

Analytical functional representation of quantum chemical potential energy curves and surfaces

Theses of the Ph.D. Dissertation

Anita Rágyanszki

Supervisor:

Dr. Béla Viskolcz

*Doctoral School of Chemistry
Faculty of Science and Informatics
University of Szeged*

Szeged, Hungary

2015



1. Short summary of the antecedents

The protein folding problem is one of the 125 largest unsolved problem in science in addition to “what is the Universe made of” or “are we alone in the Universe”. There were several attempts to solve this century old unanswered question. To solve this problem we need many hundreds of supercomputer years.

One of the reasons that protein folding has not yet been solved is due to the big data associated with it. Reducing this data set is a possible approach towards simplification. In order to make these simplifications, we first have to understand very deeply mathematical properties of conformational spaces.

The Potential Energy Surfaces (PES) as a function of two independent variables and they are playing an important role in chemistry. They are containing information about the molecular structure, thermodynamic properties and chemical reactivities. To predicting the total Potential Energy Hypersurface, we must to describe two subspaces, the reaction subspace and the conformational subspace. Although all chemical changes maybe described by a PES, nevertheless conformational PES are representing a special case which is the subject of the present dissertation.

Conformational analysis of organic molecules has been initiated before the middle of the 20th century. In 1963 Ramachandran attempted to perform conformational analysis on simple peptides and the pattern for such a double rotor –CONH-CH₂-CONH- turned out to be quite different from the propane surface. More recently, it was observed that even as a simple compound as n-pentane behaved in non-ideal fashion. Consequently, it now appears to be reasonable to assume that surface topology is a function of the rotating moieties.

Fitting mathematical functions to potential energy curves was investigated, in the 1970s, by Pople and co-workers. Their original driving force appeared to be that such functions might be useful to estimate the rate of conformational changes. Pople and his co-workers attempted to use very simple trigonometric functions to achieve a reasonable accuracy. Later, in 1988, Chung has pointed out that full Fourier expansion is needed. In the meantime Peterson et al. started to use an extended Fourier series in which Gaussian functions were augmenting the trigonometric expansion. Work has also continued using rather large multi-term Fourier expansion.

The long term aims for the present dissertation is to find mathematical functions which is describe the folding of the peptide residues. In order to find eventually the solution to solve the protein folding problem and to build up the conformational network for the folding but it is reasonable to start with the description of small compounds, and aim for a bottom up solution.

The aims for the present dissertation:

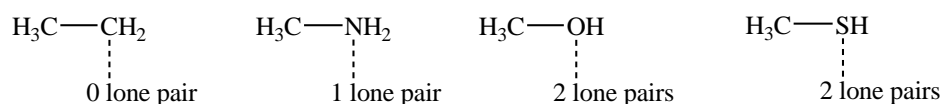
- to achieve a reasonable accuracy of the fitted functions for the folding of the internal rotation of typical organic functional groups which have only one independent variable,
- to extend the one dimensional mathematical functions to describe the folding of the molecules which have two independent torsional angles and they are include peptide bonds,
- the fitted functions, which are describe the conformational spaces, can be analyzed to obtain the critical points, minima, transition states and maxima.

2. Applied methods

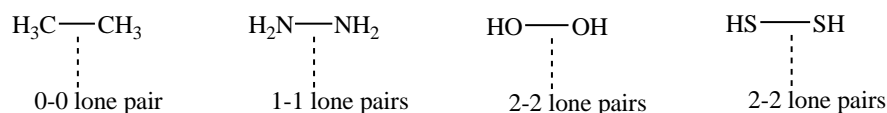
2.1 Chemical background

Organic molecules with a single rotor

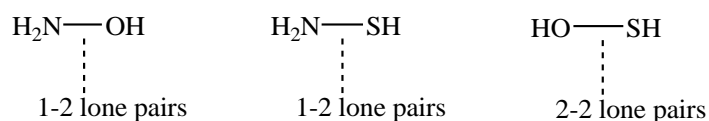
The first question we wish to investigate is how large such a nearest neighbour interaction is in compounds $\text{H}_3\text{C-X}$, where X may be chosen horizontally or vertically along the periodic table:



It will explore is if it is possible to quantify the changes in energy for the rotation about homonuclear diatomic units such as (X-X) as the number of lone pairs of X increases, as in the series:



Heteronuclear (X-Y) diatomic units were also studied.



Organic molecules with double rotors

Of the five compounds the first two (I and II) had saturated rotating moieties attached to the central CH_2 skeleton. The second set (III, IV and V) contains amide bonds (-CONH-). Compounds, III, IV and V have special structures: which is actually glycine diamide -CONH- CH_2 -CONH-, alanine diamide-CONH-CH-Me-CONH- and valine diamide -CONH-CH-CH-Me₂-CONH- .

