

[4] A. Rágyanszki, K. Z. Gerlei, A. Suranyi, A. Kelemen, S. J. K. Jensen, I. G. Csizmadia, B. Viskolcz. Big data reduction by fitting mathematical functions: A search for appropriate functions to fit ramachandran surfaces. *Chem. Phys. Lett.* 625 (2015) 91-97.

[5] M. R. Peterson, I. G. Csizmadia. Analysis of the topological features of the conformational hypersurface of n-butane, *J. Am. Chem. Soc.* 100 (1978) 6911-6916.