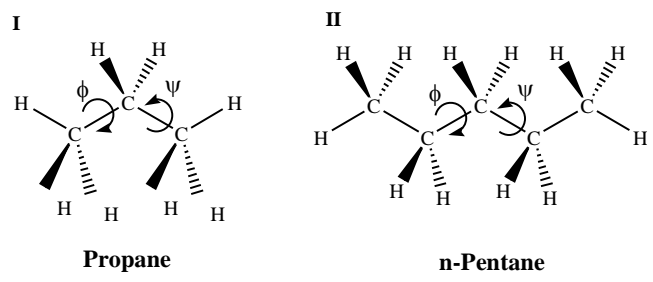


Figure 1 Molecules with rotating sp^3 ligand studied as double rotors to generate the Potential Energy Surface (PES).

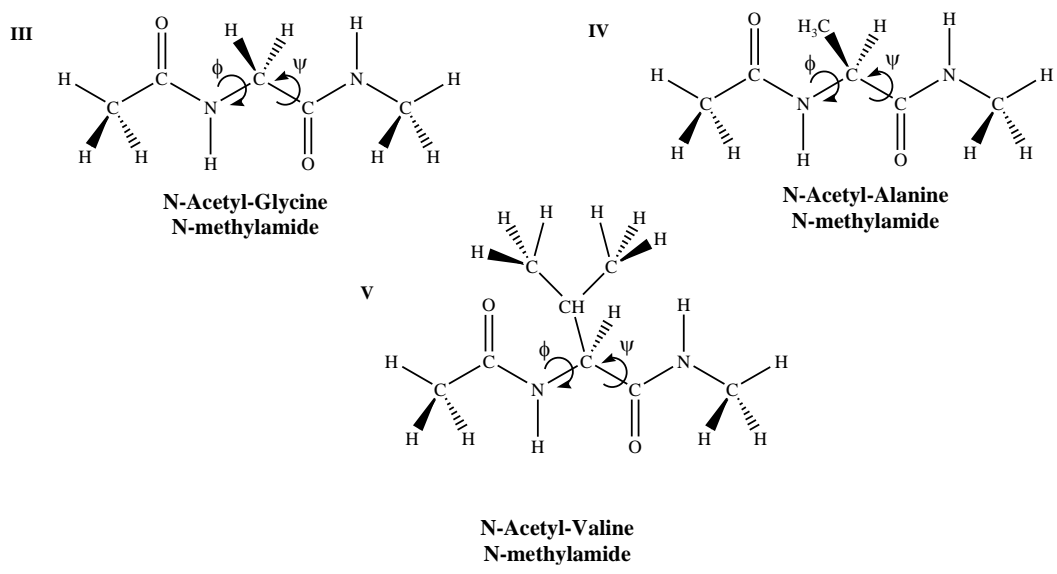


Figure 2 Molecules with rotating symmetric sp^2 ligand studied as double rotors to generate the Potential Energy Surface (PES).

Clearly, in the first two compounds (I and II) the rotating moieties connected to the central methylene carbon are of tetrahedral moieties. In contrast to that the last three compounds (III-V) have flat, trigonal planar, rotating moieties with heteroatoms.

These set of five compounds (I-V) represent increasing complexity of the potential energy surface. Of course, the increasing complexity in appearance may lead to increasing complexity of the mathematical function to be fitted.

2.2 Functional analysis

The relative energies, ΔE , associated with internal rotation were calculated by quantum mechanics using the B3LYP/6-31G(d) implementation of the density functional theory in gas phase in the Gaussian09 software package. The calculations were carried out for a range of dihedral angles ϕ and ψ of the molecules in the interval $[-\pi, \pi]$ with grid points at 15° increments in order to generate the curve, surface and hypersurface.

The functional representation of the internal rotations are based on the finite Fourier series [1] which are a periodic functions in $[-\pi, \pi]$ interval and it able to represent the periodicity of the PEC or PES. Fourier-series with the general form is shown in equation [1].

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \quad [1]$$

where $f(x)$ is a periodic function in $[2\pi]$ interval and a_0 is a constant term of the series .

The functional analyses were carried out using the Matlab software. A Levenberg-Marquardt algorithm, a nonlinear least square method with a local minimizer, was performed to fit a Fourier expansion. A nonlinear least square method is defined as an equation that is nonlinear, or a combination of linear and nonlinear in the coefficients. The model particularly sensitive to the starting points.

Potential Energy Curve fitting

Let ΔE be the energy change associated with rotation of a functional group relative to the lowest energy. ΔE can be expanded in a Fourier-series [2]:

$$\Delta E(\varphi) = a_0 + \sum_{n=1}^m (a_n \cos n\omega\varphi) \quad [2]$$

where a_0 is a constant, m is the number of terms in the expansion and ω is the fundamental frequency of the signal.

Potential Energy Surface fitting

In the case of surfaces, each surface had a minimal set of functions [3], which includes a summation up to 6, yielding $6 \times 4 = 24$ terms of simplified Fourier expansions⁶⁸ with two independent variables.

$$E_a(\varphi, \psi) = a_0 + \sum_{m=1}^6 (a_1 \cos m\omega\varphi + b_1 \cos m\omega\psi + a_2 \sin m\omega\varphi + b_2 \sin m\omega\psi) \quad [3]$$

where a_0 is the constant term in the series, ω is the fundamental frequency of the signal and m is the number of the terms.

For the sake of achieving higher accuracy, a set of Gaussian functions were fitted [4] to the recognisable critical points.

$$E_b(\varphi, \psi) = \sum_{m=1}^9 A_m e^{\left(- \left(\frac{c_\varphi (\varphi - \varphi_{0m})^2}{2\sigma_{\varphi m}^2} + \frac{c_\psi (\psi - \psi_{0m})^2}{2\sigma_{\psi m}^2} \right) \right)} \quad [4]$$

where m is the number of the terms, A is the amplitude, φ_0 and ψ_0 are define the center and the σ_φ and σ_ψ are the φ and ψ extension of the ellipsoid.

And for the sake of the transformation of coordinate system an extended two dimensional Fourier-series [5] is needed.

$$E_c(\varphi, \psi) = \sum_{m=1}^6 f_1 \cos(m\omega\varphi + m\omega\psi) + f_2 \cos(m\omega\varphi - m\omega\psi) + f_3 \cos(m\omega\varphi + m\omega\psi) + f_4 \sin(m\omega\varphi - m\omega\psi) + f_5 \sin(m\omega\varphi + m\omega\psi) + f_6 \cos(m\omega\varphi - m\omega\psi) + f_7 \sin(m\omega\varphi + m\omega\psi) + f_8 \sin(m\omega\varphi - m\omega\psi) \quad [5]$$

In the summation m is the number of terms in the expansion and ω is the number of the terms.

Determination of critical points

First derivative and the second derivatives for the classification of critical points are also determined by the methods of mathematical analysis.

3. Achieved results

T 1. The PEC of an organic molecules with a single rotor can be fitted with an even Fourier-series

T 1.1. Depending on the structural complexity of a given molecule, a 1-term or 2-term or 3-term Fourier expansion turned out to be adequate, using only cos expansion [2].

T 2. Any conformational PES can be described with an accurate multivariable mathematical function.

T 2.1. Molecules, which have two symmetric rotational groups (like propane) and their surfaces have symmetry axes, can be described with a linear combination of the single Fourier series [3].

T 2.2. In the case, at a larger molecule such as n-pentane, the surface becomes less-symmetric and the fitted functions are increasingly more complicated, thus, Gaussian functions [4] are also needed to fit the surface.

T 2.3. If the surface has only central symmetry (e.g. glycine diamide) or has not symmetry (e.g. alanine diamide or valine diamide) at all the fitting needs a dissymmetrical correction function, in the form of an extended Fourier series [5].

T 3. The critical points can be determined by functional analyses.

T 3.1. The minima and the transition states has been defined and that interconnects minima and TS are regarded as the network of the conformational subspace.

T 4. The present dissertation suggests that the potential energy surfaces and hypersurfaces of flexible molecules, such as very small peptides, containing several internal bond-rotations, may be reasonably well represented by these types of fitting method.

T 4.1. However, the amount of computation needed is increasing exponentially.

T 4.2. In the case of a pentapeptide (e.g. opiorphin) the computational requirement is beyond what is nowadays called “big data”.

4. Practical use of the results

Present dissertation is a first step to make a new, more accurate method which can be describe the full conformational subspace of a large molecule. Furthermore, the model can be useable to develop a more accurate force field.

The Dissertation is based on the following Papers

Anita Rágyanszki, Attila Surányi, Imre G. Csizmadia, András Kelemen, Svend J. Knak Jensen, Selma Yarligan Uysal, Béla Viskolcz. Fourier type potential energy function for conformational change of selected organic functional groups. *Chem. Phys. Lett.* **599 (2014) 169–174**
Impact Factor: 1.99

Anita Rágyanszki, Klára Z. Gerlei, Attila Surányi, Imre G. Csizmadia, András Kelemen, Svend J. Knak Jensen, Béla Viskolcz. Big Data reduction by fitting mathematical functions A search for appropriate functions to fit Ramachandran surfaces. *Chem. Phys. Lett.* **625 (2015) 91-97**
Impact Factor: 1.99

János J. Szórád, Eszter P. Faragó, **Anita Rágyanszki**, Franco A. Cimino Prado, Béla Fiser, Michael C. Owen, Balázs Jójárt, Claudio A. Morgado, Milán Szőri, Svend J. Knak Jensen, Imre G. Csizmadia, Béla Viskolcz. Conformation change of opiorphin derivatives. A theoretical study of the radical initiated epimerization of opiorphin. *Chem. Phys. Lett.* **626 (2015) 29–38**
Impact Factor: 1.99

Published Papers: 3

Cumulative Impact Factor: 5.97

Conference Presentations

Oral presentations

Anita Rágyanszki, Attila Surányi, Imre G. Csizmadia, András Kelemen, Svend J. Knak Jensen, Béla Viskolcz. Analytical Functional Representation of Quantum Chemical energy curves and Surfaces. I. Innovation in Science Doctoral Student Conference. (Szeged, Hungary) May 2-3, 2014

Anita Rágyanszki, Attila Surányi, Imre G. Csizmadia, András Kelemen, Svend J. Knak Jensen, Béla Viskolcz. Functional Analyses of 2D Conformational Potential Energy Surfaces. KeMoMo-QSAR. (Szeged, Hungary) May 22-23, 2014

Anita Rágyanszki, Klára Z. Gerlei, Attila Surányi, András Kelemen, Svend J. Knak Jensen, Imre G. Csizmadia, Béla Viskolcz. Functional representation of Multidimensional Conformational Potential Energy Surfaces. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) June 17-20, 2014

Klára Z. Gerlei, **Anita Rágyanszki**, Tímea Tóth, Imre G. Csizmadia, Béla Viskolcz. 4D Analytical PEHS of L-Asparagine Diamide. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) June 17-20 2014

Scientific Posters

Eszter Pongráczné Faragó, **Anita Rágyanszki**, Franco A. Cimino P., János J. Szórád, Béla Fiser, Michael C. Owen, Balázs Jójárt, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. Radical Initiated Epimerization and Conformational Change of Opiorphin: A Theoretical Study. 3rd Visegrad Symposium on Structural Systems Biology (Smolenice Slovakia) 19– 22 June, 2013

János J. Szórád, Eszter P. Faragó, **Anita Rágyanszki**, Franco A. Cimino Prado, Béla Fiser, Michael C. Owen, Balázs Jójárt, Claudio A. Morgado, Milán Szőri, Svend J. Knak Jensen, Imre G. Csizmadia, Béla Viskolcz. A Theoretical Study of the Epimerization of Opiorphin Pentapeptide in the Presence of OH[•]/H₂O₂ Redox System. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014

Joe Bedard, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. A Computational ab initio Study on the Thermodynamic Stability of the 45 Constitutional C₂H₄O₃ Isomers. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014

Keyue Chen, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. A Thermodynamic Study of Glycerol and Its 191 Isomers Using Computational Ab Initio Methods. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Zsolt Filákovics, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. A Theoretical Analysis on the Thermodynamic Stability of Various Constitutional Isomers of C₄H₆O₂. A case study of tetrahydrofuran derivatives. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Kyungseop Lee, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. Theoretical stability assessment of isomers of C₃H₄O₂ with thermodynamic properties. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Li Jian Bin, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. An ab initio study on the thermodynamic properties of $C_3H_6O_3$ isomers and its potential in synthetic and metabolic processes. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Ying Lu, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. A Theoretical Analysis on the Thermodynamic Stability of Various Constitutional Isomers of $C_3H_2O_3$. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Hyemin Park, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. Ab initio Study on the Thermodynamic Stabilities of Isomers for $C_2H_6O_3$. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Chaitnandanie Ramkissoo, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. Screening Constitutional Isomers of $C_3H_2O_2$ for thermodynamic properties obtained from Systematic Ab initio calculations. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Zsófia Borbála Rózsa, Milán Szőri, János J. Szórád, Imre G. Csizmadia, Béla Viskolcz, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, Michael C. Owen, Béla Fiser. A theoretical analysis on the thermodynamic stability of various constitutional isomers of $C_4H_8O_2$. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Dhiraj Sinha, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Rudiger Ettrich, Béla Viskolcz. An ab initio Study of Thermochemical Properties of $C_4H_4O_2$ Isomers. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Daniel Y. Wang, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. Systematic Ab initio Study of all Constitutional Isomers of $C_3H_8O_2$ Including Molecular Complexes. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

XuanTing Xu, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. Ab initio study of the thermodynamic stability of $C_3H_6O_2$ isomers. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.

Peiqi Yu, Gabriella Kaszap, Klára Z. Gerlei, **Anita Rágyanszki**, János J. Szórád, Béla Fiser, Michael C. Owen, Milán Szőri, Imre G. Csizmadia, Béla Viskolcz. Ab initio study of the thermodynamic stability of $C_3H_4O_3$ isomers. 4th Visegrad Symposium on Structural Systems Biology (Nové Hradý Czech Republic) 17– 20 June, 2014.